UNIVERSITY OF SOUTHAMPTON

FACULTY OF MEDICINE

Cancer Sciences Unit

Cellular Pathology and Molecular Diagnostics for Cancer

by

Dr Emily Clare Shaw

BM(Hons) MRCP(UK) FRCPath

Thesis for the degree of Doctor of Medicine

September 2017

ABSTRACT

UNIVERSITY OF SOUTHAMPTON: Faculty of Medicine, Cancer Sciences

Thesis for the degree of Doctor of Medicine: **CELLULAR PATHOLOGY AND MOLECULAR DIAGNOSTICS FOR CANCER,** Dr Emily Clare Shaw

Recent developments in knowledge of cancer at the molecular level have led to a growing demand for tissue-based predictive analysis to inform therapeutic decision-making. Molecular parameters are also being increasingly incorporated into traditionally morphology-based diagnostic and prognostic classification systems. The resulting broader application of molecular techniques for interrogation of tissue samples requires adaptation of cellular pathology methods. A number of large-scale initiatives underway worldwide, including the Cancer Research UK Stratified Medicine Programme, are attempting to establish the evidence base and develop the teams, processes and infrastructure necessary to deliver this approach in routine practice.

This thesis describes the findings of work in this area including collaborative efforts through the Stratified Medicine Programme and STRATFix consortium in the areas of patient consent, data, technology, tissue fixation and processing, utility of alternative tissue fixatives and pathologist or digital tumour content assessment. This work has demonstrated that acquisition of tissue surplus to diagnostic requirements for DNA-based tests is acceptable to patients, that targeted mainly 'hotspot' sequencing of up to five clinically relevant genes is feasible in a single tissue sample and that clinical data systems in their current form require a large amount of manual intervention to produce cancer data in a format compatible with the current NHS information standard. Furthermore, this research has demonstrated the variation in different aspects of tissue sample handling despite an increasing number of laboratories receiving accreditation to ISO standards, with its central focus on uniformity of process. There is also description of variation in tumour content assessment by a group of experienced pathologists using online whole slide imaging, indicating that accurate tumour quantification in samples submitted for sequencing is likely to require digital image analysis. This work shows that as a 'molecular friendly' fixative, the PAXgene® Tissue system provides tissue preservation generally suitable for morphological assessment and diagnosis, with the exception of lymphoid tissue for which further optimisation work is in progress. Histochemical techniques in use in our laboratory appear to be directly transferable to PAXgene® Tissue-fixed paraffin embedded tissue but immunohistochemistry requires protocol modification, particularly for antigens located in the cell nucleus. Double-stranded DNA yields from PAXgene® Tissue-fixed, paraffin embedded tissue are at least comparable to those obtained from matched formalin-fixed paraffin-embedded tissue and show better preservation with less DNA fragmentation.

This body of work has enabled me to develop knowledge, skills and evidence to contribute to the crucial role of cellular pathology in the implementation of stratified cancer medicine for improved patient care.

Table of Contents

List of tables

List of figures

DECLARATION OF AUTHORSHIP

I, Emily Clare Shaw, declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

The Cancer Research UK Stratified Medicine Programme (CRUK-SMP) and STRATFix consortium represent large, multi-institution collaborative research projects I have been involved in and my individual contribution to these projects, forming a major part of the work of this thesis, is as follows:

- I commenced full-time employment with Cancer Research UK as Clinical Lead for the Stratified Medicine Programme in August 2011, in the early stages of phase one, with a remit for clinical input to the running and monitoring of the programme. Through this I performed a series of site visits, advising on a range of clinical, scientific and operational issues, advised on appropriate key performance indicators, prepared guidance documents on patient eligibility, clinical dataset issues, agreed reporting nomenclature and use of sequencing results as well as coordinating and attending working groups and expert advisory groups. I was responsible for liaising with pharmaceutical and biotechnology company representatives as well as other partner organisations, preparing and presenting data and reports on programme activities for CRUK boards established as part of the programme governance structure and also submitting and presenting data at national and international conferences (see references below)

- After completing a year of full-time employment with CRUK, in August 2012 I changed to working one day per week for CRUK on the programme and combined this with a return to histopathology specialty training. I continued to advise on clinical issues, participate in working groups and advisory groups and write up the findings of the programme. This part-time employment continued until I gave notice to take up my NHS consultant post in June 2016.

- The STRATFix consortium was formed in 2013 in response to invitations to apply for collaborative R&D funding from the Technology Strategy Board (now Innovate UK) for 'Improving cell and tissue analysis for

stratified medicine'. I coordinated a group of collaborators I knew from CRUK SMP to prepare a bid with QIAGEN which was successfully funded.

- Planning of the structure and model for CRUK-SMP was already in place when I joined CRUK, with a finalised protocol, research ethics committee approval granted and the laboratory and clinical sites selected. The gene list for each tumour type had also just been agreed through a series of teleconferences with representatives from the National Cancer Research Institute Clinical Specialty Groups. I prepared detailed patient eligibility criteria and advised on the content of key performance indicators for clinical and laboratory sites (Appendix A).

- The SMP1 mutation data analysis presented in Chapter 3 is based on a single data extract I requested from colleagues working on incorporating the data into a database in Oxford. I received the requested data items in a single Excel spreadsheet, and then performed manual curation before summarising the data into the format presented here. A lot of the remainder of the content of this chapter represents insights I gained on visits to sites and represent my opinions as well as those of others I worked with on the programme, as related to me and attributed in the acknowledgements section of this thesis.

- I came up with the concept for, designed, sent out data requests and analysed and presented results of the two cross-sectional analyses of cellular pathology department sample handling and endobronchial ultrasound-guided lung samples presented in Chapter 4 (Appendix B and Appendix C).

- I acquired from surgical resections received in the Southampton cellular pathology department the tumour samples used to prepare paired PAXgene® Tissue-fixed and formalin-fixed blocks for the morphology, histochemical, immunohistochemical and molecular work presented in Chapter 5. I also designed the sample handling dataset used by all laboratory sites for the project (Appendix D). I was one of the two assessors for the morphology, histochemistry and immunohistochemistry scoring.

I identified the series of H&E specimens and determined the specification of the module created within the POSHviewer system and used for online pathologist tumour content assessment for the work presented in Chapter 5. I identified and invited the participating pathologists and retrieved, collated and analysed their responses from the system.

Cellular Pathology and Molecular Diagnostics for Cancer

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;

2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;

3. Where I have consulted the published work of others, this is always clearly attributed;

4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;

5. I have acknowledged all main sources of help;

6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;

7. Parts of this work have been published as:

Lindsay CR, Shaw EC, Blackhall F, Blyth KG, Brenton J, Chaturvedi A, Clarke N, Dick C, Evans TRJ, Hall G, Hanby AM, Harrison DJ, Johnston S, Mason MD, Morton D, Newton-Bishop J, Nicholson AG, Oien KA, Popat S, Rassl D, Sharpe R, Taniere P, Walker I, Wallace WA, West NP, Johnson PWM. Somatic Cancer Genetics in the UK: Results from Phase One of the Cancer Research UK Stratified Medicine Programme. Manuscript submitted to *British Journal of Cancer*.

Shaw EC, Hair J, Rogan J, Steele J, Speirs V, Banks RE, Beirne D, Hanby AM, Tuff-Lacey A, Beaton S, Croudass A, Jones L, Lindsay CR, Johnson PWM. Facilitating Tissue Donation for Research: Innovative consent models used in Phase One of the Cancer Research UK Stratified Medicine Programme. Manuscript submitted to *British Journal of Cancer*.

Shaw EC, Cummings R, Lindsay CR, Jones MCM, Butler S, Rehal P, MacDonald F, Griffiths M, Bell J, Smith M, O'Sullivan B, Taniere P, Gonzalez de Castro D, Lillis S, Thompson L, Mair D, Butler R, Stuart H, Ellwood-Thompson R, Brown C, Eden J, Deans ZC and Johnson PWM. The Cancer Research UK Stratified Medicine Programme: Phase One – Methods and insights for delivering routine molecular diagnostics for cancer patients. Manuscript submitted to *British Journal of Cancer*.

Shaw EC, Lees R, Reilly S-J, et al on behalf of the STRATFix consortium. Morphology assessment and suitability for diagnosis of formalin-fixed versus PAXgene® Tissue-fixed and paraffin embedded tumour tissue. Published abstract in *J Pathol* **240** suppl S1, October 2016.

Lindsay CR, Shaw EC and Johnson PWM. Stratified Medicine in the UK. Chapter in the Association of Cancer Physicians text 'Problem Solving Through Precision Oncology: A Case Study Based Reference and Learning Resource'. Oxfordshire, EBN Health, 2016.

Tuff-Lacey A, Shaw EC, Cummings R et al. A collaborative approach to enabling stratified cancer medicine in the United Kingdom. *Drug Disc Today*. 2015 http://dx.doi.org/10.1016/j.drudis.2015.07.001.

Lindsay CR, Shaw EC, Walker I and Johnson PWM. Lessons for molecular diagnostics in oncology from the Cancer Research UK Stratified Medicine Programme. *Expert Rev Mol Diagn.* 2014 Dec **30**:1-3. [Epub ahead of print]

Shaw EC, Ashton-Key M and Johnson PWM. Stratified medicine for cancer: the role of the histopathologist. Book chapter in Recent Advances in Histopathology no.23, 2014.

Shaw EC and Johnson PWM. Molecular Diagnosis of Cancer: Making it a reality. Extended contribution in Translational Genomics. Edited by Li Wan Po A and Farndon P. NHS National Genetics Education and Development Centre, Birmingham 2013

Shaw EC and Johnson PWM. Stratified medicine for cancer therapy. *Drug Discov Today* 2012 Mar; 17(5-6): 261-8.

Abstracts and poster presentations

Pathological Society Summer Meeting 2017: Belfast Pathology, Belfast, to be held June 2017. Lees R, Shaw EC, Reilly SJ, Bodey K, Machado M, Ashton-Key M. The comparison of DNA quality and integrity between PAXgene® and formalin fixed paraffin embedded tissues (abstract submitted March 2017).

The Farr Institute International Conference 2015: Data Intensive Health Research and Care, University of St Andrews, Scotland, August 2015. Shaw EC, Tuff-Lacey A, et al. Utilising stratified medicine data for research: a national approach to collection and integration of molecular and NHS clinical data.

10th National Cancer Research Institute conference, November 2014. Shaw EC, Butler R, Gonzalez de Castro D et al. A national platform for molecular diagnostics: Results of the Cancer Research UK Stratified Medicine Programme.

26th EORTC-NCI-AACR Symposium on Molecular Targets and Cancer Therapeutics, Barcelona, Spain, November 2014. Shaw EC, Lindsay C, Walker I and Johnson PWM on behalf of the Stratified Medicine Programme 1 Consortium. The Cancer Research UK Stratified Medicine Programme: From national screening to national trial.

26th European Congress of Pathology, London, September 2014. Shaw EC, Smith M, Thompson L et al. Progress in Phase Two of the Cancer Research UK Stratified Medicine Programme: A National Molecular Profiling Initiative for Lung Cancer and Shaw EC, Cane P, Taniere P et al. Handling of lung cancer endobronchial ultrasound samples at sites involved in the Cancer Research UK Stratified Medicine Programme.

American Society of Clinical Oncology 2014 Annual Meeting, Chicago, May 2014. Shaw EC, Butler R, Gonzalez de Castro D et al. A national platform for molecular diagnostics: Results of the Cancer Research UK Stratified Medicine Programme.

7th Joint Meeting of the British Division of the International Academy of Pathology and the Pathological Society of Great Britain & Ireland, Edinburgh, June 2013. Shaw EC, Collins VP, Gonzalez de Castro D et al. Variation in Cellular Pathology Department Specimen Handling Processes in the Cancer Research UK Stratified Medicine Programme.

Signed: Emily Shaw

Date: 23/09/2017

Acknowledgements

First and foremost I would like to thank my supervisors Professor Peter Johnson and Dr Meg Ashton-Key for their patience, wise counsel and encouragement without which I would not have got as far as this point.

The funders of the work contributing to this thesis are Cancer Research UK, AstraZeneca, Pfizer, Roche, Bristol-Myers Squibb and Oracle Health Sciences through the Stratified Medicine Programme and Innovate UK (formerly the Technology Strategy Board) and Qiagen for the STRATFix consortium.

Patient consent, sample preparation and clinical data collation for the Stratified Medicine Programme were performed by colleagues in teams in Birmingham (Principal Investigator Professor Dion Morton), Cambridge (Principal Investigator Professor Peter Collins), Cardiff (Principal Investigator Professor Malcolm Mason), Edinburgh (Principal Investigator Professor David Harrison), Glasgow (Principal Investigator Dr Karen Oien), Leeds (Principal Investigator Professor Andrew Hanby), London (Principal Investigator Professor Stephen Johnston) and Manchester (Principal Investigator Professor Noel Clarke). Sample molecular analysis, interpretation and reporting for the Stratified Medicine Programme were performed by teams at the molecular genetics laboratories in Birmingham (Birmingham United Molecular Pathology, Professor Mike Griffiths and Dr Phillippe Taniere), Cardiff (All Wales Molecular Genetics Laboratory, Dr Rachel Butler) and Sutton, Surrey (Institute of Cancer Research/ Centre for Molecular Pathology, Dr David Gonzalez de Castro).

Operational management and strategic development of the Stratified Medicine Programme was carried out by the fantastic team past and present at Cancer Research UK headquarters, London, including James Peach, Professor Peter Johnson, Dr Louise Jones, Dr Fiona Hemsley, Dr Rowena Sharpe, Dr Ilaria Mirabile, Helen Pitman, Jon Messer, Anne Peel, Anne Croudass, Monica Jones, Lorna Rose, Sarat Telaprolu, Dr Catrin Middleton, Tosin Sule, Steven Johnson, Dr Ian Walker, Dr Colin Lindsay, Alice Tuff-Lacey, Will Francis, Nicola Soper, Dr Sumi Subramaniam, Rachael Shaw, Tracey Powell, Flavie Belanco, Esau Moreno-Camacho, Jeremy Brown and Rupesh Vyas-Robinson. Dr Emily Smith, Tarita Turtiainen, Sarah Beaton and Dr Rebecca Cummings were also part of this team

and had responsibility for receiving, collating the patient accrual and laboratory data presented in chapter 3. We also worked closely with Dr Sandi Deans of the UK National External Quality Assessment Scheme (NEQAS) for Molecular Genetics in developing the bespoke EQA scheme for SMP1, some results of which are presented in chapter 3. I am also indebted to, and learned so much, from the various advisory and working groups and boards I interacted with through my involvement in the Stratified Medicine Programme. It was a unique personal and career development opportunity and a privilege to work on this exciting initiative which I think has greatly influenced the future direction of my career.

Collation of the CRUK SMP1 clinical dataset was overseen by colleagues from the National Cancer Registration and Analysis Service (formerly Eastern Cancer Registry Information Centre), Cambridge including Dr Jem Rashbass, Mr Andy Murphy and Ms Jess Farrimond. I am also grateful for assistance with transfer of this data to a research database and production of the data extract analysed in Chapter 3 to Professor Jim Davies and Dr Steve Harris of Oxford University of Computing Sciences. Kevin Wheeler in the University of Southampton Clinical Informatics Research Unit developed the online slide viewing system described in chapter 2 and used for the tumour content assessment exercise in Chapter 5. I would also like to acknowledge the generous collaboration of Dr Nick West at Leeds for performing the digital tumour content assessment on the SMP2 lung cancer cases, the results of which are presented and discussed in chapter 5.

An enormous amount of work in sample preparation for the STRATFix project has been carried out by Maria Machado, Amicha Afonso, Dr Kathleen Bodey, Charlie Greene, Robert Lees, Sara-Jane Reilly, Catherine Griffiths, Hamid Alagra and Mary Gudime in the department of Cellular Pathology at University Hospital Southampton NHS Foundation Trust. Sara-Jane Reilly was responsible for tissue processing, embedding and preparing slide-mounted including H&E-stained sections. I am indebted to Rob Lees for carrying out the histochemistry, immunohistochemistry, DNA quantification and quality assessment on the formalin- and PAXgene® Tissue fixed pairs of blocks reported in chapter 5, with the assistance of Sue Harris, Dr Liz Hodges, Dr Steve Doherty, Nicola Meakin and Dr Ros Ganderton in the molecular pathology department at University Hospital Southampton NHS Foundation Trust.

Clinical trials assistants Ben Johnson, Angela Blackmore, Michael Dahl, Lucy Hogan, Carine Fixmer, Roy Raphael and Jane Dare deserve a mention here for their essential role in transporting fresh specimens from the operating theatre to the cellular pathology department, allowing simultaneous sampling for the STRATFix and 100,000 genomes projects.

Oversight and development of the STRATFix consortium has been led by the industrial partner QIAGEN with invaluable input from Dr Daniel Grölz, Dr Uwe Oelmueller, Tomasz Krenz, Natasha Cant and project manager Safiya Reza and Innovate UK monitoring officer Suresh Chinwala keeping us all on track. I am grateful to Meg Ashton-Key and colleagues from the other participating institutions in Papworth (Dr Doris Rassl and Mark Southwood), Glasgow (Jane Hair and Dr Karen Oien), Birmingham (Dr Phillippe Taniere and Brendan O'Sullivan), Belfast (Professor Manuel Salto-Tellez and Dr Perry Maxwell) and London (Dr Bridget Wilkins, Dr Lilian Edmunds, Dr Gareth Gerrard and Dr Tim Forshew).

Statistical input for my thesis has been ably and patiently provided by Scott Harris of the University of Southampton Research and Development Support Unit.

My interest in thoracic pathology and skills in fresh tissue sampling and research were developed during a brief but valuable six week clinical fellow attachment in the histopathology department at the Royal Brompton Hospital, London in November – December 2015, and for this and their mentoring and encouragement I thank Professor Andrew Nicholson, Dr Alex Rice and Dr Angeles Montero.

The online tumour content assessment exercise written up in chapter 5 was administered through the CRUK SMP2 pathology working group and I would like to extend my thanks to the chair Professor Jo Martin and other members of this group: Dr Anca Oniscu, Dr William Wallace, Dr Karin Oien, Dr Craig Dick, Dr Fiona Black, Dr Anshuman Chaturvedi, Dr Helen Doran, Dr Jon Bury, Dr Nick West, Dr Eldo Verghese, Dr Phillippe Taniere, Dr John Le Quesne, Dr Jason Shannon, Dr Doris Rassl, Dr Sanjay Jogai, Professor Andrew Nicholson, Dr Paul Cane, Dr Elaine Borg, Dr Francesco Mauri and Dr Michael Sheaff.

Finally I would like to thank my unfailingly supportive, understanding and patient family, particularly my husband Kevin, daughter Genevieve, mother Penny and father Keith for making all this worthwhile.

Definitions and Abbreviations

aCGH	array comparative genomic hybridisation
ADC	adenocarcinoma
ADP	adenosine diphosphate
ARMS	amplification refractory mutation system
ATP	adenosine triphosphate
bp	base pair (length of DNA molecule)
BRAF	v-raf murine sarcoma viral oncogene homolog B1
ccfDNA	circulating cell-free tumour-derived DNA
CH	clinical hub in the Stratified Medicine Programme
COSD	Cancer Outcomes and Services Dataset for England
CRUK	Cancer Research UK
CRUK-SMP	Cancer Research UK Stratified Medicine Programme
Ct	cycle threshold for polymerase chain reaction
DIN	deoxyribonucleic acid (DNA) integrity number
DNA	deoxyribonucleic acid
DPAS	diastase periodic acid-Schiff
dsDNA	double-stranded DNA
EBUS-FNA/TBNA	endobronchial ultrasound-guided fine needle aspiration/ transbronchial needle aspiration
EGFR	epidermal growth factor receptor
ER	oestrogen receptor
ERG	ETS-related gene

ETS	erythroblastosis-virus transformation-specific
EUS	endoscopic ultrasound
EVG	elastic van Gieson
FDA	Food and Drug Administration (United States)
FF	fresh frozen (tissue)
FFPE	formalin fixed paraffin embedded
FISH	fluorescent in situ hybridisation
H&E	haematoxylin and eosin
HER2	human epidermal growth factor receptor 2
HGVS	Human Genome Variation Society
HIER	heat-induced epitope retrieval
HRM	high resolution melt
IARC	International Agency for Research against Cancer
ICC	intraclass correlation coefficient
ICD-10	10th revision of the World Health Organization International Statistical Classification of Diseases and Related Health Problems
INCa	*Institut National du Cancer* initiative in France
INN	World Health Organization's International Non-proprietary Name system for generic drug nomenclature
IPA	isopropyl alcohol
ISO	the International Organization for Standardization
KRAS	kirsten rat sarcoma 2 viral oncogene homolog
LBC	liquid-based cytology

LMP	low melting point
MAPK	mitogen activated protein kinase
MLPA	multiplex ligation-dependent probe amplification
NBF	neutral buffered formalin
NEQAS	UK National External Quality Assessment Services
NGS	next generation sequencing
NICE	the National Institute of Health and Care Excellence
NRAS	neuroblastoma rat sarcoma virus
NSCLC	non-small cell lung carcinoma
PARP	poly-adenosine diphosphate ribose polymerase
PAS	periodic acid-Schiff
PCR	polymerase chain reaction
PFPE	PAXgene® Tissue fixed paraffin embedded
PIS	patient information sheet
PR	progesterone receptor
REC	Research Ethics Committee
RIN	RNA integrity number
RTB	research tissue bank
SCC	squamous cell carcinoma
SMP	Stratified Medicine Programme
SMP1	Stratified Medicine Programme Phase 1
SMP2	Stratified Medicine Programme Phase 2

SNOMED-RT	College of American Pathologists' Systematized Nomenclature of Medicine Clinical Terms Reference Terminology
SOP	Standard Operating Procedure
SPIDIA	Standardisation and Improvement of Generic Pre-analytical Tools and Procedures for In Vitro Diagnostics
SSCP/SSCA	single strand conformation polymorphism/ analysis
ssDNA	single stranded DNA
TCD	tumour content determination
TCGA	the Cancer Genome Atlas
TH	technology hub in the Stratified Medicine Programme
TKIs	tyrosine kinase inhibitor drugs
TNM	Tumour Nodes Metastasis classification system
UK	United Kingdom
UKAS	United Kingdom Accreditation Service
US	United States of America

Chapter 1: Introduction and literature review

1.1 Introduction

There are an increasing number of novel therapeutic agents, both licensed and in clinical trials, requiring deoxyribonucleic acid (DNA)-based analyses on tumour tissue to identify patients who may benefit from treatment due to the presence of specific molecular aberrations that may be detected in the tumour material. Little is known currently about the impact of sample handling processes in departments of cellular pathology on DNA quality and therefore the likelihood of success of subsequent analysis, and there is an emerging literature in this area.

As a multi-centre pilot of the implementation of molecular profiling into cancer diagnostics in the United Kingdom, the Cancer Research UK Stratified Medicine Programme provides a unique opportunity to study the mutational profiles across common cancer types in a broad sample of the UK population, and to investigate the impact of factors such as tissue fixation and processing on DNA quality in cancer specimens, in order to contribute to the evolving evidence base and establish standards for best practice.

1.2 Background

1.2.1 The diagnostic process in cellular pathology

For over a century, morphological analysis has been the mainstay of cellular pathology, with disease classification and prediction of biological behaviour based on assessment of microscopic appearances in formalin-fixed, paraffin embedded material (FFPE) by skilled histopathologists. The tissue sample handling pathway in cellular pathology is depicted in figure 1. As well as tissue samples for histological analysis, cellular pathology laboratories also receive cytology specimens comprising exfoliated or aspirated cells in suspension in a

fluid medium. More recently, diagnosis and prognostication have been supplemented by the introduction of immunohistochemistry to demonstrate the presence or absence of particular antigenic proteins in tissue sections, and in situ hybridisation techniques for the detection of foreign (e.g. viral) DNA sequences, gene rearrangements and amplifications.

1. RECEIPT	2. FIXATION	3. DISSECTION	4. PROCESSING	5. EMBEDDING	6. MICROTOMY	7. STAINING
The specimen is received and booked in on the laboratory sample tracking/ reporting computer system.	The tissue undergoes fixation in formalin (a solution of formaldehyde).	The specimen is examined, dissected and samples placed in individual plastic cassettes, each slightly larger than a postage stamp.	The cassettes are placed into an automated tissue processing machine, usually run overnight. The tissue passes through a series of reagents to achieve dehydration, clearing and paraffin wax impregnation.	Each tissue sample is embedded in paraffin wax (a paraffin 'block') by a biomedical scientist or support worker.	4-5µm thick sections are cut and placed onto glass slides.	The sections are stained, coverslipped, labelled and booked out of the laboratory for subsequent examination and clinical reporting by a pathologist.

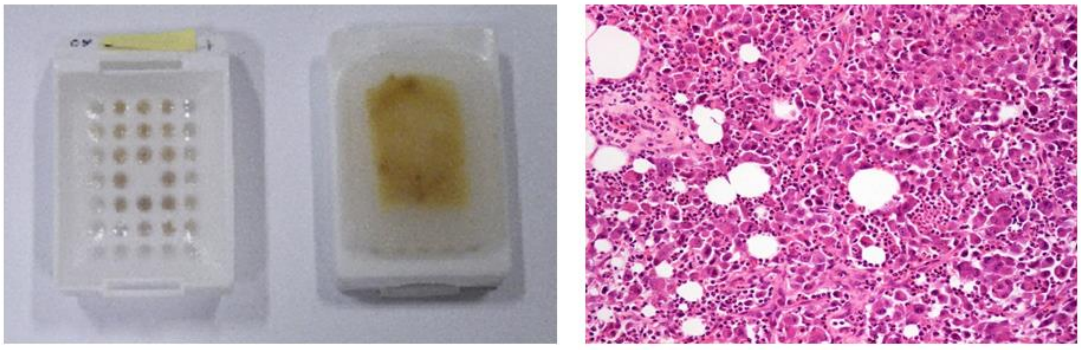

Figure 1. Specimen handling steps in a cellular pathology laboratory

Left lower image: upper and lower views of a cassette and tissue paraffin wax block; right lower image: haematoxylin and eosin-stained section of giant cell carcinoma of lung, 100x overall magnification.

1.2.2 Progress in somatic cancer genomics

Many cancer genomes have now been sequenced and published, including as part of the International Cancer Genome Consortium (http://icgc.org). Cancer genome sequencing projects have generated findings of clinical and therapeutic relevance, for example the identification of the mutated *BRAF* (v-raf murine sarcoma viral oncogene homolog B1) oncogene as a key driver in just over half of malignant melanomas[1]. The timescale from this discovery to the licensing of the BRAF inhibitor vemurafenib was encouragingly short in drug

development terms. Apart from the immunomodulatory therapy ipilumumab, this represents the first effective treatment option for patients with advanced melanoma. One of the challenges in making sense of the deluge of data from genome sequencing studies is in distinguishing key driver mutations from non-pathogenic bystander or passenger mutations, particularly since cancer is characterised by genomic instability, with each cancer containing anything from 1,000 to 100,000 different point mutations[2], some of which are private to that tumour, making assessment of their role in pathogenesis difficult. There is an increasing requirement for the cellular pathologist to be conversant in the language describing the effects of genetic abnormalities identified through cancer genome screening and some of the terminology is summarized in table 1. Mutations are conventionally described using nomenclature agreed by the HGVS (Human Genome Variation Society, www.hgvs.org) with reference to both the coding and protein changes. The following example is for the most common mutation in the *BRAF* gene, a point mutation due to a single nucleotide substitution.

Coding (c.): A description of the abnormality at DNA nucleotide level, according to the numbered nucleotide position on the sense DNA strand (5' to 3' direction) of the reference genome: e.g. c.1799T>A refers to substitution of adenine for thymine at position 1799

Protein (p.): This describes the abnormality at protein coding, amino acid level, according to the number of the affected amino acid. The reference amino acid is denoted by its one or three letter code at the start of the sequence and the mutant amino acid follows the position number at the end of the sequence: e.g. p.V600E or Val600Glu refers to coding for glutamate rather than valine at codon 600.

Recent research into cancer genomic evolution gives insights into the degree of spatial and temporal heterogeneity and complexity[3]. The clonal evolution of tumours over time, particularly in response to selection pressures generated by treatments targeted against specific genetic changes, represents a major challenge to the delivery of personalised cancer medicine through genomics. Relapse of advanced solid malignancies following encouraging initial responses to targeted therapies such as BRAF inhibitors are well recognised and documented in the literature[4 5]. It is becoming more common in clinical practice

for oncologists to request repeat biopsies from metastatic sites in order to perform molecular analysis to guide treatment decisions. The requirement for sequential biopsies is also built into many oncology trial protocols, including the National Lung Matrix Trial, in order to gain understanding of tumour evolution and drivers of disease progression and treatment resistance.
Given that biopsy procedures are not without risk and can be logistically and technically challenging, non-invasive approaches such as obtaining circulating cell-free tumour-derived DNA (ccfDNA, also referred to as 'liquid biopsy') from the plasma fraction of blood have been developed and are gaining acceptance for clinical applications. As well as the use of a ccfDNA-based assay to demonstrate acquisition of *EGFR* resistance mutations such as T790M in lung cancer patients with progressive disease on EGFR inhibitors, a recent proof of principle study has demonstrated the utility of ccfDNA from breast cancer patients for the early detection of relapsed disease[6].

Table 1. Terminology used to classify and describe the effects of gene mutations in cancer

Timing of mutation	*Germline* (constitutional): A mutation present in one parental gamete and transferred to all cells resulting from subsequent cell divisions.	*Somatic* (acquired): A mutation occurring during DNA replication and cell division during life and present only in a subset of cells in a tumour or tissue.
Effect on protein coding	*Synonymous* (silent): A nucleotide change resulting in the same amino acid, due to redundancy in nucleotide combinations coding for each amino acid e.g. coding DNA strands containing GTA and GTG would both encode the amino acid valine.	*Non-synonymous*: A nucleotide change leading to a different amino acid e.g. the DNA sequence GTA codes for valine but a change to GAA would result in glutamic acid instead.
Effect on protein function	*Activating (gain of function)*: A change leading to enhanced effect of a protein e.g. the codon 600 *BRAF* V600E mutation leads to increased activity of the BRAF protein irrespective of usual regulatory mechanisms.	*Inactivating*: A change leading to a non-functional or reduced function protein e.g. codon 594 mutations in the *BRAF* gene cause loss of kinase activity.
Functional effect	*Pathogenic*: A mutation that can be experimentally demonstrated or predicted to contribute to initiation or progression of a tumour.	*Non-pathogenic*: A genetic abnormality that does not appear to contribute to initiation or progression of tumours.
Prediction of treatment response	*Sensitising*: A mutation that has been shown in clinical trials to be associated with treatment response e.g. the L585R mutation in the *EGFR* gene is predictive of response to EGFR tyrosine kinase inhibitor drugs in patients with non-small cell lung cancer.	*Resistance*: A mutation that has been shown in clinical trials to be associated with treatment resistance e.g. the T790M mutation in the *EGFR* gene predicts a lack of response to EGFR tyrosine kinase inhibitor drugs in patients with non-small cell lung cancer. This may be seen as a primary or secondary phenomenon, possibly due to clonal selection pressures in tumours during treatment.
Importance in carcinogenesis	*Driver*: A mutation that is recurrent between different tumours, and can be functionally linked to carcinogenesis e.g. *KRAS* mutations in colorectal cancers cause increased activity of the mutated KRAS protein leading to abnormal cell proliferation.	*Passenger*: A mutation that does not appear to play a role in the initiation or progression of cancer and is likely to have occurred as a bystander effect due to genomic instability. These may be 'private' to particular tumours and are typically not recurrent.

1.2.3 Stratified cancer medicine

Stratified cancer medicine involves predictive analysis: the characterization of tumours according to the presence or absence of specific molecular abnormalities that are associated with differential treatment responses, in order to offer appropriately targeted therapy and avoid exposing patients to treatments with a low likelihood of benefit. This is being applied to an increasing number of solid tumour types to complement the traditional morphological organ or tissue of origin-based assessment of tumours. Thanks to recent advances in genomic technology that have opened up new possibilities for molecular taxonomy in cancer, this is a rapidly evolving area which is having a direct impact on histopathology practice. This represents an important part of the wider concept of delivering more personalised or precision medicine across many different disease areas in the current post-genomic era, since the elucidation and publication of the first human genome over a decade ago.

Histopathologists are accustomed to the use of molecular markers for diagnostic purposes, for example characteristic chromosomal translocations in soft tissue tumours and haematolymphoid malignancies. An exemplar for the application of newer predictive molecular markers was the introduction of HER2 (human epidermal growth factor receptor 2) testing in breast cancer to identify patients who may benefit from treatment with trastuzumab (Herceptin, Roche, New Jersey, US)[7]. Table 2 summarizes selected currently licensed therapeutic agents for which patient eligibility is determined according to the presence or absence of specific genetic aberrations within the tumour.

Table 2. Selected cancer medicines active against specific tumour genotypes

Drugs	Molecular marker predictive of response	Disease indication
ATRA (all-trans retinoic acid)	t(15;17) translocation	Acute promyelocytic leukaemia (APML)
Imatinib Dasatinib Sunitinib	t(9;22) translocation (Philadelphia chromosome); *BCR-ABL* fusion *KIT/PDGFRA* mutation	Chronic myeloid leukaemia (imatinib) Gastrointestinal stromal tumours
Trastuzumab Trastuzumab emtansine (antibody-drug conjugate) Pertuzumab	*HER2* gene amplification	Breast cancer Metastatic gastric cancer (trastuzumab only)
Cetuximab Panitumumab	wild-type *KRAS and NRAS*, i.e. lack of mutation	Metastatic colorectal carcinoma
Gefitinib Erlotinib Afatinib Osimertinib Crizotinib Ceritinib	*EGFR* mutation *ALK* or *ROS1* gene rearrangement	Non-small cell lung carcinoma
Vemurafenib Dabrafenib Trametinib	*BRAF* codon 600 mutation, especially V600E	Malignant melanoma
Olaparib	Somatic or germline *BRCA1/2* gene mutation	Ovarian cancer

The drugs have been selected as those that are approved for use in Europe and are named according to the World Health Organization's International Non-proprietary Name (INN) system, with the components indicating the type of drug: -mab, monoclonal antibody; -ib, small molecule drug with protein inhibitory properties; -tin-, tyrosine kinase inhibitor; -xi-, chimeric human-mouse monoclonal antibody; -zu-, humanised monoclonal antibody; -u-, human monoclonal antibody.

1.2.4 Progress towards routine delivery of predictive molecular analysis in solid tumours

There are a number of current initiatives around the world performing broad molecular profiling of tumours with a view to assisting treatment decisions. Since 2011, the Cancer Research UK Stratified Medicine Programme (CRUK-SMP) has been underway at a number of clinical and laboratory sites in England, Wales and Scotland (http://www.cancerresearchuk.org/funding-for-researchers/how-we-deliver-research/our-research-partnerships/stratified-medicine-programme). Phase one (SMP1) took place between 2011 and 2013 and piloted the routine delivery of mutational analysis of 4-5 prioritised genes of interest in different solid tumour types (breast, colorectal, lung, ovarian and prostate cancer as well as malignant melanoma). Nine thousand tumour samples have undergone analysis and the efforts have generated a wealth of insights into numerous aspects of this approach. Phase two commenced in summer 2013 and is currently underway with a sole focus on lung cancer. The genetic analysis involves profiling of a broader panel of genetic abnormalities using next generation sequencing technology. This should yield greater opportunities for patients to enter clinical trials and access novel treatments based on the results.

The 100,000 genomes project (https://www.genomicsengland.co.uk/the-100000-genomes-project/) is an ongoing initiative active at 13 designated genomic medicine centres in England and announced by Prime Minister David Cameron in December 2012. There are two main parts to the project, focusing on rare inherited disease and cancer. The rare disease programme involves whole genome sequencing of blood-derived germline DNA from a patient and two first degree relatives in order to identify and define disease-causing genetic aberrations. In the cancer programme, DNA extracted from a fresh-frozen or formalin-fixed tumour sample is put through whole genome sequencing with the patient's germline DNA from blood as a comparator, in order to examine somatic variations driving cancer maintenance and progression.

An increasing number of molecular profiling initiatives are underway internationally, involving solid tumour genetic analysis additional to what is

required for entry to specific clinical trials (Table 3). Most stratified medicine programmes are based at single organizations, but the INCa (*Institut National du Cancer*) initiative in France is one of few other national programmes. The French central government has provided funding since 2006 to a national network of 28 genetics laboratories in order to facilitate genetic analysis of tumour samples from any eligible patient. Similar to the CRUK-SMP approach, the laboratories were able to use a variety of methods including *in situ* hybridisation and targeted and screening sequencing techniques to detect clinically relevant mutations. The published data from this initiative details mutation and test failure rates but does not contain details of the exact scope of tests for each gene or the techniques in use at each laboratory (table 4).

Table 3. Selected international stratified medicine initiatives.

Country	Institution	Protocol/ study	Technology/ platform	Scope of analysis	Eligible tumour types	Reference
United States	Dana-Farber Cancer Institute, Boston, Massachusetts; Broad Institute, Cambridge, Massachusetts and Brigham and Women's Hospital, Boston	Profile	OncoPanel on Illumina HiSeq	645 genes	Solid tumours	http://www.dana-farber.org/Research/Featured-Research/Profile-Somatic-Genotyping-Study.aspx
	Memorial Sloan-Kettering Cancer Center, New York	Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT)	Illumina HiSeq	419 gene panel	Multiple solid tumour types, includes FFPE tissue	https://www.mskcc.org/msk-impact
	MD Anderson Cancer Center, Houston, Texas	IMPACT2: Randomized Study Evaluating Molecular Profiling and Targeted Agents in Metastatic Cancer	Foundation Medicine FoundationOne assay	315 genes	Multiple solid tumour types, includes FFPE tissue	https://clinicaltrials.gov/ct2/show/NCT02152254
	Michigan Center for Translational Pathology, Ann Arbor,	Personalized Oncology Through High-throughput Sequencing:	Illumina HiSeq	148 genes on core gene list	Multiple solid tumour types, includes FFPE tissue	http://mctp.med.umich.edu/physicians/mi-oncoseq-study

		MI-ONCOSEQ (Michigan Oncology Sequencing Center)				
	Vanderbilt-Ingram Cancer Center, Nashville, Tennessee	Personalized cancer medicine initiative (PCMI)	SNaPshot	6-8 genes	Melanoma, non-small cell lung cancer, colorectal and breast cancer	http://www.vicc.org/research/shared/translational/services/snapshot/
Norway	Nationwide	Norwegian Cancer Genomics Consortium	NGS	Whole exome	9 tumour types, both solid and haematopoietic	http://kreftgenomikk.no/
Canada	Princess Margaret Cancer Centre, Toronto	Integrated Molecular/ Community Oncology Profiling in Advanced Cancers Trial (IMPACT and COMPACT)	Sequenom Genotyping and/ or Targeted MiSeq NGS	23 genes and/or 48 genes	Multiple solid tumour types, FFPE tissue	http://www.cancergenomicsprogram.ca/about-cgp
France	Nationwide	Institut National du Cancer (INCa)	Various	Up to 8 genes	Multiple solid tumour types, FFPE tissue	http://en.e-cancer.fr/
Multinational	Worldwide Innovative Networking (WIN) consortium	Worldwide Innovative Networking Therapeutics (WINTHER) trial	Genomic and transcriptomic analysis	236 genes (DNA)	Multiple solid tumour types	Rodon et al. Challenges in initiating and conducting personalized cancer therapy trials: perspectives from WINTHER, a Worldwide Innovative Network (WIN) Consortium trial. *Ann Oncol.* 2015 Aug; 26(8):1791-8.

FFPE: formalin-fixed, paraffin-embedded tissue; NGS: next generation sequencing

Table 4. Data on mutation frequency and failure rates from the first few years of the French Institut National du Cancer (INCa) programme

Cancer type	Gene	2013			2012		2011			2010		
		Tested	Aberration detected	Failed tests	Tested	Aberration detected	Tested	Aberration detected	Failed tests	Tested	Aberration detected	Failed tests
Colorectal	*KRAS*	-	-	-	18,306	40%	17,003	40%	3%	16,581	38%	4%
	BRAF	-	-	-	-	-		9%	4%	4,457	8%	4%
Lung	*EGFR*	23,336	10%	8%	21,995	10%	20,750	10%	10%	16,800	11%	9%
	ALK	18,861	3.5%	13%	-	-		-	12%	-	-	-
	KRAS	22,958	27%	8%	-	-	-	-	-	-	26%	-
	BRAF	20,100	2%	9%	-	-	-	-	-	-	2%	-
Melanoma	*BRAF*	-	-	-	4,545	37%	3,479	38%	5%	1,835	39%	5%
	KIT	-	-	-	-	-	1,936	4%	11%	1,416	4%	8%

Data has been compiled from the INCa annual reports and publicly available meeting presentations [8-11].

1.2.5 Application of stratified medicine to selected tumour types included in the Cancer Research UK Stratified Medicine Programme

The tumour types included in SMP1 were selected because they represented common cancers - with breast, colorectal, lung and prostate cancer making up over half of all incident cancer cases in the UK each year and ovarian cancer being the fifth most common cancer in females[12]. Advanced malignant melanoma represents 4% of all new cancer diagnoses in both men and women and was included due to the link to targeted therapies[12].

Breast cancer: Breast cancer is the commonest cancer in the UK, with a total of 50,285 new diagnoses in 2011 of which 0.7% were men[12]. Most cases of are of ductal subtype, with invasive ductal carcinoma comprising 68%, ductal carcinoma in situ comprising 10% and invasive lobular carcinoma a further 10% of the total in a UK-wide audit of all new breast cancer diagnoses in 2006[13]. This audit found an oestrogen receptor (ER) positive rate of 85%, progesterone receptor (PR) positive rate of 69% and HER2 positive rate of 16%[13]. Ascertaining the HER2 status of invasive breast cancer has now been the standard of care for over a decade, and this is achieved using immunohistochemical assessment of protein expression in the majority of cases. In situ hybridisation (ISH) can be used to confirm gene amplification and is generally reserved for cases with equivocal immunohistochemistry results. Trastuzumab is used as neoadjuvant therapy and in patients with advanced HER2 positive disease, and newer developments include a further HER2 targeting agent pertuzumab as well as the licensing of the antibody-drug conjugate trastuzumab emtansine. In the past few years there has been interest in using gene expression profiling in breast cancer to provide risk stratification in addition to traditional histopathologically determined parameters such as grade, vascular invasion and lymph node involvement. Data from gene expression profiling tools such as Oncotype DX (21 genes, Genomic Health, California, US) and MammaPrint (70 genes, Agendia, California, US) can be used to identify a subset of patients with such a good prognosis that they can be spared adjuvant chemotherapy since the likely benefits would be less than risk of adverse effects[14 15]. A recent appraisal by the National Institute of Health and Care Excellence (NICE) of gene

expression arrays in breast cancer led to the approval of Oncotype DX in the UK but recommended further research to establish the utility of the IHC4 panel (immunohistochemistry for oestrogen and progesterone receptors, HER2 and the proliferation marker Ki67)[16]. In terms of molecular taxonomy, a landmark study using expression arrays led to classification of breast cancer into five molecular subtypes [17]. These were further expanded into ten subtypes in the METABRIC study published in 2012 [18], and although this work has led to mechanistic insights and the discovery of new driver mutations in breast cancer, translation to the clinic is still some way off. Trials are also underway in the setting of metastatic breast cancer of poly (adenosine diphosphate [ADP]-ribose) polymerase (PARP) inhibitor drugs, targeting deficient cellular homologous DNA repair processes, and these show particular promise in patients with breast cancer arising as a result of germline mutations in the *BRCA1* or *BRCA2* genes[19].

Lung cancer: Lung cancer is the second most common cancer in men and women with 43,463 new diagnoses in 2011 in the UK[12]. Approximately 11% of cases are classified as small cell lung carcinoma and the remaining 89% are non-small cell lung carcinoma[20]. The non-small cell carcinoma group comprises adenocarcinoma (53%), squamous cell carcinoma (34%) and a few more unusual subtypes including large cell carcinoma[21 22]. Over recent years a number of key driver mutations have been discovered in pulmonary adenocarcinoma, and those which have been clinically validated so far are *EGFR* mutations and *ALK* translocations. Novel targets also identified and linked to drugs in development or clinical trials include *KIF5B-RET* and *ROS*[23]. Progress in pulmonary squamous cell carcinomas has not been so promising, though occasional very rare cases with *EGFR* and *ALK* abnormalities have been described[24-27]. This presents a dilemma for treatment of these patients, since they are often not representative of the study population for clinical trials providing the evidence base on which drug approval is granted. Also this adds complexity to the process of determining optimal testing strategies, with economics of testing affected by the prevalence of the mutation in question. Molecular analysis in lung cancer is ahead of other tumour types in that the multiple tests now required have developed to be used sequentially, reinforced by the US model of an approved companion diagnostic test to accompany each drug. The relative anatomical inaccessibility of lung cancers and the resulting

small tissue samples compound the problem of limited tissue availability, with current analysis methods consuming significant amounts of DNA.

EGFR in lung cancer: The *EGFR* gene (epidermal growth factor receptor gene, also known as *ERBB1* or *HER1*) encodes the cell membrane-bound epidermal growth factor receptor (https://www.ncbi.nlm.nih.gov/gene/1956), and mutations in this gene determine response to tyrosine kinase inhibitors erlotinib and gefitinib in patients with non-small cell lung cancer (NSCLC)[28,29]. 90% of *EGFR* mutations are located in the tyrosine kinase binding domain (exons 18-21). The mutant EGFR protein activates cellular pathways implicated in cancer cell growth, survival, and migration. The most common activating mutations are exon 19 deletions (45-55% of mutations) and a codon 858 missense mutation in exon 21 (L858R,35-45% of all mutations)[30]. The most common resistance mutation is *EGFR* T790M, but other mechanisms of resistance, such as amplification or over-expression of *MET*, *PIK3CA* mutations and transformation to small cell lung carcinoma have also been described[29].

Clinical associations of *EGFR* mutations have been recognized and the most strongly correlated are female gender, a history of never having smoked cigarettes and East Asian ethnicity[31]. Histological features associated with *EGFR* mutations are adenocarcinoma of any growth pattern, especially well-differentiated papillary or micropapillary tumours, but with the exception of mucinous carcinomas which are instead associated with *KRAS* mutations in common with mucinous neoplasms arising in other organs[32]. *EGFR* mutant tumours invariably show immunohistochemical expression of TTF1, a commonly used marker of pulmonary adenocarcinoma [33]. *EGFR* mutations appear much less common in pulmonary neuroendocrine, mucoepidermoid and adenoid cystic carcinomas [29].

The mutant EGF receptor can be targeted using small molecule inhibitor drugs, which act inside the cell against the internal tyrosine kinase domain of the receptor, so-called 'tyrosine kinase inhibitors', TKIs, such as erlotinib (Tarceva, Genentech, California, US) or gefitinib (Iressa, AstraZeneca, Cambridge, UK). These drugs have demonstrated clinical response in lung cancer patients in clinical trials. There are also therapeutic antibodies active against the extracellular domain of the EGF receptor, cetuximab (Erbitux, Merck Serono,

Darmstadt, Germany) and panitumumab (Vectibix, Amgen, California, US), which are used in patients with head and neck or colorectal cancer[34 35].

EML4-ALK in lung cancer: *The EML4-ALK* fusion gene is derived from an inversion affecting chromosome 2 and leading to fusion of the *EML4* (echinoderm microtubule-associated protein-like 4) gene (https://www.ncbi.nlm.nih.gov/gene/27436) with the *ALK* (anaplastic lymphoma kinase) gene (https://www.ncbi.nlm.nih.gov/gene/673). This fusion gene is found in 4-7% of unselected non-small cell lung cancers and these are nearly all adenocarcinomas, but the *ALK* fusion has also rarely been detected in squamous cell carcinomas[36 37]. Alternative *ALK* fusion partners (e.g. TRK-fused gene *TFG*, *NPM* and *KIF5B*) have been described but are much less common than *EML4-ALK*[38]. *ALK* gene rearrangements generally occur exclusively of *EGFR* or *KRAS* mutations, though this may simply reflect the fact that both are relatively uncommon events and therefore statistically unlikely to co-exist. The fusion gene encodes a fusion protein with over-activity of ALK due to ligand-independent dimerization, and ALK signalling leads to cellular proliferation. Clinical correlates of the presence *ALK* mutation are never or light cigarette smoking history, younger age at onset of disease and there is also a strong association with adenocarcinoma showing a signet ring or acinar growth pattern[36 37]. The ALK/MET inhibitor crizotinib (Xalkori, Pfizer, New York, US) is a multi-targeted small molecule tyrosine kinase inhibitor, administered orally, which inhibits ALK phosphorylation and signal transduction. Crizotinib was licensed for use in NSCLC by the United States' Food and Drug Administration (FDA) in 2011. Unusually, the FDA's accelerated approval was based not on evidence of survival benefit, but instead on trial data demonstrating a response rate of up to 57% in patients with a fusion-gene positive tumour[39]. A subsequent phase III trial has demonstrated superiority of crizotinib over standard chemotherapy with an end-point of progression-free survival[40]. At the time of approving crizotinib, the FDA also licensed a specific break-apart fluorescent *in situ* hybridisation (FISH) probe (Abbott Diagnostics, California, US) as the requisite companion diagnostic for detection of the *ALK* gene translocation. Crizotinib is an orally administered drug and the major side-effect is transient visual disturbance, affecting up to two-thirds of patients, with gastrointestinal disturbance, fatigue, pneumonitis and abnormal liver function tests being less common

(https://www.pfizerpro.com/product/xalkori/hcp/safety-profile). Crizotinib resistance mutations have been detected following therapy[41]. Crizotinib is also being investigated as a treatment for aggressive and resistant forms of anaplastic large cell lymphoma and neuroblastoma in the paediatric population[42]. Second generation ALK inhibitor drugs are now available for crizotinib-resistant disease and the FDA approval of the first of these therapies, ceritinib, was remarkable for being based on the data from a phase one clinical trial, demonstrating evolution of the drug approval process in response to the success of specific targeting of therapies to pre-defined genetic aberrations in their tumours[43].

Colorectal cancer: Colorectal cancer is the third most common cancer in men and women with 41,581 registered new diagnoses in 2011 in the UK[12] and nearly all cases are adenocarcinomas. The *KRAS* gene (12p12.1) is a commonly mutated cancer gene, with mutations occurring most commonly in codon 12 but also in codons 13 and 61 and found in 30-40% of colorectal cancers as well as 8% of non-small cell lung cancers (mostly adenocarcinomas)[34]. Patients with codon 12 and 13 *KRAS* gene mutations in their tumours do not appear to show clinical response to EGFR inhibitors such as cetuximab due to downstream activation of the mutated KRAS protein, but there is some evidence to suggest that not all mutations are equal, with evidence that the G13D mutation is associated with a response to cetuximab close to that of patients with wild-type genotype[44]. Although no drugs are currently licensed that directly target mutant KRAS, strategies using newer targeted therapies in combination with chemotherapy or other targeted therapies, for example, combined PI3K and MEK inhibition, are under investigation in clinical trials.

Malignant melanoma: 13,348 new diagnoses of malignant melanoma were registered in the UK in 2011 and this represents a cancer type that has been increasing in incidence in the past decade[12]. Possible explanations for this increase are lifestyle factors such as increased exposure to ultraviolet radiation through use of sun beds for skin tanning, increased awareness and surveillance leading to earlier diagnosis, and also changes in pathological classification[45 46]. Interpreting, classifying and predicting the clinical behaviour of atypical melanocytic lesions is an accepted area of difficulty and inter-observer variability in diagnostic histopathology, and any reduction in the

threshold for histological diagnosis of melanoma in atypical melanocytic lesions would contribute to an increase in melanoma incidence[47].

The mutated *BRAF* oncogene was identified as a key driver in just over half of malignant melanomas in 2002[1]. The *BRAF* gene (7q34) encodes a serine/threonine kinase, an enzyme activated by phosphorylation and responsible for transferring phosphate groups to other proteins to modulate their function that is part of the Raf kinase family (https://www.ncbi.nlm.nih.gov/gene/673). *BRAF* mutations are found in 8% of all solid tumours, including 40-60% of malignant melanomas, 5-15% colorectal adenocarcinomas, 35% of low-grade/borderline serous ovarian tumours and 1-3% of all non-small cell lung cancers [1 48]. Over 90% of *BRAF* mutations are found in codon 600, the commonest being V600E which accounts for up to 30% of *BRAF* mutations in melanoma[49]. The BRIM3 trial provided evidence that patients with previously untreated, unresectable stage IIIC/IV melanoma with V600E mutation had improved overall and progression-free survival with vemurafenib when compared to standard dacarbazine therapy[50]. An unexpected finding was the increased risk of cutaneous squamous cell carcinoma in patients receiving vemurafenib therapy, and a possible mechanism for this is paradoxical stimulation of events in a related cellular pathway in epidermal cells with wild-type *BRAF*. Dose interruption and modification was required in 38% of patients in BRIM3 but this oral therapy is generally well-tolerated. Ongoing studies are focusing on improving the durability of response to BRAF inhibitors by trialling them in combination with other targeted agents acting on related pathways. The MAPK (mitogen activated protein kinase) pathway also shows overactivity in melanomas harbouring *BRAF* mutations, and the BRAF inhibitor dabrafenib and MEK inhibitor trametinib have recently been approved for use in patients with metastatic melanoma whose tumour shows codon 600 *BRAF* mutations[51 52].

Of cutaneous melanomas lacking *BRAF* mutations, 15-20% instead show mutations in the *NRAS* (neuroblastoma rat sarcoma virus) gene, which encodes another molecule in the MAPK pathway[53]. *NRAS*-mutated melanomas appear to carry an adverse prognosis, independent of other prognostic parameters, and clinicopathological correlates are patient age greater than 55 years, tumour location on the extremities, an increased Breslow thickness and higher mitotic rate[54-56]. In contrast, *KIT* gene mutations are characteristically found in 30% of melanomas that arise at either acral (palms, soles or sub-ungal), mucosal or

chronically sun-damaged sites, the latter defined by the presence of dermal solar elastosis, and are also associated with a lentiginous growth pattern [57]. Point mutations in exon 11 (L576P) or exon 13 (K642E) are most common and also appear to represent an independent prognostic marker, with patients with *KIT*-mutated melanomas having decreased survival compared to those with *KIT* wild-type melanomas[58].

Ovarian cancer: Ovarian cancer is the fifth most common cancer in women and is also the fifth most common cause of female cancer-related death in the UK[59]. The incidence of ovarian cancer has been increasing over the past few decades and most patients are diagnosed with an advanced stage of disease. There are five main, well-characterised different types of ovarian cancer which in descending order of incidence are high-grade serous carcinoma, clear cell carcinoma, endometrioid carcinoma, mucinous carcinoma and low-grade serous carcinoma. In recent years, two distinct pathways of ovarian carcinogenesis have been recognised leading to either low-grade or high-grade tumours. Low-grade carcinogenesis encompasses all of the above types apart from high-grade serous carcinoma and progresses slowly through borderline tumours as an intermediate step. The high-grade pathway, in contrast, leads to high-grade serous ovarian carcinoma, in which ubiquitous *TP53* mutations are found. Clinically these tumours behave aggressively and present late. Hereditary and/or acquired *BRCA1/2* mutations are also implicated in high-grade serous ovarian carcinomas, and the resulting defect in cellular DNA repair machinery represents a target for therapy through the 'synthetic lethality' route: causing cell death through impairment of a different DNA repair pathway mediated by PARP. The first PARP inhibitor drug, olaparib, received European marketing authorisation in December 2014 for use in patients with high-grade serous ovarian carcinoma and germline and/or somatic *BRCA* gene mutations, based on the results of a phase III trial[60].

Prostate cancer: Prostate cancer is the most common cancer in men and the second most common cause of cancer-related death in men in the UK after lung cancer[59]. *PTEN* and the *TMPRSS2-ERG* gene fusion have been identified as driver genes showing recurrent aberrations in prostate cancer but have not yet reached clinical utility. The *TMPRSS2-ERG* gene fusion is found in approximately 50% of prostate cancers and mutations or deletions of the *PTEN*

tumour suppressor gene are found in up to 30% [61-63]. *TMPRSS2* (21q22.3) encodes a serine protease and the gene is regulated through an androgen-dependent promoter region (https://www.ncbi.nlm.nih.gov/gene/7113). *ERG* (ETS-related gene, 21q22.2) is a member of the ETS (erythroblastosis- virus E26 transformation-specific) transcription factor family (https://www.ncbi.nlm.nih.gov/gene/2078). The downstream effects of the gene fusion include up-regulation of Wnt pathways and down-regulation of TNF/cell death pathways. Several morphological features have been associated with *TMPRSS2-ERG* fusions. These include the presence of blue-tinged mucin, a cribriform growth pattern, signet ring morphology, prominent macronucleoli, intraductal tumour spread and the presence of collagenous micronodules within the tumour. The presence of a *TMPRSS2-ERG* gene fusion has been associated with a worse prognosis, and also possibly linked to sensitivity to abiraterone and PARP inhibitor drugs[64]. Abiraterone was approved for use in patients with treatment-resistant prostate cancer in 2011, and targets the androgen/androgen receptor pathway.

The tumour suppressor gene *PTEN* (10q23.3) encodes a lipid phosphatase that negatively regulates the PI3K-AKT pathway, and loss of *PTEN* leads to constitutive PI3K-AKT signalling (https://www.ncbi.nlm.nih.gov/gene/5728). *PTEN* is inactivated in many cancers through various mechanisms. PTEN loss of heterozygosity is used as a marker of somatic deletion of the *PTEN* tumour suppressor gene, by comparing relative quantity of polymorphic components of *PTEN* gene DNA (e.g. microsatellites or short tandem repeats) in normal (ideally germline) and tumour tissue. Due to contamination by normal cells in the tumour (e.g. stroma/blood cells), a reduction of one *PTEN* allele rather than complete disappearance may be seen in the tumour. *PTEN* deletion confers potential sensitivity to PI3K inhibitor drugs.

1.2.6 Molecular analysis techniques

Detection of gene amplification or translocations may be performed in thin sections on the glass slide using in situ hybridisation, whereas detection of gene mutations or fusion transcripts requires extraction of nucleic acids and analysis using PCR and sequencing-based methods (table 5). Some screening techniques involve comparison to a known normal sample, such as high

resolution melt (HRM) analysis and single strand conformation polymorphism analysis (SSCP/SSCA). This allows identification of those samples that are not normal, for further work to characterize the precise abnormality present if required. Determination of the sequence can be achieved by conventional Sanger (dideoxy-) sequencing, which is considered to be more labour intensive and have a lower sensitivity than more modern next generation sequencing technologies. The limit of detection for direct sequencing is generally considered to be 20%, i.e. a mutation must be present in 20% of the DNA within a sample to be confident of picking it up by direct sequencing analysis[65]. It is currently uncertain what effect – if any – mutations present at a low frequency within a tumour have on overall biological behaviour, and therefore whether there is a threshold of significance. Pyrosequencing is a similar but slightly more sensitive technique for sequence determination and mutation detection, with an estimated limit of detection of 5%.

Methods for targeted analysis include amplification refractory mutation system (ARMS)[66], which is a technique in combination with real-time quantitative PCR to selectively amplify those sequences containing a defined mutation over those that don't, i.e. are 'wild-type' as well as fragment length analysis. Fragment length analysis can be used to detect insertions or deletions but not point mutations.

The analytical sensitivity of the mutation detection methods in use currently is between 75-90% for sequencing and HRM analysis and greater than 90% for pyrosequencing, SSCP, fragment length analysis, next generation sequencing and allele-specific PCR[65 67-69]. The choice of technique involves a trade-off balancing analytical sensitivity (ability to correctly detect mutation-positive cases) and limit of detection (minimum detectable percentage of mutant vs. wild-type alleles in a sample) with the specificity of the method.

Table 5. Selected techniques for mutation detection

Method	Brief description of method	DNA input required	Sensitivity	Limit of detection*	Main advantages	Main disadvantages
Sequencing- based screening methodologies (detection of all variants: known and unknown)						
Sanger (dideoxy-) sequencing	Amplification and sequencing of PCR products by selective incorporation of chain-terminating dideoxynucleotides during in vitro DNA replication.	Low (~100ng)	Lowest	10-20%	Identification of known and novel variants	Labour-intensive; may miss low-level variants
Pyrosequencing	Sequencing by synthesis: detection of the luminescence released when a pyrophosphate-labelled nucleotide molecule is incorporated during DNA synthesis	Low (~100ng)	High	5%	Can also be used for targeted mutation detection Fast method with real-time read-out	Comparatively high sequencing error rate
Next generation sequencing	Massively parallel sequencing of thousands-millions of amplified DNA molecules using capture- or amplicon-based approaches	High (~500ng) – dependent on scope of analysis	High (dependent on read depth)	10%	Highest throughput technology, enabling greater scope of analysis up to whole exome/genome	Larger input quantities of DNA required and complex data interpretation requirements

Screening methods using comparison of mutated with normal DNA (detection of all variants, known and unknown)						
High resolution DNA melting analysis (HRM)	Screening of samples using comparison of the melting curves of PCR products against known normal samples	Low (~100ng)	Low	5%	Quick, melting products can be sequenced to identify exact abnormality	Non-specific amplification of products can lead to mis-calls
Single strand conformational polymorphism analysis (SSCP or SSCA)	Heat-denatured PCR products are compared to known samples using capillary electrophoresis and analysed according to electrophoretic mobility	Low (~100ng)	High	1-10%, varies by mutation	Established technique, low-cost	Technical parameters (e.g. temperature, gel composition) must be strictly controlled
Targeted mutation detection methods (genotyping of known mutations only)						
Amplification refractory mutation system	Selective amplification of sequences containing a defined mutation over those that don't	High (~500ng)	Highest	<1%-8%, varies by mutation	Fast and sensitive technique	Only detects pre-defined hotspot mutations
Fragment length analysis	DNA fragment length analysed against size standards to detect deletions and insertions	Low (~100ng)	High	1-2%	Fast and sensitive technique	Cannot be used to detect point mutations

*% mutant alleles in wild-type background

1.2.7 Sample preparation for molecular analysis

Molecular analysis protocols have been developed for FFPE tumour tissue, derived from either tissue sections or cytological cell block preparations formed from a cell pellet. Material can be submitted as sections on glass slides or as scrolls and ideally a matched haematoxylin and eosin (H&E) stained section should also be provided, with the area containing tumour outlined on the slide and the percentage tumour nuclei content of this area assessed by the referring pathologist. Some laboratories perform macrodissection of slide-mounted sections to isolate and enrich the material for tumour nuclei and avoid analysis of any adjacent non-neoplastic tissue. The workload implications of this are significant and the development of more sensitive analysis techniques may mean that this is not required in future. Macrodissection is generally considered mandatory with current methods if the tumour content of the section is assessed as less than twenty per cent by the reporting pathologist[70]. For currently available methods, mutation analysis for mutation hotspots in up to 5 genes can be performed on DNA extracted from a single 5µm paraffin section with tumour cellularity greater than 50%. Each section can be expected to yield at least 150ng DNA, with inputs as low as 10ng yielding a meaningful result, but variability in quality of DNA due to the effects of formalin fixation may mean that only a small proportion of the extracted DNA can be amplified and that larger amounts of starting material are required in order to compensate[71].

An attempt to set thresholds for tumour content and cellularity for *EGFR* mutation testing has been reported by the molecular diagnostics team at the Royal Marsden Hospital, after evaluating the first two years of their service using a combination of targeted methods including an allele-specific PCR-based kit, fragment length analysis and direct sequencing[33]. Of 115 samples, those assessed by a pathologist as showing good overall cellularity and tumour content greater than 30% (n=64) were associated with a 91% test success rate, which was not significantly different to those assessed as showing good cellularity but less than 30% tumour content (n=13). A lower success rate (77%) and comparable mutation rate was reported for those samples assessed as showing overall poor cellularity but containing representative tumour (n=32).

The remaining six samples were assessed as scantily cellular or necrotic and of these only two yielded a result, neither of which revealed mutations[33]. This suggests that despite the understandable desire to attempt analysis and obtain a result on any available patient sample, there are a small minority in which testing can be predicted to be unsuccessful and the expense and delay of failed analysis can be avoided by rejecting the specimen. Communication of the evidence and rationale underlying this decision-making process to clinical teams responsible for acquiring and submitting samples will contribute to a better understanding of the pre-requisites for successful molecular analysis, possibly also driving up sample quality over time.

Accurate assessment of tumour content by the pathologist is therefore clearly of critical importance. Furthermore, for highly sensitive next generation sequencing approaches, the percentage tumour nuclei content in the starting material informs mutation and wild-type calling algorithms in the analysis pipeline and therefore the confidence and certainty of the result. Histopathological practice has evolved to rely on pathologist estimation of the proportion of tumour versus non-tumour nuclei in a tissue section, informally referred to as 'eyeballing', rather than any systematic method of accurately quantifying proportions of different tissue components. A recently published study compared assessment of tumour content in a series of H&E-stained lung cancer biopsy (n=24) and resection (n=24) tissue sections by experienced pathologists to tumour content designation by manual cell counting [72]. The pathologists were asked to classify tumour content into 0-5%; 6-10% and subsequent categories with increments of 9% up to 91-100%. Taking the pathologists as a group alone, there was an average range of six categories between the highest and lowest estimates per case. 33% of estimates deviated by at least three categories from the 'gold standard' result obtained by cell counting. The study also identified systematic bias between different pathologists, including a tendency to either serial under- or over-estimation as well as seemingly random errors. In a more recent publication, researchers from Belfast describe the first automated tumour content assessment system, TissueMark[73]. This application can be used to annotate the tumour area and then perform computerised image analysis to determine percentage tumour nuclear content, and in the study was applied to a series of 136 slides from lung carcinoma resections. The annotation tool showed 97.25% accuracy in

correctly determining presence of tumour in a section, with three slides misclassified as containing tumour due to the presence of necrosis, reactive pneumocytes and areas of dense lymphocytic infiltrate. There was a high level of overlap between the areas of the slide annotated by pathologists and computer analysis, as assessed by a pathologist comparing both images following independent marking up. Accurate tumour content determination by manual cell counting was performed on selected 1mm^2 areas for 10 cases, and there was good concordance for all cases, with the automated assessment value lying within 10 units of the manually determined value (correlation coefficient, $r=0.972$, $p < 0.0001$). The authors commented on how time-consuming the manual cell counts were, taking four hours per case, compared to three minutes per case for automated image analysis.

The increasing use of digital pathology for education, research and clinical applications provides an opportunity to develop and deploy digital image analysis aids to accurate tumour content assessment, contributing to the increasing complexity but hopefully also accuracy of the service provided.

As well as assessing tumour content, the pathologist can provide useful information to the molecular laboratory about other specimen-dependent factors that may influence the likelihood of success in subsequent analysis, such as the presence of inhibitors of the polymerase chain reaction (PCR) including necrosis, or melanin pigment in malignant melanoma. Specimen handling during the pre-analytical phase has an important impact on the outcome of mutation analysis and there is potential for optimisation by simple changes in practice in cellular pathology laboratories, such as use of a clean microtome blade for cutting sections from each new block in order to prevent cross-contamination of DNA and tissue between samples.

Following analytical and clinical interpretation, the output of the molecular analysis is formulated into a report for the requesting clinician or pathologist. ISO (the International Organization for Standardization), the College of American Pathologists and UK National External Quality Assessment Scheme (NEQAS) have all issued guidance on the contents of this report (summarized in table 6)[74]. In some centres the report is sent to the referring clinician only and filed in the patient record, and in others the report is received by the histopathologist and issued as a supplementary report or integrated molecular

pathology report. Since the mutational profile is an attribute of the tumour rather than the patient, the latter approach seems more logical and may allow the molecular results to be further interpreted in the context of the morphological and immunohistochemical features of the tumour.

Table 6. Requirements of a molecular pathology report

Dates: of sample receipt and report authorization **Patient information:** 3-point identifiers e.g. patient name, date of birth and reference number **Information about request:** Nature of sample, tissue and tumour type, percentage content tumour nuclei as assessed by referring pathologist, clinical indication for analysis, name and address of referrer **Information about the analysis**: technique(s) used, scope of test, sensitivity/limit of detection **Results:** Presence or absence of abnormality in gene(s) in question expressed using standard HGVS nomenclature, interpretation of clinical significance of result (may be unknown) **Contact information:** For laboratory as well as name/job title of person authorizing report

1.2.8 Challenges of formalin-fixed, paraffin embedded (FFPE) material

There are technical challenges involved with molecular analysis of FFPE material. Formalin fixation leads to cross-linking and degradation of DNA into fragments typically less than 200 base pairs in length. Both over- and under-fixation of specimens should be avoided since either can cause problems. There is a need for optimal, standardized sample handling protocols in the pre-analytical phase to maximize the potential for obtaining diagnostically useful information from DNA extracted from FFPE tumour samples. Advances in interventional techniques such as endobronchial ultrasound-guided sampling fine needle aspiration (EBUS-FNA) of lung or mediastinal lymph node lesions have enabled combined cytological diagnosis and staging, but contribute to a trend for smaller samples.

The use of formalin as a histological fixative dates back to 1892, when Ferdinand Blum made the observation that contact with formaldehyde solution hardened the skin of his fingers during experiments in Frankfurt to investigate its use as an antiseptic agent[75]. Fixation using formalin (usually buffered neutral aqueous 10% solution of 4% formaldehyde, pH neutral buffered formalin or NBF) is a crucial step in tissue handling in order to preserve cellular detail for morphological assessment. Buffers, most commonly phosphate, are added to the formalin solution in order to remove precipitates and pigments resulting from formation of formic acid during fixation.

Formalin is a cross-linking fixative, exerting its effects in preserving structural integrity of cells and tissue primarily by formation of methylene bridges within DNA and between the amino groups of proteins. As well as stabilising the tissue ultrastructure, the effects on proteins serve to inactivate enzymes that might otherwise degrade the tissue. Formalin penetrates tissue by diffusion at a rate of 0.5-1mm per hour, but fixes it slowly with ongoing cross-linking reactions observed for a period of at least two weeks under experimental conditions[76-78]. The various chemical reactions between formalin and tissue components are still incompletely understood.

Formaldehyde exists mostly as its non-reactive hydrate methylene glycol (N-methylol) in solution:

$$HCHO + H_2O \leftrightharpoons CH_2(OH)_2$$

formaldehyde *methylene glycol*

Methylene glycol leads to the formation of methylene bridges between adjacent bases through electrophilic attack on the amino base and also to mono-methyl group (-CH_2OH) additions, which can be reversed to some extent by heating the nucleic acids in the presence of buffer. The cross-linking process is accelerated at high temperatures and increased pH. Further reactions in solution lead to the formation of polymers which can precipitate out of solution as paraformaldehyde. Formalin molecules also

cross-react in solution and the equilibrium is described by the Cannizzaro reaction:

$$2[HCHO] \leftrightharpoons CH_3(OH) + CH_2O_2$$

formaldehyde — *methanol* + *formic acid*

The cross-linking effect of aldehyde fixatives including formalin increase the susceptibility of the DNA to mechanical damage, and also create a three-dimensional matrix that reduces accessibility to the enzymes used during DNA purification or extraction. Physical shearing of DNA can also occur during isolation or extraction of nucleic acids from FFPE. Exposure to acids, unbuffered formalin or high temperatures causes damage through hydrolysis, with breakage of the glycosylic bonds attaching purine bases to the ribose ring. Exposure to alkali causes hydrogen atoms to change their position within a base leading to the formation of tautomers and non-standard base pairs, which may cause the introduction of mutations during DNA replication that were not present *in vivo*. Several substances used in tissue preparation have been found to act as inhibitors of DNA polymerase and therefore amplification by PCR[79] (table 7).

Table 7. Inhibitors of PCR

Substances used in tissue preparation:

- Paraffin wax
- Cross-linking by formalin
- Haematoxylin

Substances endogenous to tissue:

- Residual fragments of low molecular weight DNA
- Melanin pigment
- Necrotic cellular material
- Iron in haemoglobin
- Calcium ions
- Collagen

Apurinic/apyrimidinic sites are produced by hydrolysis of N-glycosylic bonds. Hydrolysis of phosphodiester bonds also occurs over time, leading to strand breaks. Some of this damage is irreversible although commercial DNA repair kits are available to try and optimise DNA quality prior to use in downstream applications (PreCR® Repair Mix, New England Biolabs, Massachusetts, United States and Restorase® DNA Polymerase, Sigma-Aldrich, St Louis, Missouri, United States).

There is evidence from comparison of formalin-fixed to fresh frozen tissue or analysis with next generation sequencing methods than formalin introduces chemical sequence artefacts that can mimic mutations. This is thought to occur at a rate of one mutation per 500 bases[80], taking into account a background sequencing rate of approximately 1%[81]. The most common event is an apparent C to T mutation, and possible mechanisms proposed for this include deamination of cytosine so that it is misread as uracil by the polymerase enzyme during PCR, or cross-linking of cytosine residues on adjacent strands so that they are missed out by the polymerase enzyme. Degraded DNA fragments may also lead to 'jumping' of the polymerase enzyme between two different template molecules creating a single strand with a novel sequence[82]. Since these effects occur during PCR, they will be amplified for representation in the sequencing reaction, with their relative abundance proportional to the total amount of starting material. This highlights the importance of having adequate amounts of input DNA to overcome these artefacts and generate reliable sequencing data.

The process of DNA deterioration is accelerated when sections are cut from the block and stored, rather than leaving the block intact, due to oxidation of DNA. Loss of epitopes and antigenicity is also observed and this may cause problems with subsequent immunohistochemistry. Possible solutions to this include refrigeration or freezing of stored sections and blocks (at 4ºC), dipping slides bearing cut sections in paraffin wax to preserve antigenicity or coating them in a proprietary tape (Path Inst Corp, Japan) to prevent oxidation. For liquid cytology preparations, an induced clot containing cells, plasma and thrombin can be dropped onto a Whatman Flinders Technology Associates (FTA) card, containing a cellulose matrix and stored in this form. Further nucleic acid degradation occurs over time

within paraffin blocks and is thought that incomplete exclusion of water during tissue processing is a contributory factor, since this entrapped water leads to slow ongoing hydrolysis of nucleic acids within the tissue. Hydrolysis may also occur during tissue staining processes in the laboratory, when sections are incubated in aqueous solutions.

The pre-analytical factors that may impact on sample quality and the success of subsequent molecular testing are summarized in table 8.

Table 8. Factors in the pre-analytical phase likely to have an impact on the likelihood of success in subsequent molecular analysis

Stage	Factor	Impact
Sample acquisition	*Warm ischaemia*: time elapsed during surgical procedure from ligation of blood supply to removal of sample from patient	Longer time will lead to enzymatic modification and degradation of sample which will adversely affect subsequent analysis
	Cold ischaemia: time elapsed between removal of sample from patient and immersion in formalin ('time to fixation')	
Fixation	Volume of formalin used	Sub-optimal immersion in formalin or inadequate tissue penetration through large, intact specimen will lead to inadequate tissue preservation
	Concentration of formalin used	
	Size of sample/tissue penetration of formalin	
	pH of formalin/nature of buffer or other chemical additives such as preservatives	Phosphate buffer appears to cause least damage to nucleic acids[83]
	Total fixation time	Over- or under-fixation both adversely affect tissue and nucleic acid preservation
Tissue processing	Processing time	Over- or under-processing both adversely affect tissue and nucleic acid preservation
	Processing system/reagents used	Variability in chemicals, pressure, temperature and use of other adjuncts such as ultrasound or microwaves may affect the reproducibility of results between different laboratories Incomplete exclusion of water during dehydration stage of tissue processing leaves entrapped water molecules that cause slow subsequent hydrolysis of nucleic acids

Stage	Factor	Impact
Tumour block selection	Selection of most representative block/best preserved tumour by reporting histopathologist	Determinant of success of subsequent molecular testing
	Tumour content: Absolute number of tumour cell nuclei and also proportion of tumour vs non-tumour tissue (e.g. stroma, inflammatory cells)	Reliability of result – inadequate representation of tumour DNA in sample submitted for sequencing may lead to false negative mutation analysis.
	Presence of inhibitors of PCR e.g. necrosis, melanin	May lead to failure of subsequent molecular analysis
Storage of paraffin block	Age of block used for molecular testing	Degradation of nucleic acids may occur over time and baseline quality of the DNA in the paraffin block will be an important factor in determining the time interval before the DNA is no longer fit for testing purposes.
	Timing of cutting sections	Oxidation of exposed surface of cut sections during storage may cause further deterioration in sample quality
	Temperature of storage	Storing at too high a temperature may accelerate the deterioration of a sample
Multiple stages: opportunity for cross-contamination of DNA	Transfer of DNA between different samples at the cut-up, processing, embedding or microtomy stages	DNA contamination may lead to false positive result on mutation analysis

The importance of a thorough understanding and quality assurance of every step of the process involved in generating a result for patient care was highlighted by the Royal College of Pathologists' investigation into events at King's Mill Hospital in Nottinghamshire in 2012-3[84] on behalf of the Care Quality Commission. This followed national media coverage regarding suspected inaccuracy of oestrogen receptor immunohistochemistry performed on samples from patients with breast cancer. A potential issue was raised through a regional NHS Breast Screening Programme audit, interpreted as showing outlier status associated with lower than expected oestrogen receptor positivity rates at this hospital. This appeared to be confirmed by re-testing of cases performed by a cellular pathology department in an external Trust, but in the opinion of the investigating team from the Royal College of Pathologists this repeat analysis used an oestrogen receptor clone (6F11) that was known to be over-sensitive and prone to false positive results[85]. Also the department had a low workload compared to other units in the region, contributing small numbers to the audit and leading to wide confidence intervals and a plausible explanation for the apparently low oestrogen receptor positivity rates. The investigating team paid close attention to laboratory sample handling and recommended standardisation of breast biopsy fixation times, closer monitoring of formalin pH within the laboratory as well as acquisition of control material for oestrogen immunohistochemistry from breast resections at the time of initial specimen dissection rather than at the time of specimen disposal six weeks later.

The literature on the impact of pre-analytical phase factors contains few systematic studies of the different steps. Baloglu *et al* (2008) compared DNA yield, ease of amplification and suitability for fluorescent in situ hybridisation of 3mm diameter punch biopsies of normal colonic tissue[86]. These were obtained from three surgical resection specimens, fixed in one of six different fixatives for time points of between 1 and 48 hours and then processed and embedded in paraffin wax, with total DNA yields varied between 0.60 and 7.58μg across all samples. DNA extracted from ethanol-fixed samples consistently exceeded 100bp in length and generated strong bands on gel electrophoresis, indicating intact and high quality DNA, but one limitation of this study was the lack of assessment of morphology or compatibility with

immunohistochemistry. DNA quality on gel electrophoresis appeared optimal for samples fixed in formalin for exactly 24 hours, and inferior for all samples fixed in formalin for markedly less or more than 24 hours[86]. Despite the variable findings on electrophoresis, all samples yielded DNA suitable for PCR amplification of a 268bp fragment of the β-globulin gene. Chung *et al* (2008) studied the effect of time to fixation, fixative buffer, duration of fixation and tissue processing regimen on the quality of RNA from formalin-fixed paraffin-embedded rodent kidney tissue. They found that fixation times less than 12 hours and exceeding 48 hours gave shorter lengths and lower yields of RNA, and that phosphate-buffered formalin gave the best quality RNA with unbuffered formalin performing worse than that with Tris- or calcium chloride buffer. Longer tissue processing times (range 140-660 minutes) gave better quality RNA as assessed by BioAnalyzer traces (Agilent), branched DNA assay (QuantiGene, Panomics) and real-time quantitative PCR for *GAPDH* and *CDK4* genes. It is widely accepted that altering pre-analytical variables may adversely impact morphology assessment and necessitate changes in immunohistochemistry protocols, however a further finding in Chung's study was that the shortest fixation and processing times led to brittle tissue sections that could not be cut at the standard 4μm thickness on the microtome[83]. 6-24 hours is widely quoted in the literature and expert consensus guidelines such as those produced by the College of American Pathologists and Royal College of Pathologists[87 88] as the optimal duration for tissue fixation, but this is at risk of being an over-simplification since this restricted range does not take into account the myriad factors affecting the fixation process in the wide variety of sample types received in a cellular pathology department. There are valid reasons for wishing to fix a sample for longer, for example in the UK professional guidelines recommend fixation of colorectal cancer excision specimens for at least 24-48 hours prior to dissection[87]. This is recommended in order to facilitate identification of lymph nodes in well-fixed mesocolic/mesorectal adipose tissue and maximise the yield of lymph nodes, a factor which has been associated with more accurate staging[89-92].

Other investigators have attempted to augment the fixation process through the use of ultrasound or microwave energy. The application of ultrasound to tissue samples accelerates fixation by facilitating formalin tissue penetration,

and in a series of post mortem tissue samples from multiple organs of five different patients, with the highest DNA and RNA yields from tissue samples subjected to ultrasound and fixed in formalin for 15-30 minutes. The authors deemed morphology acceptable despite the short fixation time, but images were not provided in the paper to support this and there was a focus on autopsy and forensic applications of the technique rather than diagnostic surgical pathology[93].

The available literature on factors affecting tissue quality can be divided into papers concerning morphology, DNA, RNA or protein analysis. A recent meta-analysis[94] of published literature on the pre-analytical phase in pathology from a group at the US National Cancer Institute in Bethesda acknowledged the multitude of variables influencing the quality of formalin-fixed, paraffin embedded tissue. Despite an extensive review of published literature spanning a period of over thirty years, the authors were only able to make the following limited list of evidence-based recommendations concerning preparation of FFPE tissue from diagnostic or resection specimens for optimal DNA quality:

1. Cold ischaemia times to be limited to 1 hour for FISH analysis and 24 hours for PCR analysis
2. Size of the piece of tissue should be between 3 and 10 cubic millimetres
3. Fixation in neutral buffered formalin for a maximum of 72 hours at either 4°C or room temperature
4. Decalcification if required using EDTA rather than formic or nitric acid
5. Microwave- and ultrasound- assisted fixation does not preclude subsequent extraction and analysis of DNA
6. Embedding in paraffin wax free of beeswax
7. Avoid DNA extraction from tissue stored for greater than 10 years due to reduced length of amplifiable gene fragments with storage over time

One of the problems with interpreting the literature in this area is determining the extent to which the findings for a given downstream application (such as PCR, array CGH, targeted sequencing) are relevant to and can be extrapolated to other applications. This is a particular issue as the move towards larger gene panels and possibly even whole genome massively parallel sequencing gathers pace.

1.2.9 Assessment of nucleic acid yield and integrity from FFPE

There are several methods currently in use for assessing the yield and integrity of nucleic acids extracted from tissue samples. Important parameters are:

- Quantification
- Integrity, e.g. degree of fragmentation
- Ease of amplification

Quantification may be achieved using measurement of optical density ratios by spectrophotometry (e.g. NanoDrop™ 800, Thermo Scientific, Waltham, Massachusetts, USA), although this method is unable to differentiate between DNA and RNA (which both absorb UV at 260nm) and other light-scattering contaminants such as protein, salts or solvents and therefore tends to overestimate the nucleic acid yield by up to five times, with DNA fragmentation and incomplete paraffin removal also contributing to spuriously high readings[95-97]. The newer NanoDrop™ 3300 device employs fluorospectrometry to differentiate double-stranded DNA from other components, using a fluorescent dsDNA-specific nucleic acid dye such as PicoGreen® or Quant-iT™ dsDNA High Sensitivity assay. Other devices using fluoroscopic methods include Qubit® (Life technologies™, Paisley, UK) and Quantus™ Fluorometer (Promega, Wisconsin, US with QuantiFluor® dsDNA system), allowing quantification of DNA distinct from RNA and other substances and give a more accurate estimate of the amount available for sequencing applications. Specific detection of double-stranded DNA (dsDNA) is important since this is the required substrate for sequencing, and inability to differentiate this from single-stranded DNA (ssDNA) present in the mixture also leads to overestimation of the true 'functional yield' of a given sample for downstream sequencing applications.

A further method of DNA quantification is the DNAQuant™ luciferase-pyrophosphorylation coupled DNA Quantitation System (ProMega, Wisconsin, US) in which DNA concentration is determined using three coupled enzymatic reactions to generate a luciferase-dependent light signal related to the amount of ATP produced in the initial reaction, which is directly proportional to the amount of DNA present. The detection system is specific for linear dsDNA, although cross-reaction with dimer/hairpin structures formed by any ssDNA present may also be picked up.

Fragmentation can be assessed using either laboratory-developed multiplex polymerase chain reaction (PCR) 'ladder'-based assays, such as the standard BIOMED-2 assay developed by the EuroClonality consortium for use in lymphoproliferative disease[98], or commercially available capillary electrophoresis methods (e.g. Agilent 2100 Bioanalyzer, Agilent Technology, Palo Alto, California, US). These techniques provide a readout of the range and peak distribution of DNA fragment sizes, in increments of 100 base pairs (bps). RNA quality is conventionally expressed using a RIN, RNA integrity number and a system for establishing an equivalent DIN or DNA integrity number has recently been proposed. This is calculated by comparison with a standardized sample set and ranges from 1 (highly degraded DNA) to 10 (highly intact DNA) with a DIN of equal or greater than 3 generally accepted as suitable for NGS analysis.

Quality assessment of DNA before downstream applications has traditionally been determined using agarose gel electrophoresis of PCR products. This is a time-consuming, laborious and semi-quantitative method of assessment at best and uses toxic reagents such as ethidium bromide. True functional yield for sequencing can be assessed using measures of the dynamics of amplification during quantitative real-time PCR. The Ct value is defined as the 'threshold cycle', i.e. the number of cycles representing the point of intersection of linear and exponential phases of the PCR reaction when plotted on a graph with cycle number on the x-axis and increase in fluorescence over baseline (proportional to amount of PCR product) on the y-axis. The readout from the Asuragen SuraSeq™ DNA QFI™ assay (Asuragen, Austin, US) gives a 'quantitative functional index' as a measure of the fraction of amplifiable DNA relative to the total number of DNA copies and compared to a standard calibration curve[99].

The Agilent Tapestation (Agilent Technologies, US) uses a fluorescent dye that specifically binds to double-stranded DNA. Dye-labelled PCR products are run on an electrophoresis gel, providing information on the concentration of nucleic acid present as well as the sample size distribution of double-stranded DNA within the sample. For small samples in particular, it should be borne in mind that increasingly sophisticated assessment methods are more consumptive of DNA – for Agilent Tapestation a 1µl volume input of genomic

DNA is required. Use of the high-sensitivity D1000 ScreenTape product with the system gives quantification of DNA fragments between 35-1000bp in size and down to 5pg/μl concentration. Also, sequencing methods differ in their sample input requirements in terms of quantity and integrity. The two major approaches to library preparation for targeted next generation, massively parallel sequencing are amplicon or hybridisation based. Amplicon-based approaches use PCR amplification for library generation, which is quick but risks inaccuracy of the sequencing output through polymerase replication errors, formation of secondary structures such as dimmers or hairpins, and preferential amplification of certain sequences dependent on the GC nucleotide content. Hybridisation based approaches avoid these issues and can target larger genomic regions, although these are more labour-intensive and require greater amounts of input DNA.

1.2.10 Alternative tissue fixatives

Although formalin has been the tissue fixative of choice in pathology laboratories for over a century, in recent years there have been attempts to develop a fixative that provides superior preservation of nucleic acids, with minimal effects on the tissue and cellular morphology (and attendant artefacts) that histopathologists are trained to recognize and rely on for diagnosis. A further reason for a move away from formalin is evidence of carcinogenicity. Formalin has long been recognized as an irritant of mucosal membranes of the conjunctiva and respiratory tract, and also as an allergen of skin and respiratory tract. In 2006, a working group for the International Agency for Research against Cancer (IARC) officially classified formalin as a human carcinogen after a statistically significant increased risk of death from nasopharyngeal cancer and leukaemia was found on meta-analysis of study data[100].

Any alternative fixative must also be compatible with existing laboratory protocols for processing and staining, including the mainstay of haematoxylin and eosin staining but also special histochemical stains, immunohistochemistry and in situ hybridisation. Formalin, as an aldehyde fixative, fixes tissue by forming chemical additions and cross-links within tissue. The most common alternative fixatives are alcohol-based and exert

their effect by subtracting water from tissue and coagulating proteins. Table 9 provides details of alternative fixatives, with a summary of published data for their utility in diagnostic pathology in table 10.

Table 9. Tissue fixatives developed as an alternative to formalin for use in diagnostic cellular pathology and molecular applications

This table has been compiled from information available on the manufacturers' websites.

Fixative base	Examples	Applications and notes
Ethanol	FineFIX (Milestone Medical, Bergamo, Italy) NeoFix (Merck, Darmstadt, Germany) PAXgene® tissue system (Qiagen, Hilden, Germany) RCL2 (Alphelys, Plaisir, France)	Non-cross linking fixatives with published evidence for superior preservation of nucleic acids with acceptable morphology
	Carnoy's: ethanol, chloroform and glacial acetic acid	Lyse erythrocytes Can produce tissue hardening and shrinkage
Methanol	Methacarn: methanol, chloroform and glacial acetic acid Modified methacarn: methanol and glacial acetic acid	
	UMFix, marketed as TissueTek® Xpress® Molecular Fixative (Sakura FineTek, Torrance, United States): 90% methanol + 10% polyethylene glycol	Developed specifically for use with microwave-assisted rapid tissue processing
Glyoxal	Cell-Block (Bio-Optica, Milan, Italy) ExCell™ Plus (American MasterTech, California, United States) GreenFix (Diapath, Bergamo, Italy) GTF™ formalin substitute (StatLab, Texas, United States) Histochoice, (Amresco®, Ohio, United States) Mirsky's Fixative (National Diagnostics, Georgia, United States) Prefer (Anatech Ltd, Michigan, United States) Preserve™ (Energy Beam Sciences, Connecticut, United States) SafeFix II (Fisher Scientific, Massachusetts, United States) Shandon Glyo-Fixx™ (Thermo Scientific, Massachusetts, United States)	Cross-linking aldehyde fixatives Virtually no vapours at room temperature therefore marketed as safer alternatives to formalin
Acetone	AMeX method: fixation at -20°C overnight in acetone followed by clearing in methylbenzoate and xylene HOPE (Hepes-glutamic acid buffer-mediated organic solvent protection effect, DCS, Hamburg, Germany): Fresh tissues incubated in 'protecting solution' composed of amino acid mix, then dehydrated in acetone at 4°C	Acetone component is volatile and flammable and can cause tissue to become brittle
Picric acid	Bouin's: picric acid, formaldehyde and glacial acetic acid Hollande's: picric acid, copper acetate, formaldehyde and acetic acid	Lyse erythrocytes and remove small amounts of iron and calcium Dry picric acid is explosive Degrade nucleic acids
Mercuric chloride	B5: mercuric chloride, sodium acetate and formalin Zenker's: mercuric chloride, potassium dichromate and glacial acetic acid	Good nuclear preservation but lyse erythrocytes, past use for haematolymphoid pathology Mercury pigment forms and needs removal during slide preparation Mercury is corrosive and toxic
Zinc	ZBF: zinc-based fixative (zinc acetate and zinc chloride in Tris buffer) Z2: zinc acetate, zinc chloride and calcium chloride in Tris buffer Z7: zinc trifluoroacetate, zinc chloride and calcium acetate	Non-toxic, non-carcinogenic, thermostable and inexpensive Tissue shrinkage commonly seen Inferior tissue penetration compared to formalin.
Other	HistoFix (Richard-Allan Scientific, Michigan, United States): pyrrolid-2-one, a polyol, a urea and a zinc salt NOTOXhisto (Scientific Device Laboratory, Des Plaines, United States): complex aldehyde in 70% alcohol with antiseptic and antifungal agents	

Table 10. Published evidence for utility of selected alternative tissue fixatives for morphology and molecular analyses

Abbreviations used in table: aCGH = array comparative genomic hybridisation; *ALK* = anaplastic lymphoma kinase gene; bp = base pairs (length of DNA fragments); CISH = chromagenic *in situ* hybridisation; Ct = cycle threshold for polymerase chain reaction; ER = oestrogen receptor; FF = fresh frozen tissue at -80°C; FFPE = formalin-fixed, paraffin embedded tissue; FISH = fluorescent *in situ* hybridisation; HER2 = human epidermal growth factor receptor 2; HRM = high resolution melt; IHC = immunohistochemistry; ISH = *in situ* hybridisation; MF = molecular fixative; PAGA = polyethylene glycol, ethyl alcohol, glycerol and acetic acid; PCR = polymerase chain reaction; PFPE = PAXgene® Tissue fixed paraffin-embedded tissue; PR = progesterone receptor; q(RT-)PCR = quantitative (real-time) polymerase chain reaction; RFPE = RCL2 fixed paraffin embedded tissue; RIN = RNA integrity number; rt-PCR = reverse transcription polymerase chain reaction; ZBF: Zinc-based fixative.

a. **RCL2 fixation and paraffin embedding (RFPE)**

Number and type of samples	Comparator groups	Alternative fixative performance with different techniques				Reference
		Morphology	Immunohistochemistry	Nucleic acid integrity	*In situ* hybridisation	
22 pulmonary adenocarcinoma or squamous cell carcinoma from human clinical samples Human cell line pellets for FISH	FF FFPE	RFPE gave similar morphology to FFPE with better nuclear detail (qualitative assessment)	Less intense staining with RFPE than FFPE; resolved by minor protocol modification	RFPE and FF yielded DNA amplicons 100-600bp in size; more fragmented DNA seen with FFPE	*ALK* FISH on cell line pellets gave identical results with no adjustment of protocol	Khellaf L, Larrieux M, Serre I et al. Morphological and molecular analysis of lung cancer biopsies fixed in RCL2. *Histopathology* 2013, **63:** 137-9
49 samples from 36 fresh specimens: benign ovarian, fallopian tube, uterine, thyroid, tonsil, breast and placental tissue	FFPE	Tissue and retraction seen in RFPE, in stained sections and also in blocks, where hardening and friability made microtomy more difficult. Better representation of nuclear features in RFPE	Of 18 different antibodies assessed, β-hCG showed strong background staining in RFPE, pan-cytokeratin and progesterone receptor were heterogeneous between tissue types.	4-7 fold higher DNA yields from RFPE vs FFPE, with similar acceptable optical density ratios for purity	Fewer BCL2 cellular signals seen in RFPE tissue compared to FFPE. SISH HER2 cleaner (less background artefact) and more intense signals in RFPE than FFPE	Masir N, Ghoddoosi M, Mansor S, et al. RCL2, a potential formalin substitute for tissue fixation in routine pathological specimens. *Histopathology* 2012;**60:** 804–15

11 breast and 12 colon tumour samples	FF FFPE	Good preservation of cytology and architecture in RFPE suitable for routine diagnosis	Protocol modification necessary for CK20 antibody: different buffer at higher pH required for antigen retrieval	DNA of up to 523bp in length amplified, suitable for aCGH, *KRAS* genotyping and microsatellite analysis. RNA integrity slightly less than FF but still suitable for rt-PCR and qRT-PCR.	HER2 CISH successful; one non-clinically significant discrepant result (score 0 for FFPE and 1+ for RFPE)	Boissière-Michot F, Denouël A, Boulle N et al. The non-crosslinking fixative RCL2®-CS100 is compatible with both pathology diagnosis and molecular analyses. *Pathol Oncol Res.* 2013 Jan; **19(1):** 41-53.
MCF-7 breast cancer cell line culture (for initial RNA assessment) and tissue samples from 6 breast cancer resection specimens	FF FFPE Methacarn	All fixatives gave comparable and acceptable morphology	Antigen retrieval protocol modification required for methacarn and RCL2-fixed tissues, giving same result for ER, PR and HER2 to FFPE material	Intact genomic DNA obtained from methacarn and RCL2-fixed tissue (up to 850bp products amplified) compared to degraded FFPE derived DNA. Intact RNA obtained from methacarn and RCL2-fixed tissue on real-time rt-PCR; degraded from FFPE.	*HER2* gene amplification could be confirmed on CISH for all HER2 immunohistochemistry score 3+ tumours, irrespective of fixative	Delfour C, Roger P, Bret C, et al. RCL2, a new fixative, preserves morphology and nucleic acid integrity in paraffin-embedded breast carcinoma and microdissected breast tumor cells. *J Mol Diagn* 2006; **8:** 157–69.

b. PAXgene® Tissue fixation with formalin-free processing and paraffin embedding (PFPE)

Number and type of samples	**Comparator groups**	**Alternative fixative performance with different techniques**				**Reference**
		Morphology	**Immunohistochemistry**	**Nucleic acid integrity**	***In situ* hybridisation**	
Three rat tissue samples from each of eight different organs	FF FFPE	All comparable – increased nuclear staining intensity and cytoplasmic eosinophilia noticed in PFPE in liver, heart and brain tissue.		RINs of 4.0-7.2 for FFPE; 6.4-7.7 for PFPE and 8.0-9.2 for FF. FFPE-derived RNA showed slower migration and less intense ribosomal peaks on electrophoresis and		Groelz D, Sobin L, Branton P et al. Non-formalin fixative versus formalin-fixed tissue: A comparison of histology and RNA quality. *Exper and Molec Pathol* 2013; **94**: 188-94.

		Erythrocyte lysis also evident.		performed less well than FF and PFPE in RT-PCR.		
70 samples of 4mm thick pieces of human tissue of multiple types including tumour and non-tumour	FF FFPE	Liver, thyroid, adrenal and skeletal muscle: more contrast and better definition. Increased eosinophilia with difficulty of lineage recognition in bone marrow. Inferior morphology in lung, stomach, prostate, spleen and small intestine. Seminoma tissue: dissociation of cells. Erythrocytes appeared lysed and empty.	Decreased immune reactivity on PFPE with several antigens that could be overcome by changes to protocol such as omission of antigen retrieval, adjusting antibody concentration and using an alternative clone	Not assessed	Stronger HER2 ISH signal in PFPE breast cancer tissue causing no problems in detection of amplification	Kap M, Smedts F, Oosterhuis W et al. Histological assessment of PAXgene® tissue fixation and stabilisation reagents. *PLoS One* 2011; **6**: e27704.
12 clinical melanoma biopsies from human subjects	FF FFPE	PFPE comparable to FFPE	5/11 antibodies showed statistically significant less intense staining, requiring addition of detergent to protocol to increase membrane permeability in tissue	Larger mRNA amplicon sizes with PFPE and less DNA fragmentation observed when compared to FFPE		Belloni B, Lambertini C, Nuciforo P et al. Will PAXgene® substitute formalin? A morphological and molecular comparative study using a new fixative system. *J Clin Pathol* 2013; **66**: 124-35.
Four benign prostate glands obtained post mortem and four radical	FF FFPE	PFPE comparable to FFPE with increased eosinophilia of staining in PFPE	Comparable expression of p63, PSA and P504S between PFPE and FFPE	DNA of comparable quality from PFPE and FF tissue. Extracted DNA fragments ~25% longer from PFPE than from FFPE. RNA	Not performed	Gillard M, Tom WR, Anitic T et al. Next-gen tissue: preservation of molecular and morphological fidelity in prostate tissue. *Am J Transl Res* 2015; **7:** 1227-35.

prostatectomy specimens from patients with prostate cancer				amplified with 8-fold efficiency from PFPE tissue compared to FFPE.		
Rat kidney and human tissue: benign breast, stomach, liver, adipose, intestine, kidney, spleen and leiomyosarcoma, cholangiocarcinoma and colorectal adenocarcinoma	FF FFPE	Similar morphology between PFPE and FFPE	Less requirement for antigen retrieval techniques with PFPE than FFPE	PFPE extracted RNA yield and qRT-PCR performance similar to that of FF tissue and superior to FFPE. PFPE extracted DNA similar molecular mass and performance on multiplex PCR, Sanger and pyrosequencing reactions to FF tissue	Not assessed	Viertler C, Groelz D, Gündisch S et al. A new technology for stabilization of biomolecules in tissues for combined histological and molecular analyses. *J Mol Diagn* 2012; **14**: 458-66
13 human breast cancer clinical samples	FFPE	Not described	Not performed	Not performed	Weak centromeric (CEN17)/ HER2 signals in PFPE tissue but could be restored to be comparable with FFPE by post-fixation of slide-mounted sections in formalin for 18-24 hours following deparaffinisation	Oberauner-Wappis L, Loibner M, Viertler C et al. Protocol for HER2 FISH determination on PAXgene®-fixed and paraffin-embedded tissue in breast cancer. *Int J Exp Path* 2016; **97**: 202–206.

c. UMFix/ MF (molecular fixative)

Number and type of samples	Comparator groups	Alternative fixative performance with different techniques				Reference
		Morphology	Immunohistochemistry	Nucleic acid integrity	*In situ* hybridisation	
Mouse liver tissue and human tissue (adrenal, breast, colon, eye, oesophagus, kidney, liver, lung, lymph node, skeletal muscle, pancreas, parathyroid, parotid, prostate, skin, small intestine, soft tissue, spleen, thyroid, tonsil, uterus)	FF FFPE	Comparable morphology between FFPE and UMFix tissue. Moderate erythrocyte swelling in tissues left in UMFix for more than 48 hours	More intense staining in UMFix tissue with 27 different antibodies, except hepatocellular marker HepPar 1 which showed less strong staining in UMFix tissue	No significant differences between UMFIX and FF tissues on PCR, rt-PCR, qRT-PCR, or expression microarrays. Higher cycle threshold values for FFPE compared to FF or UMFix tissues, increasing with length of time in formalin	Not assessed	Vincek V, Nassiri M, Nadji M et al. A tissue fixative that protects macromolecules (DNA, RNA, and protein) and histomorphology in clinical samples. *Lab Invest* 2003; **83**: 1427–35.
Benign human colonic, uterine myometrial and liver tissue from three specimens	FF FFPE	MF-fixed tissue showed increased eosinophilia and swelling of erythrocytes.	Not assessed	MF-fixed tissue with rapid processing and fixation periods of up to 7 days had higher success rates on PCR and rt-PCR than FFPE, similar to those achieved with FF tissue.	Not assessed	Turashvili G1, Yang W, McKinney S et al. Nucleic acid quantity and quality from paraffin blocks: defining optimal fixation, processing and DNA/RNA extraction techniques. Exp Mol Pathol. 2012 Feb; 92(1):33-43.
Block and biopsy-sized samples from 16 human high-grade serous ovarian cancer specimens	FF FFPE	Comparable morphology between FFPE and UMFix tissue	Additional optimisation required for WT1 in UMFix tissue, expression similar for all antibodies	Greater DNA yield, longer fragment size and more accurate copy-number calling using shallow whole-genome sequencing (WGS) in UMFix tissue	Not assessed	Piskorz AM, Ennis D, Macintyre G et al. Methanol-based fixation is superior to buffered formalin for next-generation sequencing of DNA from clinical cancer samples. Annals of Oncology 27: 532–539, 2016

d. FineFIX

Number and type of samples	Comparator groups	Alternative fixative performance with different techniques				Reference
		Morphology	**Immunohistochemistry**	**Nucleic acid integrity**	***In situ* hybridisation**	
Cell block samples derived from 15 effusions and 38 fine needle aspirates of lung, liver, breast, lymph node, thyroid and subcutaneous lesions	FFPE cell blocks Fresh unfixed cell suspension	Similar morphology in FineFIX and formalin preparations. Moderate increase in erythrocyte size in samples fixed in FineFIX for >48 hours	Similar immunoreactivity apart from higher sensitivity of FineFIX material with vimentin	Degraded DNA obtained from FFPE cell block –smear at lower molecular weights on gel and maximum 199bp amplifiable on PCR, compared to amplification of all fragments up to 2361bp in FineFIX extracted DNA. Only DNA derived from FineFIX also suitable for *EGFR* mutation screening using HRM. RNA from FineFIX (RIN 6.3) closer to that of fresh material (RIN 9.6) than FFPE (RIN 2.1).	Not assessed	Gazziero A, Guzzardo V, Aldighieri E and Fassina A. Morphological quality and nucleic acid preservation in cytopathology. J Clin Pathol. 2009 May; 62(5):429-34.
Tissue samples from 5 colon cancer specimens	FFPE FF	Comparable and acceptable morphology on immunostained slides.	More intense staining noted with KL-1 keratin antibody in FineFix tissue	Yields of DNA and RNA from FineFix tissue approximately twice that from FFPE tissue, with amplification of >1000bp lengths from FineFix tissue only.	Not performed	Stanta G, Pozzi Mucelli S, Petrera F et al. A Novel Fixative Improves Opportunities of Nucleic Acids and Proteomic Analysis in Human Archive's Tissue. *Diagn Mol Pathol* 2006; **15:** 115-23.

e. Z7

Number and type of samples	Comparator groups	Alternative fixative performance with different techniques				Reference
		Morphology	Immunohistochemistry	Nucleic acid integrity	*In situ* hybridisation	
Mouse tissue from multiple organs	FFPE FF Z2 HOPE	Superior morphology on FFPE material.	Z7 best fixative for IHC without need for antigen retrieval	Fragments of DNA 2.4kb in length and RNA up to 361bp in length obtained from Z7 fixed tissue and performed well in PCR-based applications.	Not performed	Lykidis D, Van Noorden S, Armstrong A et al. Novel zinc-based fixative for high quality DNA, RNA and protein analysis. Nucleic Acids Res. 2007; 35(12):e85. Epub 2007 Jun 18.

f. **Hollande, B5, Bouin, Zenker**

Number and type of samples	Comparator groups	Alternative fixative performance with different techniques				Reference
		Morphology	Immunohistochemistry	Nucleic acid integrity	*In situ* hybridisation	
Punch biopsies of normal human colon tissue from three surgical resection specimens	FFPE 70% ethanol fixed tissue	Not assessed	Not assessed	All samples yielded sufficient DNA for PCR amplification of β-globin gene. DNA fragments predominantly >100bp from ethanol and FFPE; <100bp from B5 and Hollande	Zenker, B5 and Bouin fixed tissues not suitable for HER2 FISH analysis	Baloglu G, Haholu A, Kucukodaci Z et al. The effects of tissue fixation alternatives on DNA content: a study on normal colon tissue. Appl Immunohistochem Mol Morphol 2008; 16: 485–492

g. **FineFix and RCL2 with Peloris rapid tissue processing**

Number and type of samples	Comparator groups	Alternative fixative performance with different techniques				Reference
		Morphology	Immunohistochemistry	Nucleic acid integrity	*In situ* hybridisation	
Multiple human tissue types: adrenal, breast, brain, colon, gallbladder, kidney, liver, lung, lymph node, oesophagus, placenta, small intestine, soft tissue, spleen, stomach, thyroid, tonsil.	FFPE	Partial tissue disintegration and cellular degranulation of granulocytes, mast cells and intestinal Paneth cells with FineFix. Soft, slippery tissue that was difficult to cut and pigment deposition in bloody tissues with RCL2. Both caused shrinkage artefacts and erythrocyte lysis.	Inadequate ER and sub-optimal S100 staining despite protocol modification with both RCL2 and FineFix	Higher DNA and RNA yield and quality from FineFix and RCL2 tissue. Concordant results on *EGFR* mutation and microsatellite instability analysis.	Concordant results on CISH and FISH for *CEP15*, *BCL6* and *BCL2*. FineFix showed brightest signals and least background.	Moelans CB, Oostenrijk D, Moons MJ, et al. Formaldehyde substitute fixatives: effects on nucleic acid preservation. *J Clin Pathol* 2011; **64**: 960–7 and Moelans CB, Hoeve N, van Ginkel JW, et al. Formaldehyde substitute fixatives. Analysis of macroscopy, morphologic analysis, and immunohistochemical analysis. *Am J Clin Pathol* 2011; **136:** 548–56.

h. **RCL2, Z7, PAXgene®, RNAlater or Allprotect**

Number and type of samples	Comparator groups	Alternative fixative performance with different techniques				Reference
		Morphology	Immunohistochemistry	Nucleic acid integrity	*In situ* hybridisation	
21 clinical samples of benign human ovarian, uterine or fallopian tube tissue	FF FFPE	Similar and acceptable for PAXgene®, RCL2 and Z7; poor for Allprotect and RNAlater (both excluded from IHC/FISH analyses)	Similar and acceptable for PAXgene®, RCL2 and Z7 with no modification of standard FFPE protocol	FF and PFPE performed better than other samples with RIN and Ct for RT-PCR. No statistically significant difference in Ct for DNA	Similar and acceptable for PAXgene®, RCL2 and Z7 with no modification of standard FFPE protocol	Staff S, Kujala P, Karhu R et al. Preservation of nucleic acids and tissue morphology in paraffin-embedded clinical samples: comparison of five molecular fixatives. *J Clin Pathol* 2013; **66**: 807-810.

i. **RCL2, PAGA, ZBF, Z7, CellBlock, FineFix, Carnoy and B5**

Number and type of samples	Comparator groups	Alternative fixative performance with different techniques				Reference
		Morphology	Immunohistochemistry	Nucleic acid integrity	*In situ* hybridisation	
200 samples from surplus material in routine surgical pathology and autopsy cases	FFPE	Alcohol-based fixatives showed higher stain affinity and tissue shrinkage but superior preservation of nuclear features Shrinkage also seen with zinc-based fixatives	Protocols already optimised in laboratory for routine use of FineFixx with Carnoy's for neuropathology	Highest RINs for NeoFixx and RCL2. Cell-Block same as formalin and PAGA inferior to formalin All samples suitable for qRT-PCR	Not performed	Zanini C, Gerbaudo E, Ercole E et al. Evaluation of two commercial and three home-made fixatives for the substitution of formalin: a formaldehyde–free laboratory is possible. *Environ Health*. 2012; 11: 59.

The PAXgene® Tissue fixative and stabilisation system (PreanalytiX™, a collaboration between Qiagen, Hilden, Germany and BD Biosciences, Erembodegen, Belgium) is the most extensively studied novel fixative in the past few years, with much of the initial work taking place in European academic institutions as part of the European Union Seventh Framework Project (FP7)-funded SPIDIA (***S***tandardisation and Improvement of Generic ***P***re-analytical Tools and Procedures for ***I***n Vitro ***DIA***gnostics) consortium project to improve pre-analytical procedures. Funding has recently been granted by Innovate UK (formerly the Technology Strategy Board) for a collaborative project involving NHS and academic institutions in partnership with Qiagen to investigate its application to diagnostic biopsy, fine needle aspiration cytology and circulating cell-free DNA samples, ongoing work that forms part of this thesis. PAXgene® Tissue is a non-cross-linking, non-carcinogenic fixation reagent containing a mixture of different alcohols, acid and other soluble organic compounds. Immersion in the fixation reagent is followed by use of the stabilisation reagent, a storage and transport medium consisting of a mixture of alcohols.

1.2.11 Tissue processing

Modern tissue processing takes place in automated tissue processors. Following an initial step to complete fixation, dehydration of the tissue occurs using ethanol solutions of different concentrations. The next step involves clearing using xylene and finally paraffin impregnation. The main drawbacks of this approach are the time taken, since most laboratories use an overnight run for standard tissue processing and therefore sections cannot be cut until the following day, and the large volumes and toxicity of the reagents used. Tissue processing also contributes to nucleic acid degradation[76] and there is a risk of cross-contamination of tissue (and therefore DNA) between different samples in the cassette. The phenomenon of tissue 'carryover' is familiar to all pathologists, where an extraneous tissue fragment is included in a section in which it clearly does not belong, due to transfer into the cassette and incorporation in the block at some point during sample preparation.

In recent years, there has been a move towards xylene-free processing in some laboratories in an attempt to minimise the use of toxic reagents. A further

technical development has been rapid tissue processing, in which high temperatures are used to reduce the processing time. However, this may not be optimal for samples requiring downstream molecular analysis since high temperatures contribute to hydrolysis of nucleic acids. In the presence of buffers, heat can also be used to reverse nucleic acid monomethylol group additions induced by formalin. Other approaches to accelerating tissue processing are the application of energy in the form of either microwave or ultrasound, but these have also been shown to have detrimental effects on subsequent histologic and molecular analyses[101].

A recently proposed solution to these issues is the '2+2' fixation/processing schedule proposed by Chafin et al (Ventana Medical Systems, Tucson, Arizona, US) with collaborators from the University of Washington, US[101]. The method involves fixing tissue in formalin for 2 hours at 4°C followed by 2 hours at 45°C. This method provided acceptable tissue morphology, comparable to that of material fixed for longer periods (up to 24 hours) in formalin, and similar results on immunohistochemistry and in situ hybridisation. This protocol overcame the problem of the central part of the tissue becoming disrupted before adequate fixation could occur, as is often the case with high temperature fixation, and the authors propose that the initial lower temperature allowed formalin penetration right into the centre of the tissue, due to the relative paucity of cross-linking reactions occurring during this period of low temperature fixation. The paper does not report any details of nucleic acid extraction or quality assessment resulting from this method.

1.2.12 Fresh frozen tissue as an alternative to FFPE

Extraction of nucleic acids from fresh frozen tissue is an attractive alternative to fixed tissue since it avoids the chemical fixation and processing steps, thus yielding superior quality nucleic acids without the fragmentation and chemical modifications characteristic of FFPE material. Since the workflow and processes in cellular pathology laboratories are currently configured to accommodate formalin-immersed tissue, there are significant logistical challenges associated with the provision of fresh tissue. In particular there is a requirement to have staff ready to deliver the sample from theatre to the pathology laboratory, as well as having appropriately trained personnel available to receive, book in and

identify the sample as requiring immediate attention. One possible solution that has been proposed to this is holding the specimen in a fresh state under vacuum-packing in a refrigerator at 4°C, which has been trialled and subsequently incorporated into routine practice by pathologists in several institutions in Italy[102] as well as forming part of the ongoing experimental pathway for the 100,000 genomes project.

If sampling from a resection specimen is required, it must be borne in mind that the resection specimen is primarily intended for diagnosis, staging and prognostication of the tumour and it is critically important that fresh specimen sampling does not compromise this assessment. This is particularly relevant to assessment of margins or surface involvement/capsular breach which may be damaged by incision of the fresh specimen, and also the need to preserve enough tumour to make a diagnosis, grade the tumour accurately and identify important prognostic features such as lymphovascular or perineural invasion which may be only focally present. According to current cellular pathology guidelines, tumours smaller than 20mm in maximum diameter (the T1/T2 boundary in many solid tumour types in the TNM classification system[103]) are required to be submitted in their entirety for histopathological analysis and are therefore unlikely to contain sufficient tumour for fresh sampling under direct visualisation, although one or more core biopsies may carefully be taken from the fresh specimen as an alternative.

A further important step that must not be omitted in fresh tissue sampling is the morphological assessment of the sample. This is required in order to confirm that tumour is present, determine the ratio of tumour to non-tumour nuclei and also the relative quantity, presence or absence of other substances that might inhibit PCR, such as melanin or necrosis.

1.3 Summary

In this chapter recent advances in knowledge of the molecular pathogenesis of cancer have been reviewed and used to illustrate how this is being applied to solid tumour samples. This is leading to refined characterisation of tumours beyond what is possible through traditional morphological assessment, in order to inform the development of new therapies and identify patients who are likely to show objective clinical response to selected treatments. A further implication of this is the evolution of a new taxonomy of cancer based on classification according to molecular aberrations rather than organ of origin[104]. This approach requires a significant departure from the processes and methods established in cellular pathology during the past century, and there are clinical and logistical challenges to be addressed and overcome in making the transition to the delivery of stratified medicine as part of routine clinical care in the health service. The different options for acquiring tissue for molecular analysis are outlined in figure 2 in comparison to the current standard of care. In the following chapters analysis of the feasibility of this approach is presented, with a focus on areas requiring evidence and the development of new skills and expertise in the future workforce.

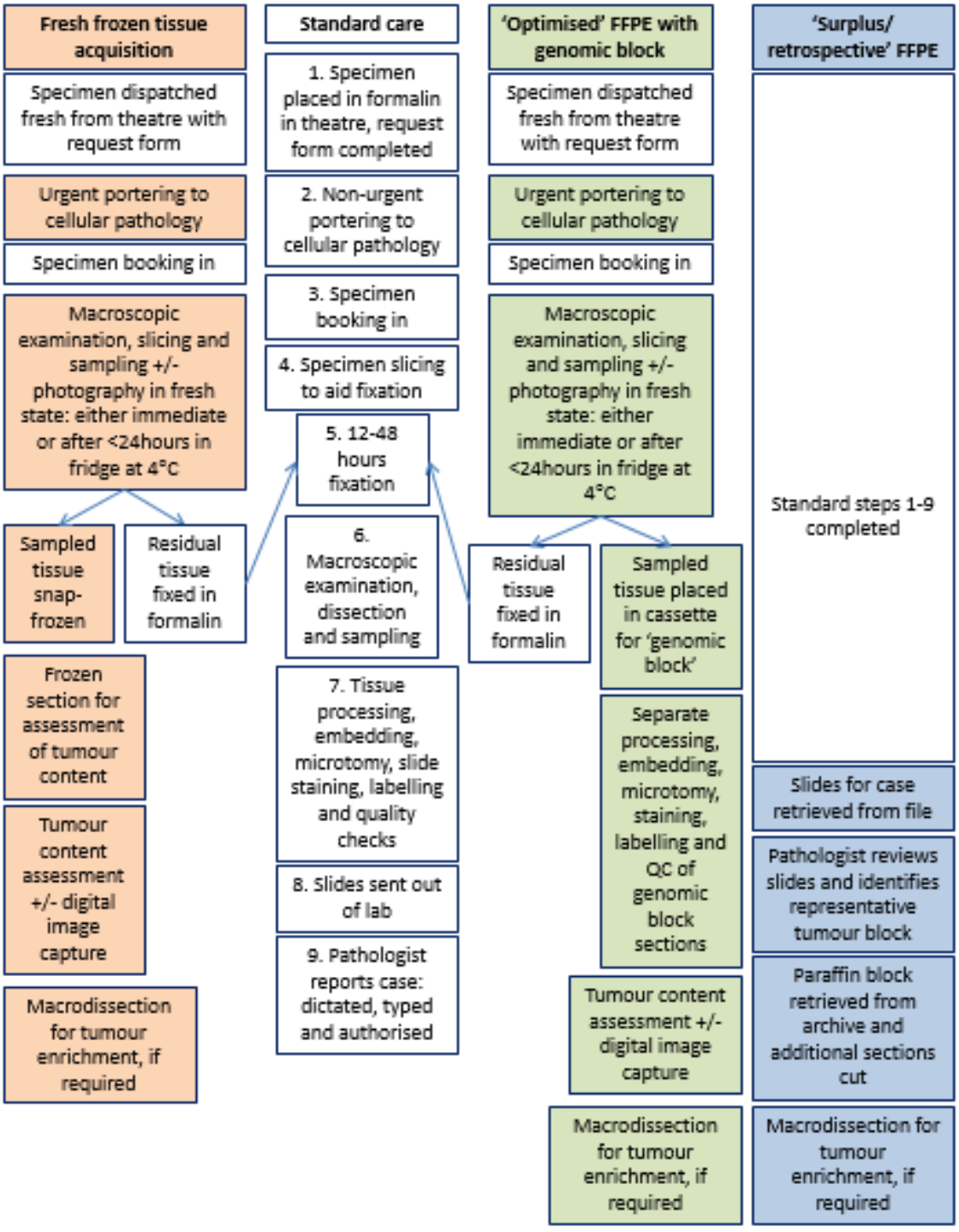

Figure 2. Options for tissue acquisition for molecular analysis from a resection specimen

The shaded boxes highlight extra steps and processes over and above standard care for each alternative pathway. FFPE = formalin-fixed paraffin-embedded (tissue).

Chapter 2: **Methods**

2.1 Phase One of the Cancer Research UK Stratified Medicine Programme

During CRUK SMP1, patients with selected tumour types (breast, colorectal, prostate, lung or ovarian cancer, or advanced malignant melanoma), were approached through one of 26 hospitals forming part of a network of clinical hubs (CHs) coordinated through Cancer Research UK-funded Experimental Cancer Medicine Centres. Consent was sought for centralised molecular testing of surplus material from resections or biopsies of tumour tissue from a primary or metastatic site, performed as part of routine clinical care.

Formalin-fixed paraffin-embedded tissue sections were forwarded with a peripheral blood sample from each patient to one of three technology hubs (THs) for analysis of a small panel of abnormalities determined according to primary tumour type. The range of molecular aberrations detected included point mutations by PCR-based sequencing methods, gene rearrangements by fluorescent in situ hybridisation and loss of heterozygosity by microsatellite analysis (tables 11 and 12).

Table 11. Prioritised genes for each tumour type included in SMP1

Tumour	Genetic aberrations sought
Colorectal carcinoma	Mutations in *KRAS, BRAF, NRAS, PIK3CA* and *TP53*
Breast carcinoma	Mutations in *PIK3CA, TP53, BRAF* and *PTEN* Loss of heterozygosity in *PTEN* by microsatellite analysis
Prostate carcinoma	Mutations in *BRAF* and *PTEN* mutation Loss of heterozygosity in *PTEN* LOH by microsatellite analysis *TMPRSS2-ERG* fusion by fluorescent in situ hybridisation
Lung carcinoma	Mutations in *EGFR, KRAS* and *BRAF* Mutations in *DDR2* (squamous cell carcinoma subtype only) *ALK* rearrangement by fluorescent in situ hybridisation
Ovarian carcinoma	Mutations in *PIK3CA, TP53, BRAF* and *PTEN* Loss of heterozygosity in *PTEN* by microsatellite analysis
Malignant melanoma	Mutations in *BRAF, KIT, NRAS* and *PIK3CA*

Table 12. Genetic regions of interest and techniques used for analysis during SMP1

Gene	Scope of test	TH1	TH2	TH3
BRAF	Initially codons 599-601 in exon 15, then extended to encompass all of exons 11 and 15 from September 2012	Qiagen TheraScreen Pyromark BRAF Assay, in-house pyrosequencing assay, Roche Cobas® 4800 system, direct sequencing as required	In-house pyrosequencing assay	CE-SSCA +/- subsequent sequencing if required
		All do exon 11 by sequencing		
DDR2	Exons 3-18; added in September 2012	Sequencing		
EGFR	Exons 18-21	Qiagen TheraScreen Pyromark EGFR Assay	Pyrosequencing or fragment length analysis	CE-SSCA
EML4-ALK	Confirm presence of breakpoint in *ALK* gene	FISH (break apart probe)		
KIT	Exons 11, 13 and 17	Sequencing	Sequencing	CE-SSCA
KRAS	Codons 12, 13, 61 and 146	Qiagen TheraScreen Pyromark KRAS Assay; Qiagen KRAS TheraScreen Assay	Pyrosequencing	Cobas 4800
NRAS	Codons 12, 13, 61	Qiagen TheraScreen Pyromark NRAS Assay	Pyrosequencing	CE-SSCA
PIK3CA	Exons 9 and 20	Pyrosequencing, Snapshot or Qiagen ARMS kit	Pyrosequencing	CE-SSCA
PTEN LOH	Confirm loss of heterozygosity	Same protocol – microsatellite analysis using three different STR markers		
PTEN mutation	Exons 2-10	Sequencing	HRM with sequencing of variants	CE-SSCA
TMPRSS2-ERG	Confirm presence of rearrangement	FISH (triple colour probe)		
TP53	Exons 4-9	Sequencing	Sequencing	CE-SSCA

CE-SSCA: capillary electrophoresis-single strand conformation analysis; FISH: fluorescent in situ hybridisation; HRM: high resolution melt analysis; STR: short tandem repeat; TH1-3: technology hub 1, 2 or 3

Results were transmitted electronically to clinical centres for inclusion in medical records. A clinical dataset including diagnostic, treatment and outcome data was collated for all patients and submitted to the lead national cancer registry for England. This clinical dataset was based as far as possible on existing NHS information standards, including attributes drawn from the enhanced cancer registration dataset for England, the Cancer Outcomes and Services Dataset, COSD. This included data item definitions from the NHS data dictionary and used accepted coding systems such as ICD-10 (10th revision of the World Health Organization International Statistical Classification of Diseases and Related Health Problems and SNOMED-RT (Systematized Nomenclature Of Medicine Clinical Terms Reference Terminology).

Death registration data including the date and cause of death for patients who died during the course of the initiative was requested at six monthly intervals by the National Cancer Registration Service from the Office of National Statistics. Following removal of demographic data items, an extract of collated clinical, pathological, treatment and outcomes data for SMP patients was sent to the Department of Computing Sciences at the University of Oxford and incorporated into a single Microsoft Access database. This database was composed of a number of linked tables according to a data model.

SMP1 provided an excellent opportunity to assess the feasibility and acceptability of genetic analysis of tumour samples in the UK population. The eligibility criteria (table 13) were deliberately designed to be broad and inclusive to maximise the relevance of the findings to the broader UK population.

Table 13. SMP1 patient eligibility criteria

- Adult patient aged 18 years or more
- Able to give informed consent
- Diagnosis of one of the following types of invasive malignancy:
    - Carcinoma of the breast including ductal, lobular and other subtypes
    - Adenocarcinoma of the colon or rectum
    - Carcinoma of the lung including both small cell and non-small cell subtypes but excluding malignant mesothelioma and carcinoid tumours
    - Advanced malignant melanoma *(i.e. stage III or IV disease with at least regional lymph node involvement),* cutaneous as well as less common mucosal sites
    - Carcinoma of the ovary
    - Carcinoma of the prostate
- Formalin-fixed paraffin-embedded (FFPE) tissue sample from resection or biopsy specimen with surplus tissue available, from either the primary tumour or a site of metastasis

For the purpose of this analysis, an extract of a subset of data items in an Excel spreadsheet was requested from the central Microsoft Access database held at the Department of Computer Sciences at the University of Oxford. This comprised the following data items:

1. Anonymized unique patient identifier
2. ICD code from the diagnosis table
3. ICD code from the pathology table
4. SNOMED morphology code from the diagnosis table with decoded textual meaning from an associated look-up table
5. SNOMED morphology code from the pathology table with decoded textual meaning from an associated look-up table
6. Gene
7. Test status
8. Test result
9. Source sample identifier

The raw data extract comprised a spreadsheet of 33,639 rows excluding the table header and representing records for 7,813 unique patients. Due to the structure of the data with a one-to-many relationship between a patient and their multiple diagnostic codes and genetic test results, cross-products were formed in the exported data whereby multiple rows of data existed for each patient. In order to allow aggregated analysis through Excel, it was necessary to condense the results to a single row per patient. Firstly using the anatomical site ICD codes, the data was sorted into multiple spreadsheet tabs as follows:

- One for each of the 6 disease-based cohorts: breast, colorectal, lung, ovarian and prostate cancer and malignant melanoma
- One for all queries (cases that could not be allocated to a disease-based cohort using the ICD code, SNOMED code or genetic data)
- One for patients with ineligible diagnoses (e.g. malignant mesothelioma of the pleura, adenocarcinoma of the small intestine)

Following this step, the data was manually condensed into one row per patient. A minority of patients had more than one tumour sample submitted for analysis, and these pairs of samples were removed to a separate spreadsheet per tumour type for separate analysis, preserving the sample with the earliest sequential alphanumeric source sample identifier in the main spreadsheet. The manual condensation of data involved rationalising the SNOMED codes from the diagnosis and pathology tables where they were different, leaving one SNOMED code per row that provided the most detail (e.g. leaving 'M81403,

adenocarcinoma' in place of 'M80103, carcinoma not otherwise specified' where both were present). In the majority of cases the SNOMED codes in the different tables were identical. The format of the genetic data was simplified to display the predicted protein change only where it was possible to determine this (e.g. for an *EGFR* result reported as 'c.2573T>G (p.L858R)' this was simplified to 'L858R' to facilitate aggregated data analysis).

Once the data had been condensed in individual tabs to one row per patient, it was possible to perform data analysis for the proportions of different histological subtypes present per tumour type, the frequency of different genetic aberrations by histological subtype, the occurrence of multiple aberrations in the same gene or sample and to calculate the proportion of samples failing some or all of the genetic tests as well as correlates between mutations and some of the clinical and pathological data. This data is presented in chapter 3.

Additional data items were collected during the course of SMP1 from clinical and technology hubs as part of the process of monitoring a series of key performance indicators determined at the outset of the programme. The key performance indicator data items were completed on an Excel spreadsheet template (Appendix A) by the operational lead at each site, using cumulative values for numerical data items, and submitted to the central CRUK programme management team as an e-mail attachment by 5pm on 7th day of each month. Analysis of some of these data items is also presented in chapters 3 and 4 to give an insight into test failure rates and turnaround time during SMP1, and possible correlations with sample handling processes in cellular pathology laboratories.

2.2 Cross-sectional analysis of variation in sample handling by cellular pathology laboratories

Cellular pathology departments providing diagnostic services to the NHS are required to obtain accreditation from the UK Accreditation Service (UKAS) and participate in external quality assurance schemes. Part of the accreditation involves use of an over-arching quality management system including establishing and regularly reviewing and updating standard operating

procedures (SOPs) with which all relevant staff are expected to be familiar with and adhere to. Despite a rolling programme of laboratory inspections by a panel of external peers, data are not centrally collated within the NHS on how much variation there is between different departments in the multiple processes involved in preparing tissue for histopathological analysis. The Royal College of Pathologists has published an increasing number of cancer reporting datasets and tissue pathways for non-neoplastic disease over the past fifteen years (https://www.rcpath.org/profession/publications/cancer-datasets.html) and these contain some guidance on how to handle, sample and report a range of specimen types, increasingly with a view to sample preparation for molecular analysis, but these advise rather than mandate tissue handling techniques and are far from comprehensive. It is therefore likely that there is sufficient variation in specimen handling between different NHS cellular pathology departments, and that these differences may affect the success of subsequent mutation analysis on DNA extracted from the tissue. The aim of this work was to establish a baseline and assess variation in current NHS cellular pathology department specimen handling, as an essential step in developing optimised protocols in anticipation of the increasing requirement to analyse nucleic acids from formalin-fixed, paraffin-embedded tumour specimens.

A series of questions covering all aspects of routine tissue sample handling was compiled and sent out by e-mail to a named laboratory contact at all departments of cellular pathology participating in SMP (Appendix B). Individual responses were returned by e-mail and collated onto an Excel spreadsheet. These responses have been summarised and presented in chapter 4.

2.3 Cross-sectional analysis of variation in handling of endobronchial ultrasound-guided lung samples

One of the major challenges of delivering Phase 2 of the Stratified Medicine Programme (SMP2) compared to SMP1 was the move away from molecular analysis of nucleic acids extracted from resection specimens containing plentiful tumour to focusing on small biopsies and cytology cell block preparations which represent the typical samples obtained for diagnosis and/or staging from patients with advanced lung cancer. Endobronchial

ultrasound transbronchial needle aspiration (EBUS-TBNA) techniques have been increasingly adopted over the past few years as the preferred modality for combined diagnosis and staging in such patients. There is a lack of formal guidance on how the small fine needle aspiration cytology specimens obtained from these procedures should be handled in the endoscopy suite and pathology laboratory, and as a result of this there is considerable variation in practice. Following on from the cross-sectional analysis of general cellular pathology department processes, it was decided to explore this further by conducting a similar analysis of the pathology departments participating in SMP2. The aim of this work was to compare sample handling practices and work towards achieving consensus on optimal preservation of the available diagnostic material including nucleic acids. A questionnaire covering aspects of EBUS service provision and sample handling was sent to a named cellular pathologist collaborator at each of twelve NHS organizations involved in the programme (Appendix C). Individual responses were returned by e-mail and collated onto an Excel spreadsheet. These responses have been summarised and presented in chapter 4.

2.4 Experimental work on fixation using alternative PAXgene® Tissue fixation system

2.4.1 Project details

This work relates to work carried out in Southampton as part of a project proposal submitted to and approved for funding by Innovate UK (formerly the Technology Strategy Board) as part of the Stratified Medicine Innovation Platform for industrial-academic collaboration to improve diagnosis on cell and tissue samples. Through CRUK-SMP collaborators, a consortium was formed to embark on a project to investigate the application of the novel PAXgene® Tissue fixative system to preservation of tumour samples for morphological and molecular analysis. The consortium members were Qiagen, University Hospital Southampton NHS Foundation Trust, Papworth NHS Trust, University Hospital Birmingham NHS Foundation Trust, Guy's and St Thomas Hospital NHS Foundation Trust, Queen's University Belfast, NHS Greater Glasgow and Clyde, University College London and the Royal National Orthopaedic Hospital.

University Hospital Southampton NHS Foundation Trust was the lead applicant in obtaining research ethics committee (REC) approval for the study (REC reference number 15/YH/0221).

2.4.2 Tissue sampling

Sampling of all available surgical specimens submitted to the cellular pathology department in Southampton in the fresh state was carried out by an appropriately qualified specialty registrar or consultant pathologist, providing the specimen originated from a patient over 18 years of age and arrived in the department with a copy of the trust NHS procedure consent form completed to indicate patient approval for use of tissue surplus to diagnosis for research purposes. Following removal of the organ or tissue from the patient, the specimen was placed in an opaque, white plastic container of an appropriate size labelled with the patient's details and a cellular pathology specimen request form was completed by a member of the surgical team. The specimen was transferred immediately from the operating theatre to the cellular pathology department by a specimen porter or clinical trials assistant, according to standard trust operating procedures. On arrival at cellular pathology specimen reception, each surgical specimen was booked in by a biomedical support worker, with checking of details on the request form and specimen container to make sure three unique points of identifying patient data were present on each and that they matched. The next available specimen accession number was then allocated to the specimen, in the format 16HSnnnnnA where '16' represents the last two digits of the year of receipt, n is a sequentially allocated 5 digit number and A represents a final alphabetical check letter present on the pre-printed barcoded stickers. A copy of the sticker was placed on the container and request form, and details of the specimen request were entered in electronic form into the laboratory information management system, Labcentre (v12.0 Clinisys Solutions Ltd, Chertsey, UK). The specimen was then transferred to the frozen section room and examined by a pathologist in a Thermo Scientific™ MSC-Advantage™ Class II Biological Safety Cabinet (Thermo Fisher Scientific, Waltham, US). After incision or opening of the specimen using scissors or a scalpel as required to reveal the tumour, two pieces of tissue of similar size and each measuring up to 15 x 15 x 4mm in maximum dimensions were sampled under direct visualisation using

a scalpel. Sequential study numbers were allocated to each case in the format yySFUSnnn where 'yy' are the last two digits of the year, 'SF' is for STRATFix, 'US' is for University Hospital Southampton NHS Foundation Trust and 'nn' is a sequential number, such that the first specimen sampled was 15SFUS001 and the pieces of tissue were placed into cassettes printed with the case number suffixed with 'FA' for the formalin block and 'PA' for the PAXgene® block. If more than one sample was taken into each fixative (e.g. in a specimen with two synchronous tumours or including more than one organ) then additional samples were suffixed with FB/PB, FC/PC etc. Detailed information was recorded for each pair of samples contemporaneously, about multiple sampling handling parameters including key time points throughout the fixation and processing stages (Appendix D).

2.4.3 Preparation of tissue using PAXgene® Tissue System

The PAXgene® Tissue System was provided by Qiagen through the Innovate UK-funded project, in the form of single-use, dual chamber PAXgene® Tissue containers. Each specimen pot consisted of two separate chambers, one containing PAXgene® Tissue Fixation Reagent, a mixture of different alcohols including methanol, acetic acid and a soluble organic compound, and a second chamber containing PAXgene® Tissue Fixation Stabilization Reagent. The tissue in the cassette for the PAXgene® Tissue System block was placed into chamber 1 (PAXgene® Tissue Fixation Reagent) of a dual chamber PAXgene® Tissue container (PreAnalytiX, Hilden, Germany) at ambient temperature and then after an interval of between 3 and 24 hours was moved to chamber 2 (PAXgene® Tissue Fixation Stabilization Reagent) according to the manufacturer's instructions. Due to the importance of avoiding formalin contamination of PAXgene®-fixed tissue, a formalin-free processing system was established in the laboratory for processing of PAXgene®-fixed and stabilised tissue. During the weeks it took to establish a formalin-free processing workflow for the PAXgene® Tissue fixed samples, the tissue was held in stabiliser solution in the laboratory freezer at -20ºC as per the manufacturer's guidance. The tissue was processed in a Sakura VIP® (Tissue-Tek® Vacuum Infiltration Processor 5, Sakura, Alphen aan den Rijn, The Netherlands) on a 12 hour processing program using the reagents and timings described in table 14, following the PAXgene® manufacturer's example tissue processing

protocols as described in the PAXgene® Tissue System product circular. Following this the processed and paraffin-impregnated tissue was embedded in Paraplast Xtra low melting point (LMP) paraffin at a Leica EG1150 embedding station at just below 60°C. The wax was allowed to cool slightly in a slight modification to the usual laboratory procedure of embedding at 65°C, to mirror the lower temperature used during PAXgene® Tissue fixed paraffin-embedded (PFPE) tissue processing.

Table 14. Processing conditions for PAXgene® Tissue-fixed samples

Step	Reagent	Time (hh:mm)	Temperature	Vacuum/ Pressure	Mix
1	80% ethanol	01:00	Ambient	On	Slow
2	90% ethanol	01:00	Ambient	On	Slow
3	100% ethanol	01:00	Ambient	On	Slow
4	100% ethanol	01:00	Ambient	On	Slow
5	100% ethanol	01:00	Ambient	On	Slow
6	Isopropanol	01:00	Ambient	On	Slow
7	Isopropanol	01:00	Ambient	On	Slow
8	Xylene	01:00	Ambient	On	Slow
9	Xylene	01:00	Ambient	On	Slow
10	Paraplast Xtra low melting point (LMP) paraffin	01:00	56ºC	On	Slow
11	Paraplast Xtra LMP paraffin	01:00	56ºC	On	Slow
12	Paraplast Xtra LMP paraffin	01:00	56ºC	On	Slow
Total program time		12:00			

2.4.4 Preparation of formalin-fixed tissue

The tissue in the cassette for the formalin block was placed in a pre-filled 120ml GentaFix pot of 10% neutral buffered formalin (Genta Medical, York, UK) and then after a period of 6-48 hours fixation was processed on an overnight (12 hour) tissue processing run in one of the laboratory's main Thermo Shandon Excelsior™ AS tissue processors (table 15). Following this the processed and paraffin-impregnated tissue was embedded in Paraplast paraffin (Leica Biosystems, Nussloch, Germany) at a Leica EG1150 embedding station at 65°C according to laboratory standard operating procedures.

Table 15. Processing conditions for formalin-fixed samples

Step	Reagent	Time (hh:mm)	Temperature	Vacuum
1	10% formalin	00:40	Ambient	Off
2	70% ethanol	00:30	30˚C	On
3	90% ethanol	00:30	30˚C	On
4	100% ethanol	01:00	30˚C	On
5	100% ethanol	01:00	30˚C	On
6	100% ethanol	01:00	30˚C	On
7	100% ethanol	01:00	30˚C	On
8	Xylene	01:00	30˚C	
9	Xylene	01:00	30˚C	On
10	Xylene	01:00	30˚C	On
11	Paraplast paraffin wax	01:00	62ºC	On
12	Paraplast paraffin wax	01:00	62ºC	On
13	Paraplast paraffin wax	01:20	62ºC	On
Total program time		12:00		

2.4.5 Preparation of sections for morphology assessment

4 micron thick sections (or 2-3 micron thick sections for lymph node and spleen) were cut by a biomedical scientist using a Thermo Scientific™ HM 325 rotary microtome (Thermo Scientific, Massachusetts, United States). These were mounted on glass slides (Thermo Scientific, Germany) and stained with haematoxylin and eosin and coverslipped on a Dako CoverStainer (Dako, Cambridgeshire, UK) The slides were labelled using stickers with coded identifiers, so that the assessing pathologist was unable to tell from the slide label which fixative had been used and could perform their assessment in a blinded manner.

2.4.6. Evaluation of morphology

Tissue morphology for formalin-fixed and PAXgene®-fixed tissue sections was assessed by two consultant histopathologists using a scoring system adapted from the National External Quality Assessment Scheme for haematoxylin and eosin and other tinctorial histochemical-stained slides and the study by Craft *et al.* comparing various fixatives[105] (table 16). The equivalence of the two types of material for morphology preservation and assessment was compared using a Bland-Altman plot performed on GraphPad Prism 7 for Windows (version 7.01, GraphPad Software, Inc.).

Table 16. Scoring system for assessment of quality and suitability for diagnosis of haematoxylin and eosin-stained tissue sections fixed in either formalin or PAXgene® Tissue

Feature	Criteria		Score
Nucleus	Sharp nuclear membrane; chromatin pattern clear; nucleolus, when present is distinct		**4**
	Slight degradation in chromatin pattern, nucleolus when present, less distinct but discernable, sharp nuclear membrane		**3**
	Less distinct nuclear membrane; fuzzy chromatin pattern, nucleolus when present is difficult to discern		**2**
	Fuzzy nuclear membrane, chromatin pattern difficult to determine, nucleoli cannot be detected		**1**
	Nucleus cannot be differentiated from cytoplasm		**0**
Cytoplasm	Normal cellular morphology easily determined		**4**
	Intracytoplasmic details fuzzy		**3**
	Only rare evidence of normal intracellular structures		**2**
	Increased cytoplasmic pallor or increased cytoplasmic eosinophilia assessed as detrimental to assessment of morphology		**1**
	Cytoplasm homogenously pale or eosinophilic with no evidence of organelles		**0**
Cell membrane	Cells have distinct cellular/intercellular membranes; any normal substructures, if present, are easily distinguished		**4**
	Loss of substructures (if present) in some cells; slight loss of intracellular details		**3**
	Loss of substructures (if present) in most cells; obvious blurring of many cellular borders		**2**
	No substructures detected; significant blurring of most cellular borders		**1**
	Not possible to distinguish between adjacent cells		**0**
Total score for all above categories (maximum /12)			
Stroma	Free text comment re. staining quality or other issues affecting stromal tissue compartment		
Other	Free text comment re. presence of exogenous deposits e.g. pigment, precipitates		
Section quality	Staining even across section?	Yes	No
	Section thickness regular?	Yes	No
	Adhesion to slide satisfactory?	Yes	No
	Tissue cracking or other disruption present?	Yes	No
Overall impression of tissue quality for diagnostic purposes: Suitable for diagnosis?		Yes	No

Adapted from scoring system used in: Craft WF, Conway JA and Dark MJ. Comparison of histomorphology and DNA preservation produced by fixatives in the veterinary diagnostic laboratory setting *Peer J.* 2014; **2**: e377.

2.4.6 Preparation of sections for assessment of other histochemical stains

Several histochemical stains were prepared on matched sections of formalin- and PAXgene® Tissue fixed tissue, shown in table 17. 4 micron thick sections were cut by a biomedical scientist using a Thermo Scientific™ HM 325 rotary microtome (Thermo Scientific, Massachusetts, United States) from PFPE and FFPE blocks by a biomedical scientist, and placed onto coated glass SuperFrost™ Plus slides (Thermo Scientific, Germany). These were either stained on an Artisan staining system (Dako, Ely, Cambridgeshire, UK) or by hand (elastic Van Gieson, EVG, only) with histochemical stains selected to be informative according to the tissue type (table 14). The hand-stained Miller's EVG was prepared by rinsing sections with distilled water, treating with acidified potassium permanganate solution for 5 minutes, washing in distilled water then bleaching in 1% oxalic acid for 2 minutes and rinsing again in distilled water. Following a rinse in 95% ethanol, the sections were stained in Miller's elastin solution (VWR International Limited, Leicestershire, UK) for 3 hours. The sections were removed from the solution, rinsed again in 95% ethanol and washed in distilled water. Sections were counterstained with Van Gieson for 5 minutes, blotted dry and rapidly dehydrated in two changes of absolute (100%) ethanol without rinsing. Finally sections were cleared in xylene, mounted and coverslipped.

Table 17. Histochemical stains performed by tissue/ tumour type

Tissue type	Histology	Histochemical staining
Lung	Adenocarcinoma only	EVG
Pleura	Any	EVG
Pancreas	Any	PAS, DPAS
Spleen	Any	Reticulin

DPAS, diastase periodic acid-Schiff; EVG, elastic Van Gieson; PAS, periodic acid-Schiff.

The expected staining results of the histochemical stains performed are shown in table 18.

Table 18. Expected staining pattern of histochemical stains

Stain	Tissue component reaction
Diastase periodic acid-Schiff (DPAS)	Stromal or cytoplasmic mucin – magenta
Elastic Van Gieson (EVG)	Elastin – dark brown/black Muscle, erythrocytes, cytoplasm – yellow Collagen - red
Periodic acid-Schiff (PAS)	Carbohydrates including glycogen and mucin – magenta
Reticulin (Gordon and Sweet method)	Reticulin – black Other tissue components – pink/brown

2.4.7 Evaluation of histochemical stains

Assessment of the quality of slides prepared using histochemical staining was performed separately and in a blinded manner by two consultant histopathologists, using a modified version of the criteria in use by UK NEQAS for Cellular Pathology Technique, in which slides are assessed as satisfactory or unsatisfactory for diagnostic purposes based on a range of technical parameters (table 19).

Table 19. Scoring system for assessment of quality of histochemical stain preparations

Score	Criteria
0	**No section available**
1	**Fail** – no staining
2	**Borderline fail** – unsatisfactory demonstration based on method employed with unsatisfactory results
3	**Pass** - appropriate demonstration based on method employed and expected staining results
4	**Good** – good appropriate demonstration based on method employed and expected staining results
5	**Excellent** - excellent demonstration based on method employed and expected staining results

2.4.8 Preparation of sections for immunohistochemical assessment

4 micron thick sections were cut from PFPE and FFPE blocks by a biomedical scientist using a Thermo Scientific™ HM 325 rotary microtome (Thermo Scientific, Massachusetts, United States), and placed onto coated glass SuperFrost™ Plus slides (Thermo Scientific, Germany) labelled with the patient study ID. Slides were dried at 60 degrees for 30 minutes. The sections were then incubated with commercially available antibodies selected to be informative according to tissue or tumour type (table 20), according to the manufacturer's instructions and usual laboratory protocols on a Dako Link Autostainer 48 (Dako, Cambridgeshire, UK) with pre-treatment for formalin-fixed paraffin-embedded tissue using heat and a high pH buffer (heat-induced epitope retrieval, HIER). The usual antigen retrieval protocol incorporates some reverse cross-linking activity induced by the heat and high pH, but we hypothesised that this would not be required for tissue prepared using the PAXgene® Tissue system since the experimental fixative is non-cross linking. For this reason duplicate sections of PFPE tissue were prepared and processed with and without HIER.

For each slide, a label with a unique identifier was produced using the Dako Autostainer by entering details of the study ID, primary antibody, antigen retrieval system and buffer. Pre-treatment steps for slides undergoing HIER were carried out on the Dako PT Link as follows. The slides were placed in racks in the pre-treatment tanks of the Dako PT link (pre-treatment module) and incubated with Dako EnVision™ FLEX High pH (pH 9) Target Retrieval Solution, except for Ki67 where our usual departmental protocol is for use of Dako EnVision™ FLEX Low pH (pH 6.1) Target Retrieval Solution. The slides were heated to 98°C for 20 mins, with a 20 minute preceding heating stage and a subsequent 20 minute cooling stage. After completion of the pre-treatment racks of slides were submerged in Dako EnVision™ FLEX Wash Buffer for 5 minutes, then rinsed to remove any pre-treatment solution. The racks of pre-treated slides were transferred to the Dako Autostainer Link instrument with the slides not requiring HIER. The Dako Autostainer Link was pre-loaded with bulk reagents according to usual laboratory protocols and the specific antibody reagents outlined in table 21 were added. Visualization was performed with an automated program and the EnVision™ FLEX ready to use kit, in the stages

outlined in table 22, with intervening wash phases in which a fixed volume of buffer (or in later stages distilled water) was pipetted over the slide.

Following staining the slides were removed from the machine, transferred to racks for the coverslipper machine and left in running tap water in a sink for 5 minutes to enhance the blue haematoxylin counterstain. The slides were taken by hand through a series of solutions forming an alcohol gradient into xylene (water, 70% ethanol, 100% ethanol, 100% ethanol, xylene, xylene) with 2 minutes in each solution and some agitation. The slides were mounted using Pertex® (CellPath, Powys, Wales) and coverslipped on a Leica CV5030 automated glass coverslipper machine (Leica Biosystems, Nussloch, Germany).

Table 20. Immunohistochemical antibodies performed by tissue/tumour type

Tissue type	Histology	Antibody
Lung	Adenocarcinoma or squamous cell carcinoma	p63 TTF1 MNF116
Lung	Typical carcinoid tumour	CD56 Chromogranin A Synaptophysin Ki67
Thymus	Any	MNF116 CD3 CD20 TdT
Lymph node	Reactive	CD20 CD3 CD10 Ki67
Lymph node	Hodgkin lymphoma	CD30 CD15 Ki67
Any	Metastatic malignant melanoma	S100 Melan A HMB45

Table 21. Antibody information
Leica is Leica Biosystems, Milton Keynes, UK; BD is BD Biosciences, Oxford, UK and Dako is Dako UK Ltd, Cambridge, UK.

Antibody	Cellular expression in STRATFix sample series	Cellular localisation and expression pattern	Clone, type and manufacturer	Dilution
CD3	T lymphocytes	Membrane	LN10, mouse monoclonal, Leica	1/200
CD10	Follicle centre cells	Cytoplasm/ membrane	56C6, mouse monoclonal, Leica	1/50
CD15	Hodgkin and Reed-Sternberg cells Granulocytes	Membrane and paranuclear golgi accentuation Cytoplasm/ membrane	MMA, mouse monoclonal, BD	1/40
CD20	B lymphocytes	Membrane	L26, mouse monoclonal, Dako	1/250
CD30	Hodgkin and Reed-Sternberg cells	Membrane and paranuclear golgi accentuation	Ber-H2, mouse monoclonal, Dako	1/50
CD56	Carcinoid tumour cells	Cytoplasm/ membrane	CD564, mouse monoclonal, Leica	1/50
Chromogranin A	Carcinoid tumour cells	Cytoplasm/ membrane	5H7, mouse monoclonal, Leica	1/30
Cytokeratin (MNF116)	Epithelium	Cytoplasm	MNF116, mouse monoclonal, Dako	1/300
Melanosome (HMB45)	Cells of melanocytic lineage	Cytoplasm	HMB-45, mouse monoclonal, Dako	1/50
MIB1 (Ki67)	Proliferating cells in cycle	Nucleus	Ki-67, mouse monoclonal, Dako	1/150
Melan A	Cells of melanocytic lineage	Cytoplasm	A103, mouse monoclonal, Leica	1/200
p63	Squamous cells	Nucleus	7JUL, mouse monoclonal, Leica	1/50
S100	Cells of melanocytic lineage	Cytoplasm	Polyclonal, rabbit, Dako	1/4000
Synaptophysin	Carcinoid tumour cells	Cytoplasm	27G12, mouse monoclonal, Leica	1/100
TdT	Thymocytes	Nucleus	SEN28, mouse monoclonal, Leica	1/100
TTF1	Lung epithelial cells	Nucleus	SPT24, mouse monoclonal, Leica	1/300

Table 22. Program used for automated immunohistochemistry on the Dako Link Autostainer 48 platform

Stage		Reagents	Time
1	Wash	Dako buffer	0
2	Blocking of endogenous peroxidase	EnVision™ FLEX blocking reagent	5 minutes
3	Wash	Dako buffer	0
4	Application of primary antibody	As per table 20	20 minutes
5	Wash	Dako buffer	0
6	Application of poly-HRP anti-mouse/rabbit polymer	EnVision™ FLEX mouse/rabbit linker depending on antibody	20 minutes
7	Wash	Dako buffer	0
8	Application of peroxidase	EnVision™ FLEX horseradish peroxidase	20 minutes
9	Wash	Dako buffer	0
10	Wash	Dako buffer	0
11	Application of DAB chromagen	Substrate working solution	5 minutes
12	Application of DAB chromagen	Substrate working solution	5 minutes
13	Wash	Dako buffer	0
14	Counterstain with haematoxylin	EnVision™ FLEX haematoxylin	5 minutes
15	Rinse	Deionised water	5 minutes
16	Wash	Dako buffer	0
17	Rinse	Deionised water	5 minutes

2.4.9 Evaluation of immunohistochemistry

The quality of immunohistochemistry was assessed separately and in a blinded manner by two consultant histopathologists, and scored using the system in use for the UK NEQAS immunohistochemistry scheme (table 23).

Table 23. Scoring system for assessment of quality of immunohistochemistry preparations

Score	Criteria
0	**No section available**
1/2	**Overall not clinically readable** – very weak/ no demonstration of requested antigen; false positive/ negative staining; non-specific or inappropriate staining; uninterpretable staining; excessive morphological damage; excessive haematoxylin
3	**Although clinically interpretable/ readable, improvements can still be made in the staining** – weak demonstration of requested target antigen; background staining; diffuse staining; slightly weak/ excessive haematoxylin
4/5	**Good/ excellent demonstration of requested target antigen**

2.4.10 DNA extraction from PFPE sections

DNA was extracted from eight 10µm thick paraffin scrolls using an extraction kit, reagents and protocol supplied by Qiagen (PAXgene® Tissue DNA Kit, PreAnalytiX GmbH, a QIAGEN/BD company, Switzerland) as follows. Scrolls were cut from each paraffin block by a biomedical scientist using a Thermo Scientific™ HM 325 rotary microtome (Thermo Scientific, Massachusetts, United States), using a sterile technique and placed into a 1.5ml microcentrifuge tube (Thermo Scientific, Massachusetts, United States) labelled with the sample study ID. 200µl of Qiagen blue deparaffinisation solution was added and the tube contents were vortexed for 15 seconds then centrifuged for up to 10 seconds to collect the paraffin and cellular material at the bottom of the tube, in a Heraeus™ Biofuge™ Pico™ centrifuge (Thermo Scientific™, Paisley, UK). The tube was incubated at 56°C for 6 minutes and then left to cool to room temperature for approximately 10 minutes. 200µl of Buffer TD1 were added to the tube which was then vortexed for up to 30 seconds until a visibly homogeneous solution was produced. The tube was then centrifuged for 60 seconds at 11,000g (10,000 rpm). 35µl proteinase K was added to the lower clear phase and then the sample mixed by inserting a pipette and gently pipetting the tube contents up and down. The was incubated at 56°C for up to 60 minutes until the sample was completely lysed, with further incubation at 80°C for 60 minutes. The tube contents were then centrifuged for 10 seconds to remove droplets from inside the lid and the lower, clear phase was

transferred using a pipette into a new 1.5ml microcentrifuge tube labelled with the sample study ID. 200µl of Buffer TD2 were added to the tube which was then vortexed for up to 30 seconds until a visibly homogeneous solution was produced. 200µl of 100% ethanol was added to the tube which was then vortexed for up to 30 seconds until a visibly homogeneous solution was produced. The tube contents were then centrifuged again for 10 seconds to remove droplets from inside the lid. The sample was loaded onto a PAXgene® DNA spin column placed in a 2ml processing tube. This was centrifuged for 60 seconds at 6,000g (8,000 rpm). The spin column was next placed in a new 2ml processing tube and the old processing tube with flow-through was discarded. 500µl Buffer TD3 was added to the PAXgene® DNA spin column and the tube was centrifuged for 60 seconds at 6,000g (8,000 rpm). The spin column was placed in a new 2ml processing tube and the old processing tube with flow-through was discarded. 500µl Buffer TD4 was added to the PAXgene® DNA spin column and the specimen was centrifuged for 60 seconds at 6,000g (8,000 rpm). The spin column was placed in a new 2ml processing tube and centrifuged for 3 minutes at full speed (20,000g or 14,000rpm) to dry the membrane completely. The old processing tube with flow-through was discarded and the PAXgene® DNA spin column placed in a new 1.5ml microcentrifuge tube. 100µl of Buffer TD5 was added directly to the PAXgene® DNA spin column membrane and this was centrifuged for 60 seconds at full speed (20,000g or 14,000 rpm) to elute the DNA.

2.4.11 DNA extraction from FFPE sections

DNA was extracted from eight 10µm thick paraffin scrolls using the QIAamp DNA FFPE Tissue kit (Qiagen, Germany) as follows. Scrolls were cut from each paraffin block by a biomedical scientist using a Thermo Scientific™ HM 325 rotary microtome (Thermo Scientific, Massachusetts, United States), using a sterile technique and placed into a 1.5ml microcentrifuge tube (Thermo Scientific, Massachusetts, United States) labelled with the sample study ID. 200µl of Qiagen blue deparaffinisation solution was added and the tube contents were mixed by vortexing for 15 seconds then centrifuged for 10 seconds to collect the paraffin and cellular material at the bottom of the tube. The tube was incubated at 56°C for 6 minutes and then left to cool to room temperature for approximately 10 minutes. 200µl of ATL buffer were added to

the tube which was then vortexed for up to 30 seconds until a visibly homogeneous solution was produced. The tube was then centrifuged for 60 seconds at 11,000g (10,000 rpm). 35µl proteinase K was added to the lower clear phase and then the sample mixed by inserting a pipette and gently pipetting the tube contents up and down. The was incubated at 56°C for up to 60 minutes until the sample was completely lysed, with further incubation at 90°C for 60 minutes. The tube contents were then centrifuged for 10 seconds to remove droplets from inside the lid and the lower, clear phase was transferred using a pipette into a new 1.5ml microcentrifuge tube labelled with the sample study ID. 400µl of AL buffer were added to the tube which was then vortexed for up to 30 seconds until a visibly homogeneous solution was produced. 400µl of 100% ethanol was added to the tube which was then vortexed for up to 30 seconds until a visibly homogeneous solution was produced. The tube contents were then centrifuged again for 10 seconds to remove droplets from inside the lid. The sample was loaded onto a QIAamp column placed in a 2ml processing tube. This was centrifuged for 60 seconds at 6,000g (8,000 rpm). The spin column was next placed in a new 2ml processing tube and the old processing tube with flow-through was discarded. 500µl AW1 buffer was added to the QIAamp column and the tube was centrifuged for 60 seconds at 6,000g (8,000 rpm). The spin column was placed in a new 2ml processing tube and the old processing tube with flow-through was discarded. 500µl AW2 buffer was added to the QIAamp column and the specimen was centrifuged for 60 seconds at 6,000g (8,000 rpm). The spin column was placed in a new 2ml processing tube and centrifuged for 3 minutes at full speed (20,000g or 14,000rpm) to dry the membrane completely. The old processing tube with flow-through was discarded and the QIAamp column placed in a new 1.5ml microcentrifuge tube. 100µl of ATE buffer was added directly to the centre of the QIAamp column membrane and this was centrifuged for 60 seconds at full speed (20,000g or 14,000 rpm) to elute the DNA.

2.4.12 Assessment of extracted nucleic acids

Extracted nucleic acids were assessed for purity using the NanoDrop® ND-1000 spectrophotometer and double-stranded DNA (dsDNA) quantified using the Qubit® 3.0 fluorometer with broad-range (BR) dsDNA assay (Life

technologies™, Paisley, UK). Fragmentation of DNA was assessed using the BIOMED2 protocol control PCR reactions, to assess the relative abundance of amplified DNA of known fragment lengths.

For spectrophotometry assessment, 1.5µl of the patient sample was loaded onto the NanoDrop measurement pedestal and measured, after zeroing the machine using 1.5µl of deionised water. Measurements of the concentration of DNA present in the sample (in ng/µl), the sample absorbance at 260nm (A260), sample absorbance at 280nm (A280) and absorbance ratios (A_{260}/A_{280} and A_{260}/A_{230}) were read from the screen using the associated software and recorded on a Microsoft Excel spreadsheet against the sample ID.

For fluorometry assessment, after calibration using prepared DNA standards in thin-walled clear 500µl PCR tubes (Axygen Scientific 0.5ml PCR tubes with flat cap, VWR, Pennsylvania, US), 200µl of solution (comprising 2µl of the patient sample combined with 198µl of the DNA working solution) was incubated in a 500µl PCR tube for 2 minutes and was then placed into the sample chamber of the fluorometer. The concentration of dsDNA (ng/mL) in the assay tube and calculated dsDNA concentration of the original sample were read off the instrument screen. The values for each sample were recorded on a Microsoft Excel spreadsheet against the study sample ID.
For fragmentation analysis, molecular weight markers resulting in a ladder of five fragments (100, 200, 300, 400, and 600 bp) and based on the BIOMED-2 control gene PCR protocol were assessed, according to the methods developed by members of the EuroClonality/ BIOMED-2 consortium and previously published elsewhere (van Dongen JJM, Langerak AW, Bruggemann M et al. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: Report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia* (2003) **17**, 2257–2317). Briefly, for each sample, 45µl Specimen Control Size Ladder master mix (IdentiClone™ kit, Invivoscribe Technologies® Inc, San Diego, US) was placed in a 1.5ml microcentrifuge tube (Thermo Scientific, Massachusetts, United States) and 1.25µl (0.25µl at 5units/µl) of AmpliTaq Gold (Applied Biosystems, California, US) was added. 200ng of sample DNA was added and pipetted up and down several times to mix. A 25µl aliquot of each sample then underwent PCR on a Peltier Thermal

Cycler (MJ Research PTC-220 DNA Engine Thermal Cycler, Bio-Rad, Hemel Hempstead, UK) using the following program: 10 minutes at 94°C (pre-activation) followed by 35 cycles of: 94°C for 1 minute (denaturing), 60°C for 1 minute (annealing), 72°C for 1 minute (extension) and then final extension at 72°C for 10 minutes and held at 8°C. The tubes were then removed from the thermocycler. To prepare the PCR products for fluorescent analysis, they were diluted 1:10 in formamide by adding 1μl of PCR product to 9.5μl (Hi-Di) formamide and 0.5μl ROX-400 heteroduplex analysis internal standard. The tube was then vortexed to mix. The PCR product was then denatured at 94°C for two minutes before cooling at 4°C for one hour. Analysis by GeneScanning was performed on the 3500xL Genetic Analyser (Applied Biosystems, California, US), involving separation of denatured, single-stranded PCR products by length in a high-resolution capillary sequencing polymer and detection by automated laser scanning. Electropherograms were then produced and printed using GeneMapper v4.1 software (Applied Biosystems, California, US). Three control samples were run in parallel to the STRATFix samples (two derived from peripheral blood and one from fresh skin).

2.4.13 Statistician input

The advice of a medical statistician has been sought regarding a power calculation to determine the sample size necessary to demonstrate a statistically significant (i.e. $p<0.05$) difference in DNA quality between DNA extracted from either PFPE or FFPE tissue, measured using the abundance of difference fragment sizes as a marker of DNA integrity and suitability for downstream sequencing applications. Using nQuery Advisor® (version 5, Statistical Solutions, Cork, Ireland), it was calculated that with a standard 5% significance level, 2-sided (paired) t-test and 80% power a sample size of 20 would be able to detect differences of at least 45 base pair lengths, based on the assumption that a mean of 150bp for fragment length is observed and a common standard deviation of 68, using a paired samples t test performed using GraphPad Prism 7 for Windows (version 7.01, GraphPad Software, Inc.).

2.5 Work on assessment of tumour content by cellular pathologists

Assessing the composition of the starting material for molecular analysis is a critical step that informs whether to perform the planned analysis or not and also the validity and certainty of the result. This is especially relevant when attempting to confirm the absence of a mutation in a gene (wild-type status), for example to inform clinical trial entry. Histopathological practice has evolved to express this as percentage tumour content of the sample as a proportion of the total; i.e. percentage neoplastic cell nuclei versus other cell nuclei in the tissue such as stromal or inflammatory cells, or adjacent benign tissue. Another possible method is to estimate the volume of tissue (in mm^3) by calculating area of tissue on the slide (in mm^2) multiplied by section thickness (micrometres converted to millimetres) and the number of sections from which tissue has been macrodissected prior to molecular analysis. With wider use of macrodissection of material from the slide, it is more meaningful to limit estimation of the percentage of neoplastic: non-neoplastic nuclei to the area for macrodissection marked on the accompanying H&E stained slide. Others (Gonzalez de Castro *et al.*) have proposed estimation of the total cellularity of the tissue section and classification into a number of different categories though this is not likely to be an intuitive system for many cellular pathologists. Published work has confirmed the variability in pathologist estimation of tumour content and highlighted this as an important area for further study[72 73].

The hypothesis for this research is that there is significant intra- and inter-observer variation in estimation of tumour content by histopathologists when compared to automated digital analysis or cell counting techniques. Intra-observer variability is defined as the same pathologist making a different assessment of one or more cases when presented to them on different occasions, whereas inter-observer variability is defined as different pathologists reaching a different conclusion of tumour content in a series of the same cases.

For this project, pathologists at centres participating in SMP2 as part of the pathology working group were asked to assess the tumour content of a set of

eleven lung cancer slides from patients enrolled into the programme at one centre (Southampton). The slides were scanned at x200 magnification to obtain a whole slide image in both Olympus virtual slide image (.vsi) and big tiff (.btiff) file format, using an Olympus Dotslide scanning system (Olympus Life Science, Tokyo, Japan) and uploaded into the POSHviewer online slide viewing portal which was created by Kevin Wheeler, a programmer and developer in the Cancer Research Informatics Unit at the University of Southampton and has previously been used to study pathologist intra- and inter-observer variation in assessment of scanned breast cancer slides in a large breast cancer cohort study with the results of this assessment published by our group [106]. The system uses Zoomify software (Zoomify, California, US) to produce 'tiles' of the slide scan image which are sequentially displayed as the user navigates around the screen in a similar manner to the software behind Google Maps (Google, California, US). An account was set up for each user on the online system and they were provided with a unique username and password to access the system and record their scores for the cases. In order to increase the number of data points, represent the practice of macrodissection for tumour enrichment and attempt to ascertain whether a smaller area was easier to score, for nine of the cases a thumbnail image with an annotated area marked for macrodissection was also displayed on the screen. The user was provided with instructions asking them to assess and provide a score for the whole slide and then the marked area only, estimating the percentage tumour cell nuclei content (as a percentage of all nuclei present) and assigning it a score according to the following categories: unable to assess; no tumour; 0-5%; 6-10%; 11-20%; 21-30%; 31-40%; 41-50%; 51-60%; 61-70%; 71-80%; 81-90%; 91-100%. Figures 3 and 4 are screenshots from the top and bottom halves of the screen, showing a user's view of the system during scoring of a case.

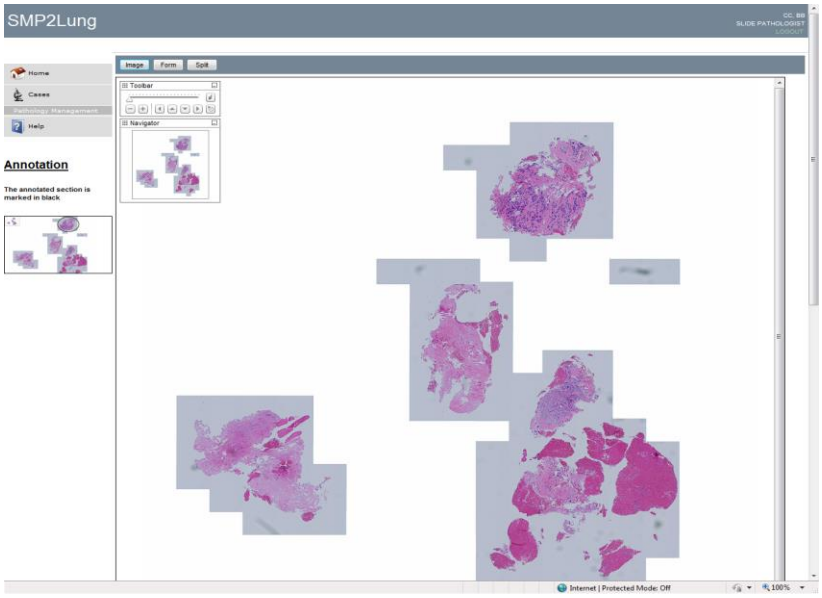

Figure 3. Screenshot of a scanned SMP2Lung slide in the POSHviewer application

The large pane contains a 3-dimensional scanned image on which the zoom function can be used to enlarge an area for examination at higher magnification. A small navigation pane can be seen in the top left corner and to the left of this pane is a thumbnail overview image with an annotated area for separate assessment.

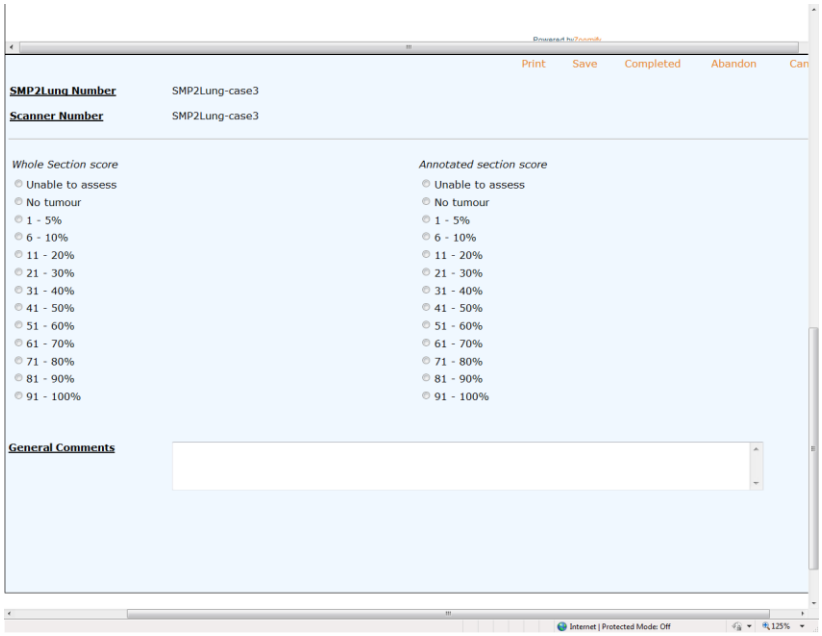

Figure 4. Screenshot of the lower half of the webpage for SMP2Lung data entry in the POSHviewer application

The pathologist is requested to allocate each case, both the whole slide area and annotated area, to a category based on their estimation of tumour nuclear percentage content.

The tumour content of the slides has been determined quantitatively by Dr Nicholas West in Leeds, using the RandomSpot cell counting technique developed at the Leeds Institute of Molecular Medicine by Dr Darren Treanor and colleagues[107]. Each case was subjected to a manual point-counting process by a pathologist, completed in approximately 30 minutes per case and performed separately for annotated and whole slide areas. The RandomSpot technique involves the insertion of between 285 and 315 random spots on the image using locally developed software, and then manual determination of the tissue component beneath each point, which was classified as either:

1. non-cellular (including all non-nucleated tissue)
2. tumour
3. non-tumour
4. necrosis

Once all points had been classified, the overall percentage tumour, percentage non-tumour and percentage necrosis were calculated.

Results of assessment of the tumour percentage content between different pathologists have been expressed graphically using Microsoft Excel® 2007 (Microsoft, Washington, USA) on hybrid dot plots, using ranges and the standard deviation to describe the spread of the raw data, with the lower and upper limits of the 95% normal range (limit of agreement) calculated as follows:

Lower limit = mean difference – (1.96 x standard deviation)

Upper limit = mean difference + (1.96 x standard deviation)

Since the assessing pathologist was asked to select a category containing a range of values for tumour percentage, the numerical mid-point has been used for calculations. The degree of inter-observer variability and comparison to digital tumour content determination performed by Leeds has been assessed using the kappa statistic, interpreted according to the suggested categories of Landis and Koch (table 24) [108], and intraclass correlation coefficient (ICC) which for the multiple raters in this situation is used as an alternative to a weighted kappa. Data analysis has been performed using SPSS (Statistical Package for the Social Sciences, v21.0, Chicago, USA).

Table 24. Landis and Koch interpretation of the kappa statistic

Kappa statistic value range	Interpretation
Less than 0.21	Slight agreement (negative values indicating systematic disagreement)
0.21 – 0.40	Fair agreement
0.41 – 0.60	Moderate agreement
0.61 – 0.80	Substantial agreement
0.81 – 1.00	Almost perfect agreement

2.5.1 Statistician input

Since colleagues who had previously tried to convince busy pathologists to perform online histopathological review and scoring reported difficulty in getting sufficient assessors to complete cases, the planned methodology was discussed with a statistician and it was agreed that there was feasibility justification to make a pragmatic judgement on sample size. A decision was made to scan ten cases and extend the invitation to include the entire pathologist membership of the CRUK SMP2 pathology working group, i.e. 24 pathologists.

2.6 Summary

Completion of SMP1 demonstrated the feasibility and acceptability of provision of molecular analysis on solid tumour samples by a network of testing laboratories. Analysing the collated histological and mutation data demonstrates how representative this cohort of patients is of the wider UK patient population and gives an indication of the generalisability of the findings as well as the prevalence of 'actionable' mutations. Carrying out cross-sectional analysis of multiple different aspects of laboratory handling of samples prepared for molecular analysis at multiple participating sites allows an insight into how much variability there is between supposedly 'standard' operating procedures and the possibility of linking this to test failure rates. Investigating the application of an alternative tissue fixative provides the

opportunity to assess the extent to which this may be compatible with morphology assessment for diagnosis, usual protocols for histochemical and immunohistochemical staining as well as the quantity and quality of extracted nucleic acids for molecular applications, using matched formalin-fixed paraffin-embedded tissue samples as a comparator. Finally, use of an online slide viewing assessment system to assess the level of concordance and variability between lung pathologists' assessment of tumour content in a series of lung cancer samples submitted for molecular analysis allows me to explore the requirement and readiness for a further change in practice required for routine molecular analysis of tumour samples.

Chapter 3: Findings from Phase One of the Stratified Medicine Programme

3.1 Patient accrual and consent

Between August 2011 and July 2013, 10754 patients consented to analysis of material surplus to diagnostic requirements from their tumour sample through SMP1, with 9010 samples sent for analysis by the end of SMP1 (figure 5). The consent rate was consistently high throughout the programme, with an overall average of 98% of those approached consenting to participate in the initiative, demonstrating the acceptability of this approach to patients.

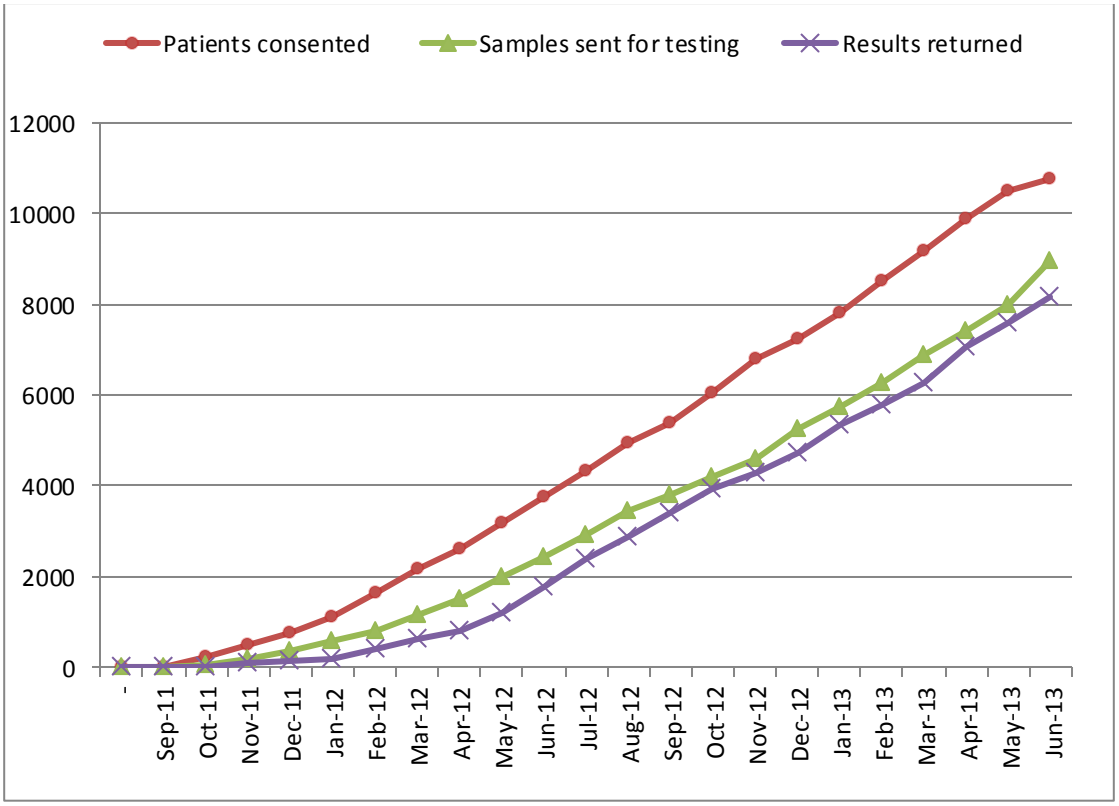

Figure 5. SMP1 patient and sample accrual

Although at the outset of the programme most participating hospitals had some provision for research use of tissue incorporated into standard consent forms in routine use for clinical procedures, only two out of eight allowed an additional blood sample and permission to report back findings of potential clinical significance. A dedicated consent form and patient information sheet

(PIS) were therefore created and granted research ethics committee (REC) approval for use in SMP. It was also possible for participating sites to seek agreement to use their existing biobanking consent forms, following central REC review and approval of all paperwork. This led to differing regional models of obtaining consent across the programme sites and the opportunity to share examples of innovative and successful practice (table 25).

Table 25. Variations in the process of obtaining consent for SMP1 across participating sites

Area of practice	Different approaches observed	Comments
Timing of consent	Before diagnosis of cancer, surgical pre-assessment, on day of surgery, post-operatively	Timing of approach should be chosen with care, with risk of information overload and consideration of other concerns and decisions facing patient
Method of initial approach	Study information sent in advance of appointment Approach in person on day of clinic	Provision of information in advance allows patient to consider in their own time
Professional background of person taking consent	None (electronic) Research nurse, physician, biobank technician, clinical trials or research assistant	Consider training requirements, competing pressures and time availability of different staff groups
Method of recording consent	Paper or electronic	Electronic record of consent/ authorisation facilitates access to wider clinical and research team
Format of study information	CRUK SMP consent form and patient information sheet, Existing institutional biobank consent form and patient information sheet, Other biobanking literature (e.g. Breast Cancer Now)	Opportunity taken to reduce duplication and overlap of information for different studies where possible
Other	Dedicated blood sample for the study in a separate container at time of drawing other bloods required for clinical care, acquisition of surplus blood from haematology	Patient participation likely to be facilitated by reduction in additional requirements to participate in study over and above requirements of clinical care

One area of concern with SMP1 was to what extent information of potential clinical significance should be communicated back to the patient via their treating physician. Following consultation with patients and physicians including clinical geneticists, it was agreed that it would be unethical not to disclose information related to the current condition, for example the presence of a molecular marker that may enable entry into a clinical trial. After further consultation, including with members of the National Cancer Research Institute Consumer Liaison Group, the following wording was agreed for inclusion in the patient literature:

I understand that my samples may be used in research aimed at understanding the genetic influences on diseases and that I will not receive the results of this research. If any research tests on my samples might have impact on my care during the course of my treatment I agree to my clinical care team being informed.

3.1.1 Innovative approaches to gaining consent for tissue donation

A problem consistently encountered at clinical sites participating in SMP1 was a lack of resources to identify all potentially eligible patients and offer them the opportunity to participate in the programme. Contacting patients in person requires staff time as well as adequate space, both of which are in short supply in busy clinics. Removing this requirement by providing patients with information outside of the clinic setting, and even the ability to provide consent remotely, has the potential to increase the number of patients approached. For instance one Scottish site was able to demonstrate an innovative electronic authorisation process, where prospective broad agreement for research use of surplus tissue is recorded electronically as part of the patient's medical record, and instantly updated and accessed by members of the clinical team in different locations. For over 1000 patients approached at pre-operative assessment in this manner, the acceptance rate was 98%, with only 18 out of 1005 patients declining (and one further patient withdrawing at a later date). The process involves providing a single patient information leaflet describing the process of giving permission for future use of pseudonymised tissue samples, including genetic analysis and the

possibility of returning information to the patient's healthcare record. The concept of 'authorisation' as an alternative to consent was introduced in the Human Tissue (Scotland) Act 2006, following the findings of the Scottish Review Group on Retention of Organs at Post-Mortem (http://www.sehd.scot.nhs.uk/scotorgrev/Final%20Report/ropm-09.htm), recognising that parents should be able to give permission for a hospital post mortem examination to be carried out on their child without having to receive detailed information about what the procedure would involve.

Recently, there has been a move towards patient-driven consent, ownership and access to medical information, with the aim of maximising patient participation in clinical care and also access to clinical research opportunities. The National Institute for Health Research's 'OK to Ask' campaign that encourages patients to ask their clinicians about opportunities to participate in trials started in 2013 and has had a high level of support, with indications that momentum for research was boosted at participating organisations (http://www.nihr.ac.uk/newsroom/get-involved-news/its-ok-to-ask-about-clinical-research-on-international-clinical-trials-day/2814). Thus patients are already interested in clinical research, and with access to the right information can be empowered to actively seek opportunities to participate. Using simple, easily-accessible technology to engage willing patients in their own time will give them the opportunity to approach a trial directly and ultimately save staff time. The Moffit Cancer Centre in Florida, US runs the Total Cancer Care programme which, similar to CRUK-SMP, asks patients to consent to their clinical data and excess tissue samples being stored and made accessible to researchers (https://moffitt.org/clinical-trials-research/clinical-trials/total-cancer-care/patients/). The patient-focused section of the website links to a video that explains the importance of patient participation for the programme, with input from clinicians, researchers and patients. Using simple, easily-accessible technology to engage patients gives them the opportunity to approach a trial directly, obtain information at a time of their own choosing and potentially reduces the time needed for face to face explanation at the time of obtaining consent.

3.1.2 Experience of regulatory and ethical requirements

Researchers at one participating Trust started the process of amending their main procedure consent form in 2009 to incorporate appropriate HTA-compliant information and facilitate consent for sample donation. This would provide any patient undergoing investigation or treatment with the opportunity to donate samples to the tissue bank for research. The intention was that all patients would receive a PIS in advance, usually sent out with the pre-operative clinic appointment letter, so that by the time they were consented for their surgery they would have received sufficient information to allow a discussion about sample donation and associated consent at the same time. The ethics committee initially advised that the research and procedure consent forms should be kept separate, involving two signatures. This led to concerns that fully separate consent documentation might reduce the number of patients consenting to tissue donation, be prone to omission, and cause unnecessary administrative burden and lack of clarity with patients and consenting staff regarding simultaneous tissue donation for a number of other approved projects. Since introduction of the new combined form in 2010, the team have collected samples from over 2,500 patients. The consent also covers access to surplus tissue held in the pathology archive from previous procedures. The completed and signed paper consent form is scanned onto the electronic patient record and a copy is usually also sent with the specimen from theatre.

3.1.3 The patient's perspective

At many sites involved in SMP1, collection of samples from patients was already ongoing for various research projects or tissue banks. One of the key aims therefore was to maximise integration of consent processes where possible, in order to avoid consenting patients twice and maximise the utilisation of donated tissue. In Leeds, this was achieved by utilising the existing PIS and consent forms for the Leeds Multidisciplinary Research Tissue Bank (RTB) for patients with ovarian cancer or colorectal cancer. In these cases the consent was for use of tissue (including archival diagnostic tissue) and blood samples, with broad consent obtained for use across a range of studies subject to approval by the RTB Management Committee and so samples could be released to SMP1. Conversely patients with melanoma were not already being recruited to donate

samples to this bank and the SMP1-specific information sheet and consent form were used for that cohort. In addition, patients with breast cancer were already being consented for research use of their samples through the existing Leeds Breast Cancer Now (formerly Breast Cancer Campaign) RTB. These patients were consented using the existing PIS and consent forms for that RTB, following approval being sought from the BCC, and with samples then being released from the RTB to the Stratified Medicine Programme. PIS and consent forms were sent out to patients in advance of their clinic appointment. Overall the response from patients was very positive, as the following examples of comments made by patients demonstrate:

- "You can have it, as it's of no use to me."
- "Great that this can benefit my children, grandchildren and future generations."
- "I thought you were able to just keep the tissue anyway."
- "Basically I think it should be compulsory that we donate surplus tissue after surgery to research."
- "Good to know that I might have contributed to helping others in the future."

An important observation made independently by staff at several clinical sites involved in the programme was under-representation of certain ethnic groups among participants enrolling to research studies. Other than melanoma, a cancer which many ethnic groups tend to suffer from less, involvement of ethnic groups in SMPI was very consistently reported at 6-7% in all other cancers assessed: even with significant amounts of unreported data in this category, this percentage appears relatively small compared to the approximately 13% population of UK ethnic groups quoted in the 2011 census (Table 26). There are many possible contributory factors to this situation, including cultural beliefs and language barriers, which must be overcome in order to ensure that the findings of clinical trials are representative and reproducible in the general clinical patient population, particularly given the increasing awareness and acknowledgement of the influence of genetic factors on disease and treatment response. Recruiting consenting staff from different cultural and ethnic backgrounds may help to address this issue.

Table 26. Recruitment of patients from different ethnic groups in SMP1

Disease Group	Reported White British (%)	Reported Other Ethnic Groups (%)
Lung	1280 (67.9)	127 (6.7)
Colorectal	1071 (66.7)	95 (5.9)
Melanoma	416 (77.8)	18 (3.4)
Prostate	730 (53.7)	92 (6.8)
Ovary	358 (64.3)	33 (5.9)
Breast	1215 (64.9)	124 (6.6)
Total	5,070 (65.9)	489 (5.9)

3.1.4 Importance of clinical engagement and collaboration

At another site, prior to involvement in SMP1, trained biobank technician staff were approaching and consenting approximately 60% of all lung cancer patients at the time of admission for surgery. The timing of patient approach for consent was changed to outpatient clinic attendance, leading to an approximately 30% increase in number of patients consenting for biobanking and SMP1. Timing of approach has to be handled sensitively at a very difficult time in a patient's life, with a new cancer diagnosis and risk of information overload, however the acceptability of this approach to patients was indicated by the consistently high consent rates throughout the two year period: above 96% with a total of 10,754 patients who gave consent to participate.

Data collected at one of the clinical sites involved in the programme offered an insight into factors affecting study participation. At this site there was a consent rate of 94% overall for the programme across 4 cancer types. The highest opt-out rate for any disease group was for patients with lung cancer (table 24). The main reason for not consenting to the programme was down to anxiety before the operation rather than any specific objections to the research itself. A fifth of patients who declined did so because they objected to the pharmaceutical support for the study. Only 3% of patients declined because they were against the genetic testing element of the study (figure 6).

Table 27. Consent rates for different disease indications at one clinical site

Disease Group	Total Consent	Total Declined (%)
Lung	649	47 (7.2)
Colorectal	231	11 (4.8)
Melanoma	80	1 (1.3)
Prostate	144	2 (1.4)
Total	1104	61 (5.5)

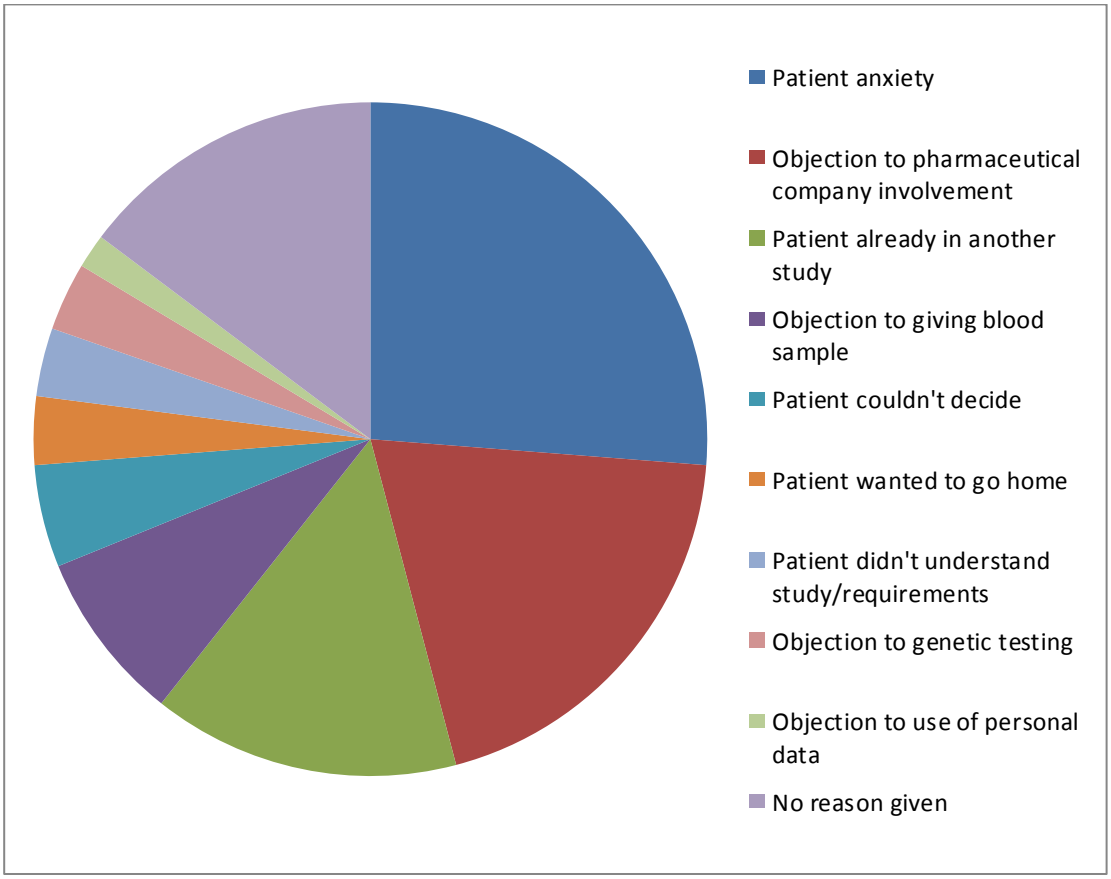

Figure 6. Reasons given by patients to consent staff for not wishing to participate in the Stratified Medicine Programme at one clinical site

The overall experience reported by staff taking consent from patients was that the process generated very few questions. Some examples of specific issues raised are as follows:

- Is any additional tissue going to be taken?
- Will my personal data be safe?
- Will the researchers be able to identify me from the sample or the data?

At another site, a decision was made to send out a follow up letter of thanks to people who had consented to donate tissue for research. In response to more than 400 letters sent, many patients took the opportunity to voice their ongoing support and only one patient replied indicating a wish to withdraw their consent. The team commented on the importance of sending out letters promptly and were also able to check each patient's current details on an electronic clinical portal, comprehensively linked to primary and secondary care, in order to avoid inadvertently sending out a letter after a patient's death.

3.1.5 Issues around pathology and access to archival tissue blocks

Most stratified medicine research programmes, including SMP1, are heavily reliant on access to tissues stored in clinical diagnostic archives. It is widely recognised that these samples represent a valuable research resource. Techniques for the isolation of reasonable quality DNA and RNA from fixed tissue are now well developed and should continue to improve. In the UK, consent is not legally required by the Human Tissue Act for the storage and use of tissue for ethically approved research from living persons not identifiable to the researcher. In practical terms it is relatively easy for ethically approved tissue banks to access diagnostic archives and ensure researchers receive non-identifiable tissue for scientifically valid approved research. However, when the researcher is also the treating clinician who has access to clinical databases, and particularly when smaller disease-specific cohorts are involved (making individual patients more easily identifiable), this can become complex to manage. Currently, individual institutions have set up their own guidelines, but there is a strong case for attempting to establish standardised procedures acceptable to all relevant authorities. A further important consideration is that research use of surplus diagnostic tissue does not exhaust the available stored tissue, in case it becomes necessary to revisit this in clinically relevant circumstances such as review of the diagnosis or the need for further analysis to inform patient management using immunohistochemical or molecular genetic techniques.

3.2 Clinical data

Data completeness within the submitted data proved variable across the sites and for different data items (table 28). There was a degree of overlap and redundancy between data items allowing some gaps to be filled, for example the integrated stage could be determined if the individual components of TNM (tumour/nodes/metastasis) classification had been separately submitted. Although there was an aspiration at the outset of the programme for the dataset to be populated by automated data extraction from electronic patient records, informal feedback during the programme indicated that due to a lack of standardisation across NHS systems instead there was a requirement for intensive compilation of data manually from multiple different clinical systems.

Table 28. Overall SMP1 data completeness by patient disease cohort

Data item	Patient cohort						Overall
	Breast cancer	Colorectal cancer	Lung cancer	Malignant melanoma	Ovarian cancer	Prostate cancer	
Total number of patients	1873	1605	1885	535	557	1359	**7814**
Gender	100%	99%	98%	96%	N/A	N/A	**98%**
Year of birth*	100%	100%	100%	100%	100%	100%	**100%**
Year of diagnosis	79%	75%	52%	74%	67%	69%	**69%**
Ethnic category	71%	73%	75%	81%	70%	60%	**72%**
Histological subtype (SNOMED morphology)	100%	99%	77%	92%	97%	92%	**93%**
Histological grade**	83%	88%	N/A	N/A	62%	53%	**72%**
Pathology TNM T classification***	92%	69%	91%	33%	35%	50%	**62%**
Pathology TNM N classification***	86%	81%	89%	31%	24%	35%	**58%**
Pathology TNM M classification***	24%	74%	77%	54%	79%	33%	**57%**
Integrated stage***	92%	89%	94%	71%	84%	55%	**81%**

Percentage of patient records containing valid and informative data according to the stipulated attributes in the clinical dataset. *Date of birth and date of diagnosis were recorded at patient level but truncated to 'year of' as an information governance measure to maintain confidentiality. **Not mandatory where this is not a core RCPath dataset reporting item. For prostate cancer the percentage refers to overall completeness of Gleason score components requested in separate data items. *** Alternative staging systems used as follows with completeness given in integrated stage field: FIGO for ovarian cancer, AJCC version of TNM7 for melanoma. TNM7 has been used in all cases apart from colorectal cancer where TNM5 is currently used in the UK according to RCPath guidance. TNM5/7 = l'Union Internationale Contre le Cancer (UICC) Tumour/Node/Metastasis Classification of Malignant Tumours 5th/7th edition; FIGO = International Federation of Gynaecology and Obstetrics; AJCC = American Joint Committee on Cancer, SNOMED = Standard Nomenclature of Medicine.

3.3 Molecular data

Of a total of 7813 samples with data available at the time of this analysis, 53% had at least one aberration detected despite the relatively limited scope of genetic analysis in this pilot study. 45% of the samples were wild type for the genes and regions analysed and the remaining 2% of samples failed all gene tests. This low total fail rate indicates that targeted mutation analysis is feasible in formalin-fixed paraffin-embedded clinical tissue samples using the variety of methods employed in SMP1. Overall results of the molecular analysis broken down by tumour type are displayed in table 29.

Table 29. Summary results of molecular analysis performed during SMP1 by tumour type

Tumour type	Breast cancer	Colorectal cancer	Lung cancer	Malignant melanoma	Ovarian cancer	Prostate cancer
Number of samples	1873	1605	1885	535	557	1359
Failed all tests	5.3%	0.9%	2.8%	3.6%	3.2%	4.0%
Wild type for all genes	45%	19%	64%	31%	40%	52%
Aberration in more than one gene	7%	33%	0.5%	2%	4%	2%

A more detailed breakdown of the results by individual tumour type and gene is given in the following sections, including a comparison of the mutation rates compared to those in the scientific literature. Direct comparison is complicated by numerous potential confounding factors such as a bias towards either early or late stage disease or particular histological subtypes in these studies. Although the SMP1 eligibility criteria were deliberately broad, there is likely to be bias in the SMP1 cohort due to selection of patients undergoing surgical resections where tissue would be more plentiful and thus likely to represent early stage disease, and also those not being approached and consented for other studies. This could be avoided where consent for SMP1 utilised local biobanking consent paperwork, minimising duplication and the risk of

information overload. The implicit bias in the SMP1 cohort is illustrated by a comment made by one of the breast cancer team at the Royal Marsden clinical hub site during one site visit, indicating that they would predominantly focus on patients with ER negative disease since they had more clinical trials open to offer patients with ER positive breast cancer. This selection bias, which limits the generalisability of the SMP1 mutation analysis findings, can also be borne out by comparison of histological subtype breakdown to the wider population of patients. In the tumour-specific tables, a mutation is considered as 'clinically actionable' if its presence will confer patient access to a licensed therapy in the appropriate clinical situation. This does not take into account clinical trial recruitment options that may exist for patients whose tumours show specific mutations, such as the molecularly stratified FOCUS4 trial in colorectal cancer and the National Lung Matrix Trial or the increasing number of histology-agnostic 'basket studies' involving therapies directed against particular genetic mutations.

3.3.1 Lung cancer

Of 1885 lung cancers, 36% had at least one abnormality, and only 0.5% had more than one. *KRAS* was most often mutated (23%), followed by *EGFR* (7.5%), *ALK* rearrangement (1.9%) and *BRAF* (1%) (table 30 and figure 7). A range of tumours of different histological subtype were present in the SMP1 cohort (figure 8). Analysis by histological subtype enriched for certain mutations, with 922 pulmonary adenocarcinomas showing 36.6% *KRAS* mutations, 11.6% *EGFR* mutations and 2.5% *ALK* gene rearrangements (figure 9). These results are comparable with those obtained by The Cancer Genome Atlas (TCGA) initiative in recently published data from analysis of 230 primary pulmonary adenocarcinomas with a mutation frequency of 33% for *KRAS*, 14% for *EGFR* and *ALK* gene rearrangements in 1.3%[109]. *EGFR* and *KRAS* mutations were mutually exclusive in the TCGA cohort, as in other series[110 111], but in the SMP1 data there were 4 cases in which both *EGFR* and *KRAS* mutations co-existed. This does not preclude use of a stepwise testing strategy with *KRAS* analysis performed first, since the presence of a *KRAS* mutation in the tumour would contraindicate EGFR inhibitor therapy. As per the reported literature, the majority of mutations in the *EGFR* gene were found in exon 19 and exon 21 with the occurrence of exon 20 mutations in this cohort higher than reported

elsewhere, for reasons that are not clear (figure 10). Over half of the detected *KRAS* mutations were in exon 12 (figure 11) and due to the different methodology used by the different laboratories in some cases it was possible to detect a mutation but not to determine whether it was in codon 12 or 13.

Table 30. Gene mutation rates detected in the SMP1 lung cancer patient cohort compared to those in the scientific literature

Gene	Published prevalence of aberration	Reference	Prevalence in SMP1 cohort	Potentially clinically actionable?
EGFR	7-15%*	[112]	7.5%	Yes – EGFR tyrosine kinase inhibitor therapy with erlotinib or gefitinib
KRAS	16%	[113]	23%	No
ALK translocation	2-7%	[114]	1.9%	Yes – ALK inhibitor therapy with crizotinib or ceritinib
BRAF	2%	[113]	1%	No

* Published mutation frequency in Caucasian patients.

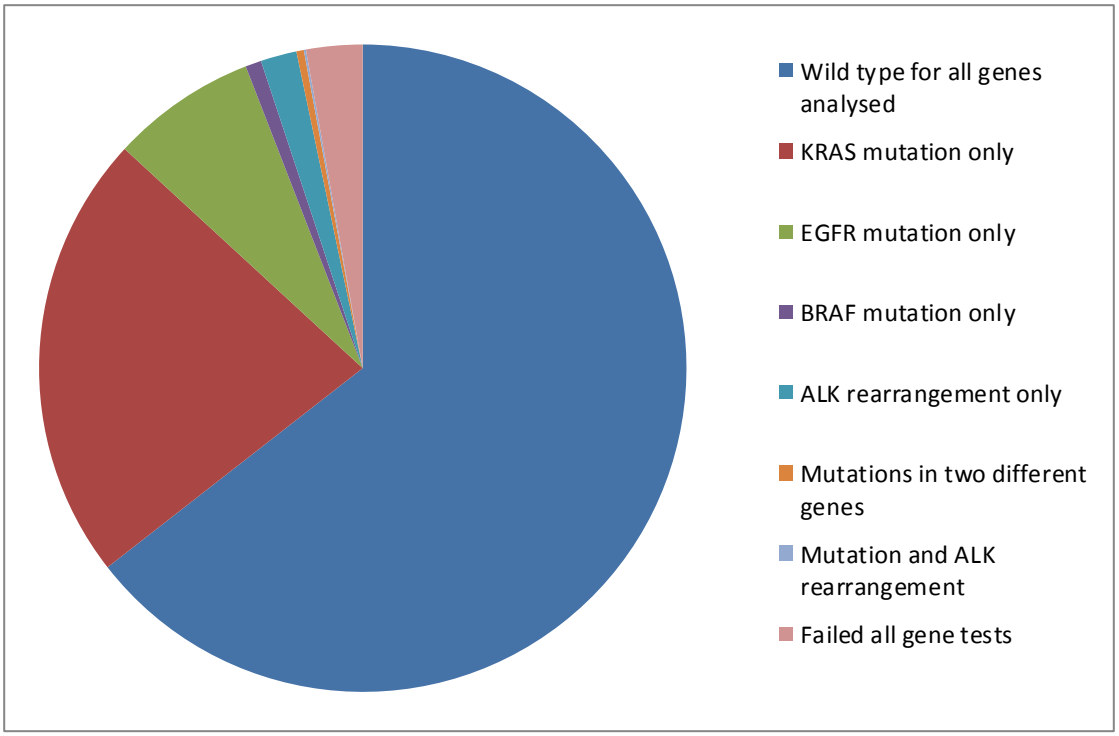

Figure 7. Mutations in tumours from the SMP1 lung cancer patient cohort as a whole

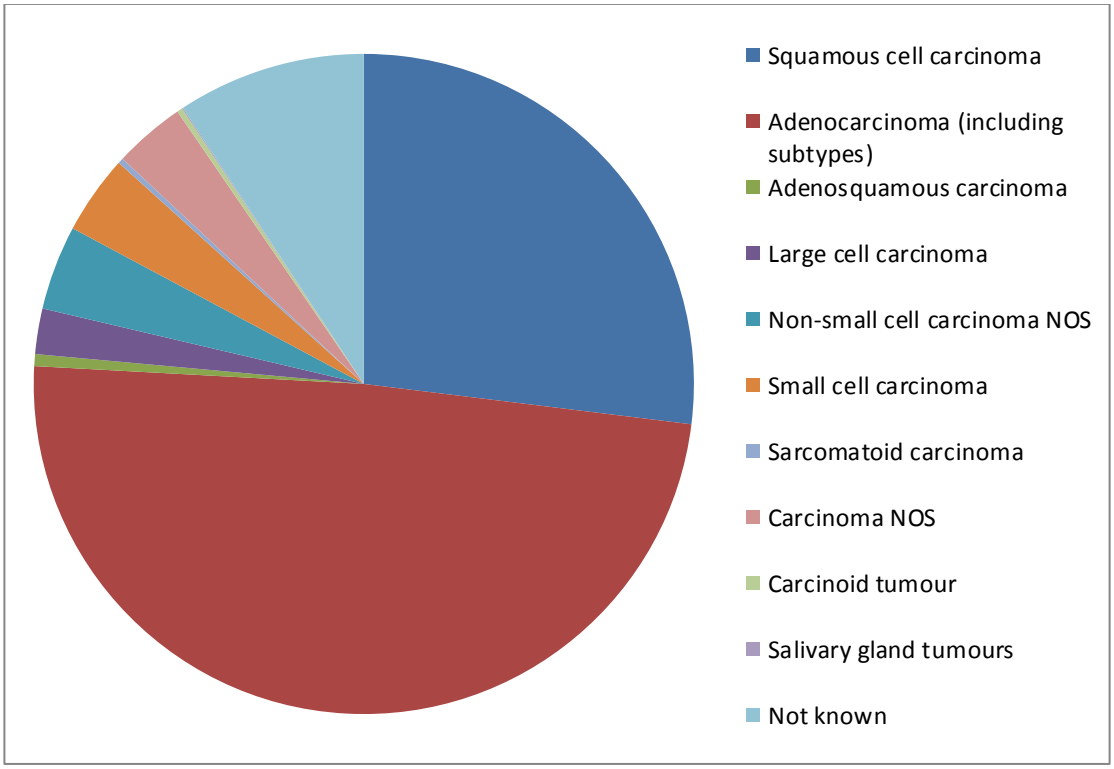

Figure 8. Histological subtype of tumours in the SMP1 lung cancer patient cohort

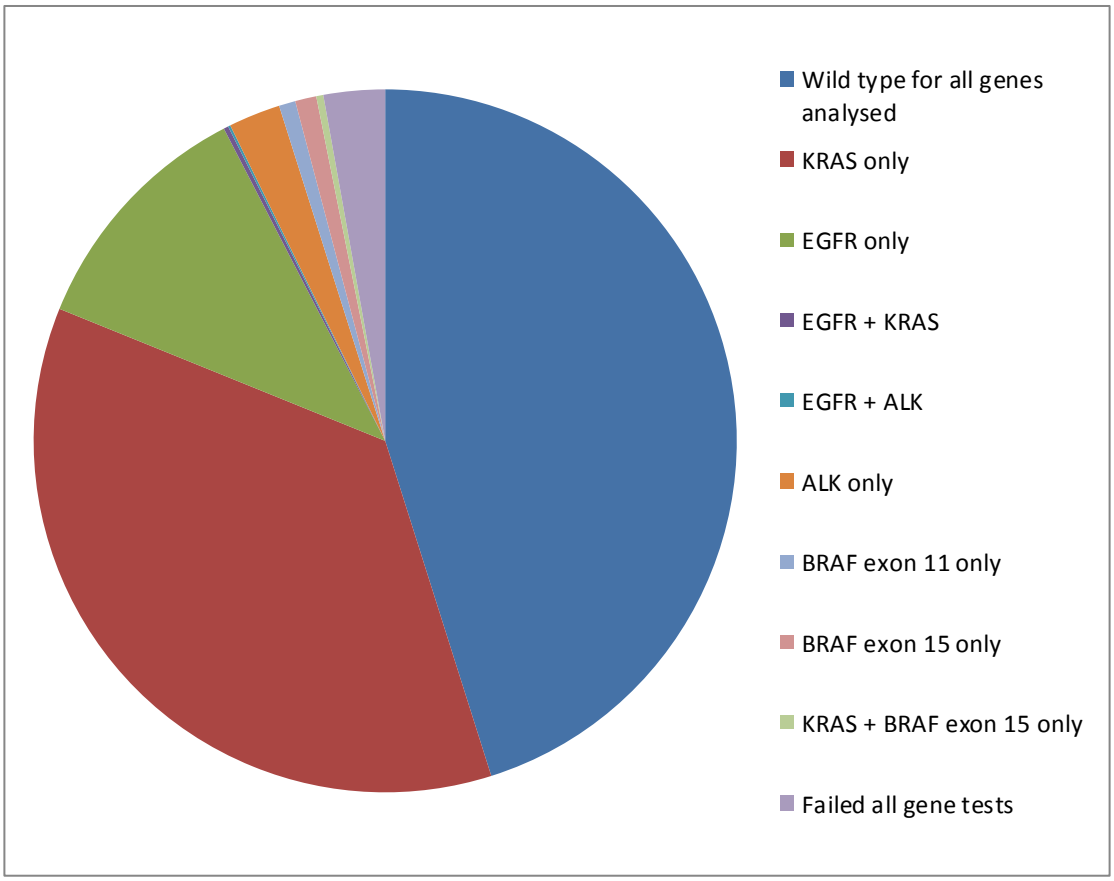

Figure 9. Mutations in the SMP1 pulmonary adenocarcinoma subgroup

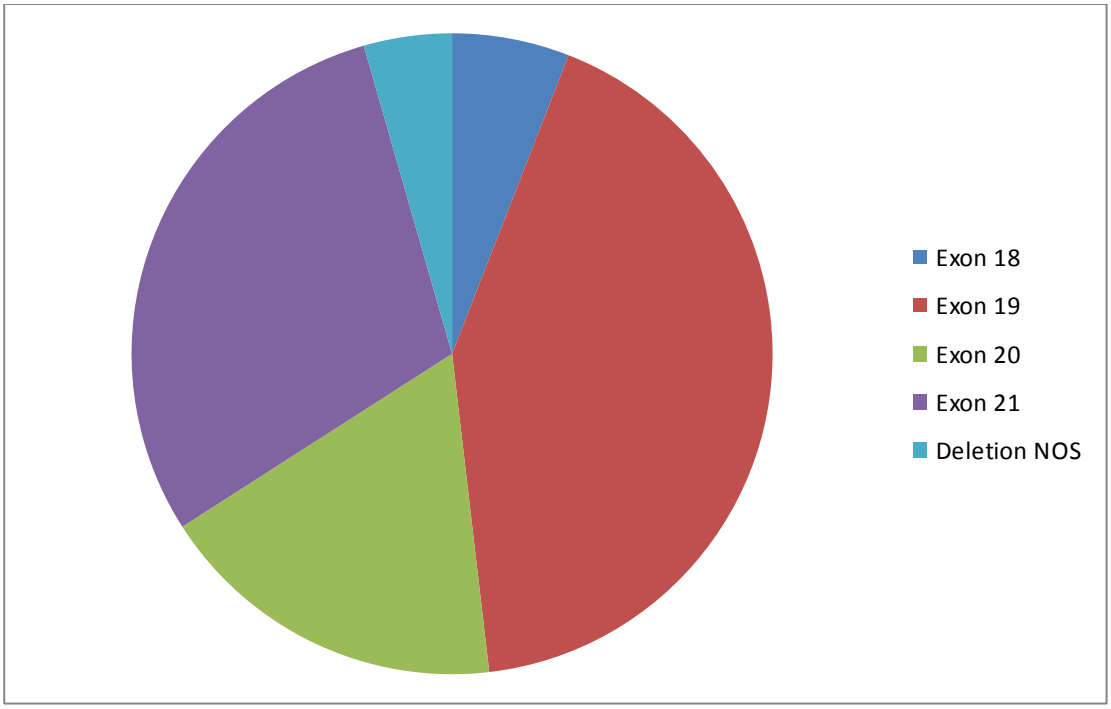

Figure 10. Distribution of *EGFR* mutations in the SMP1 lung cancer patient cohort

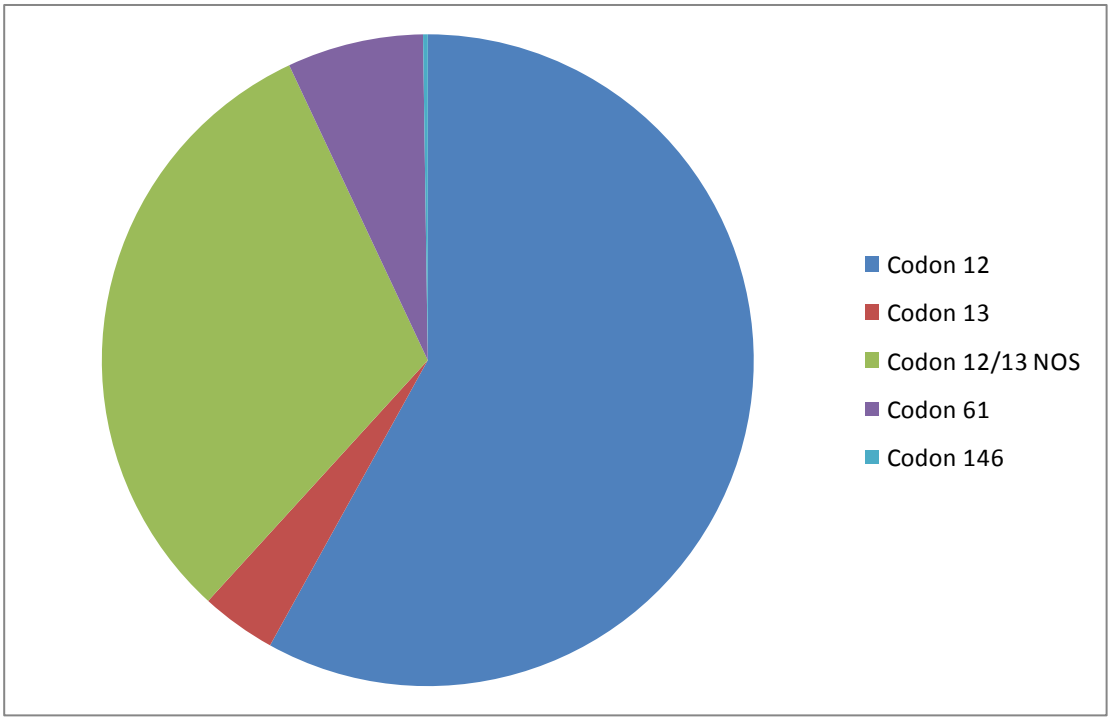

Figure 11. Distribution of *KRAS* mutations in the SMP1 lung cancer patient cohort

3.3.2 Breast cancer

In the breast cancer patient cohort, the most common histological subtype was invasive ductal carcinoma of no special type (76%), with pure invasive lobular carcinoma in 9.5% and mixed carcinoma subtypes in 6% (table 31).These findings suggest that this group was broadly representative of the overall breast cancer patient population, at least in terms of histological subtype. Of 1873 breast cancers, 43% had at least one genetic abnormality in the panel tested, 6% had 2 abnormalities, and 0.48% had three. Mutation rates were comparable to those found in other cohorts, with *PIK3CA* the most frequently mutated gene (29%), followed by *TP53* (23%), *PTEN* mutation (3.5%) and *BRAF* (0.07%) (table 32).

Table 31. Histological subtype of tumours in the SMP1 breast cancer patient cohort

Histological subtype	Total
Invasive ductal carcinoma of no special type	1424
Invasive lobular carcinoma	179
Mixed carcinoma subtypes	111
Mucinous carcinoma	25
Tubular carcinoma	35
Apocrine carcinoma	1
Medullary carcinoma	5
Metaplastic carcinoma	4
Papillary carcinoma	4
Not known	85
TOTAL	**1873**

Table 32. Gene mutation rates detected in the SMP1 breast cancer patient cohort compared to those in the scientific literature

Gene	Published prevalence of aberration	Reference	Prevalence in SMP1 cohort	Potentially clinically actionable?
PIK3CA	16-18%	[115]	29%	No
TP53	23%	[113]	23%	No
PTEN mutation	3.5%	[113]	4.8%	No
BRAF	1%	[113]	0.07%	No

PIK3CA mutations appear more common in the SMP1 cohort than in other published series to date. Analysis of the *PIK3CA*-mutant cases by histological subtype (table 33) indicates that tubular carcinoma is relatively over-represented in this group compared to the overall cohort. This is in keeping with reports in the literature that *PIK3CA* mutation is associated with breast cancers of a lower grade and with better prognosis, including tubular carcinomas, and although data on tumour grade were not available for this analysis the data may indicate a bias towards lower-grade tumours in the SMP1 cohort.

Table 33. Histological subtype of *PIK3CA*-mutant breast cancers in the SMP1 cohort

Histological subtype	Number of *PIK3CA*-mutant cases (%)	Total number in SMP1 breast cohort	Percentage of total represented in *PIK3CA*-mutant cohort (%)
Invasive ductal carcinoma of no special type	420 (76)	1424	29
Invasive lobular carcinoma	60 (11)	179	34
Mixed carcinoma subtypes	25 (4.5)	111	23
Tubular carcinoma	17 (3)	35	49
Mucinous carcinoma	3 (0.5)	25	12
Medullary carcinoma	1 (0.25)	5	20
Papillary carcinoma	1 (0.25)	4	25
Not known	25 (4.5)	85	29
TOTAL	**552**		

3.3.3 Colorectal cancer

In 1605 colorectal cancer samples, mutations in *TP53* were found in 54%, *KRAS* in 39%, *BRAF* in 10%, *PIK3CA* in 11%, and *NRAS* in 4% (table 34). Multiple gene mutations were fairly common and occurred in 33% of cases (figure 12), with the commonest being double mutated *TP53* and *KRAS* (17%), *TP53* with *BRAF* (4%) and *KRAS* with *PIK3CA* (4%). Subgroup analysis of 56 mucinous colorectal adenocarcinomas as expected enriched for *BRAF* mutations which were present in 36% and *KRAS* mutations which were identified in 43%.

Table 34. Gene mutation rates detected in the SMP1 colorectal cancer patient cohort compared to those in the scientific literature

Gene	Published prevalence of aberration	Reference	Prevalence in SMP1 cohort	Potentially clinically actionable?
KRAS	35-45%	[34] [116]*	39%	Wild-type status confers eligibility for EGFR inhibitor therapy e.g. cetuximab
BRAF	5-10%	[116]*	10%	Negative predictor of response to EGFR inhibitor
NRAS	2.2%	[118]*	4%	No
PIK3CA	10-30%	[116]*	11%	
TP53	64%	[119]	54%	

* Indicates that in the cited reference the patient population was focused on those with advanced disease, i.e. stage III-IV.

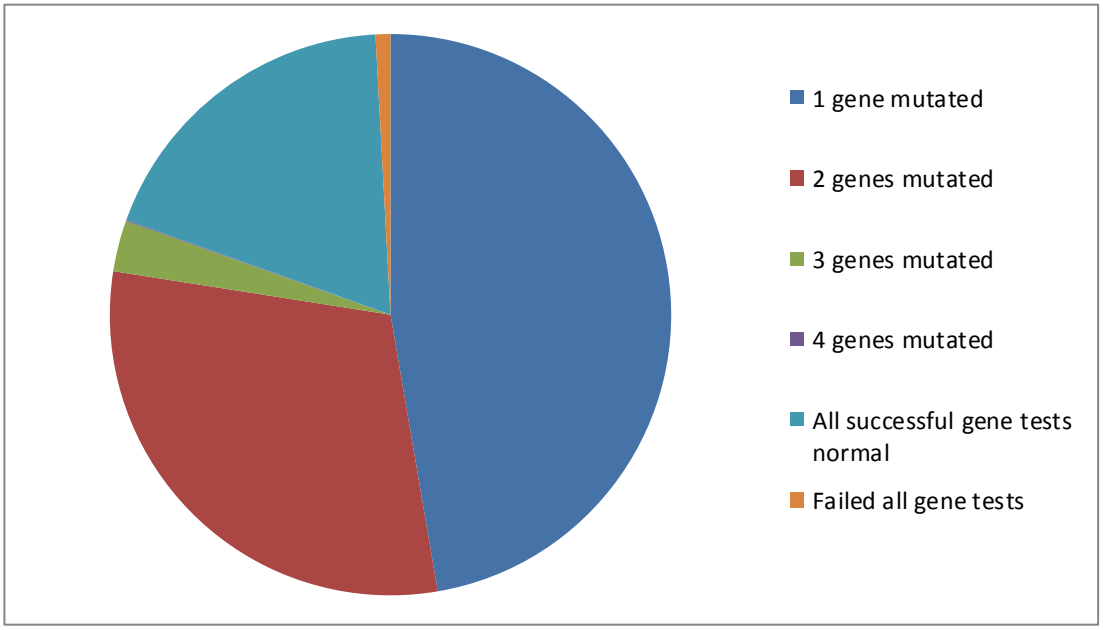

Figure 12. Mutations in tumours from the SMP1 colorectal cancer patient cohort as a whole

3.3.4 Ovarian cancer

The ovarian carcinoma patient cohort comprised mixed histological subtypes but predominantly represented serous carcinoma (figure 13). Of 557 ovarian cancer samples, 57% had a mutation in at least one gene, with *TP53* mutations in 48%, *PIK3CA* in 6.3%, *PTEN* in 4.3% and *BRAF* in 2.3% (table 35). Multiple

gene mutations occurred in 4% of cases, with the commonest being double mutated *TP53* and *PIK3CA* in 2% of tumours.

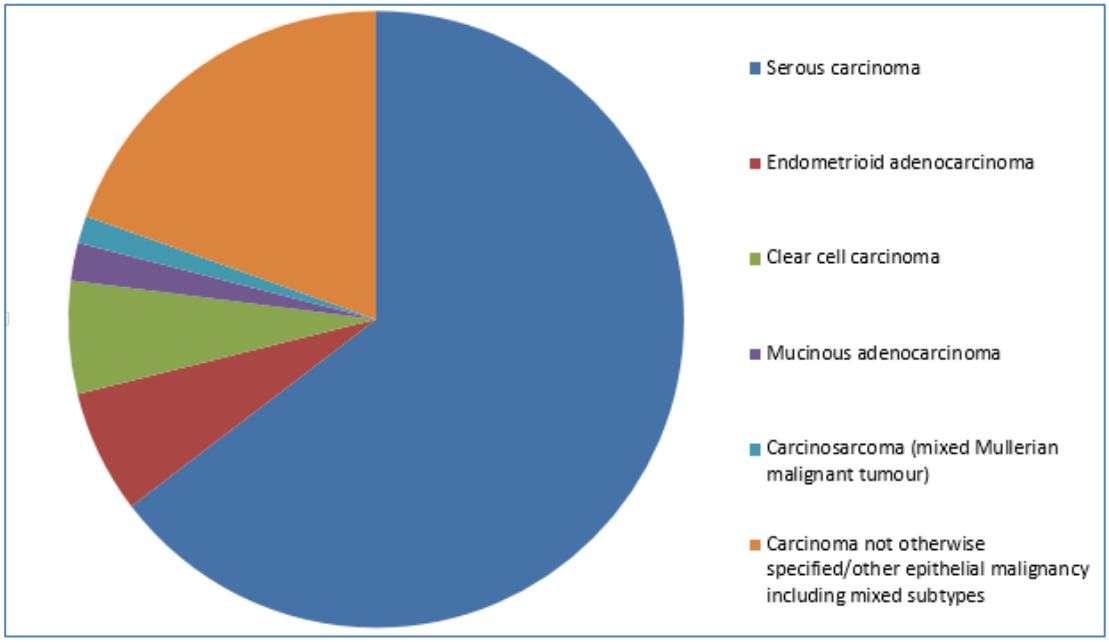

Figure 13. Histological subtype of tumours in the SMP1 ovarian cancer patient cohort

Table 35. Gene mutation rates detected in the SMP1 ovarian cancer patient cohort compared to those in the scientific literature

Gene	Published prevalence of aberration	Reference	Prevalence in SMP1 cohort	Potentially clinically actionable?
TP53	96%*	[120]	48%	No
PTEN mutation	20%	[121]	4.3%	No
PIK3CA	12%	[122]	6.3%	No
BRAF	11%	[121]	2.3%	No

*The cited reference is based on a series of high-grade serous ovarian carcinomas studied for The Cancer Genome Atlas.

Analysis by histological subtype in the SMP1 patient cohort enriched for the presence of particular mutations. In the subset of 360 patients with ovarian serous carcinoma, *TP53* was mutated in 51%, increasing to 66% if samples that failed the analysis were excluded. In the literature, *TP53* mutations are ubiquitous in high-grade serous (type II) carcinomas and although these were

by far the most common mutation detected in the ovarian cancer cohort, the lower prevalence is likely to reflect both inclusion of other histological subtypes, including possibly some low-grade serous carcinomas, in this group and also the limited scope of analysis of the *TP53* gene in SMP1 (exons 4-9 only) and comparatively low sensitivity of the technology used in each laboratory leading to missed mutations. Clear cell carcinomas are characterised by *PIK3CA* mutations (33-46%) in the literature [123] and in the SMP1 cohort (27% in 33 cases) whereas *PTEN* mutations are more common in endometrioid carcinomas (20%)[124] [125] in the literature and also in the SMP1 cohort (22% in 36 cases).

3.3.5 Prostate cancer

The 1359 prostate samples submitted were all of adenocarcinoma subtype and 39% had a *TMPRSS2-ERG* rearrangement on fluorescent in situ hybridisation (table 36), with interpretation complicated by the presence of complex rearrangements and gene copy number aberrations. Since the majority of samples submitted to SMP1 were derived from surgical resection specimens, the lower prevalence of the *TMPRSS2-ERG* fusion in our cohort is likely to represent a bias towards patients with early-stage disease who are offered radical prostatectomy. *BRAF* mutations were seen in 3.8% and *PTEN* mutations were found in 5% of samples. The co-occurrence of *TMPRSS2-ERG* rearrangement with either *BRAF* or *PTEN* mutation was seen in only 2%.

Table 36. Gene mutation rates detected in the SMP1 prostate cancer patient cohort compared to those in the scientific literature

Gene	Published prevalence of aberration	Reference	Prevalence in SMP1 cohort	Potentially clinically actionable?
TMPRSS2-ERG fusion	50%	[62] [63] [126]	40%	No
PTEN mutation	3.6%	[113]	5.2%	No
BRAF	1.4%	[113]	1.2%	No

3.3.6 Malignant melanoma

The histological subtype of 535 malignant melanoma samples sequenced through SMP1 are shown in table 38. No mutations were found in the five acral

melanomas. Three of the ten (30%) lentigo maligna melanomas represented in this cohort had codon 61 *NRAS* gene point mutations, in contrast to the reported propensity for *KIT* gene mutations in melanomas arising in sun-damaged skin since no *KIT* mutations were found in these samples. Two further lentigo maligna melanomas were wild-type for *NRAS* but were found to have the BRAF V600E mutation, giving an overall mutation rate for the genes and regions analysed of 50% in this small subset of ten cases. One of the dual mutation tumours was a spindle cell melanoma with *PIK3CA* H1047R and *BRAF* V600E mutations.

Table 37. Histological subtype of SMP1 malignant melanoma samples

Malignant melanoma by histological subtype	Number of cases
Melanoma not otherwise specified	326
Superficial spreading melanoma	88
Nodular melanoma	54
Lentigo maligna melanoma	10
Spindle cell melanoma	6
Acral melanoma	5
Amelanotic melanoma	2
Desmoplastic melanoma	1
Epithelioid melanoma	1
Not stated	42
TOTAL	**535**

In 535 melanomas there were 43% *BRAF* mutants, 23% *NRAS* mutant and only 2.4% with double abnormalities (table 35). 28 pairs of melanoma samples sent for analysis showed 100% concordance in *BRAF*, *NRAS* and *KIT* gene mutations between the two samples. The mutation frequencies are broadly comparable to the published literature including a meta-analysis of studies including 4493 patients and reporting mutation characteristics and associations in malignant melanoma[127]. This meta-analysis did report that *BRAF* and *NRAS* mutations are mutually exclusive, however in our cohort there were samples from two patients in which both *BRAF* codon 600 mutation and *NRAS* gene mutations were detected (*BRAF* V600E with *NRAS* G62E and *BRAF* V600K with *NRAS* G13C). In keeping with previously published data, 94% of the confirmed *BRAF* mutations in 232 cases were V600E (excluding a further two cases where the Roche 4800 cobas® test was used, since this is not able to discriminate between V600E and other codon 600 mutations).

Table 38. Gene mutation rates detected in the SMP1 malignant melanoma cancer patient cohort compared to those in the scientific literature

Gene	Published prevalence of aberration	Reference	Prevalence in SMP1 cohort	Potentially clinically actionable?
BRAF	41%	128*127	43%	Yes – BRAF and MEK inhibitor therapy
NRAS	18%	128*127	23%	No
KIT	Less than 5%	57 129	1.3%	
PIK3CA	3-6%	130 131	1.5%	

3.4 Turnaround times for mutation analysis

Recognising the need for clinically relevant turnaround times, the aim during SMP1 was for a result to be available within 15 working days from sample receipt at the TH. This proved difficult to achieve with 65% of samples taking longer than 16 working days to report, 24% returned in 10-15 days and only 10% returned in 6-10 days (figure 14). Repeat testing of failed genes in a particular sample often led to a delay in returning a result, so from April to May 2013, all labs were asked to adopt a policy of not re-testing samples. During this period there was an improvement in turnaround times (figure 15), with a large reduction in the number of reports returned after 16 days to an average of 36% of samples. The number of samples returned in 6-10 days doubled to an average 21%, but only 1% of results were returned in less than 5 working days. Other factors contributing to longer reporting times included a higher number of genes and the use of different testing modalities within a tumour type panel, since the report was only issued when a result was available for all genes in the panel. Batch delivery of samples from clinical hubs led to unpredictable workload and poor sample quality also increased the time taken for analysis when repeat testing of failed samples was performed.

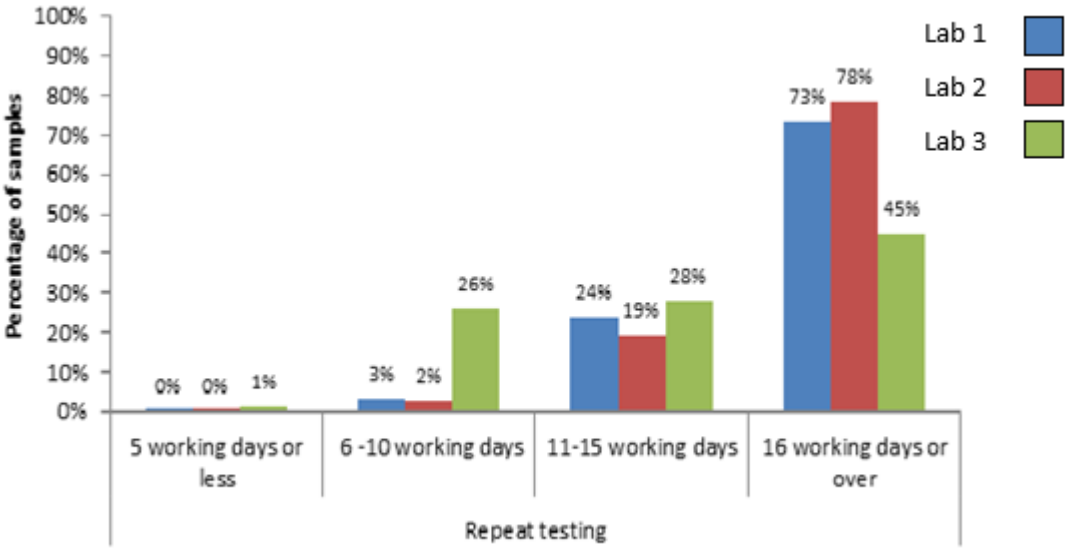

Figure 14. Turnaround time with repeat testing

Turnaround time for each sample, from receipt of sample at the TH to the analysis results for all genes in the panel test being returned to the CH. The initial repeat testing policy varied between laboratories: laboratories 1 and 2 repeated the analysis of a sample that failed the original analysis and laboratory 3 did not.

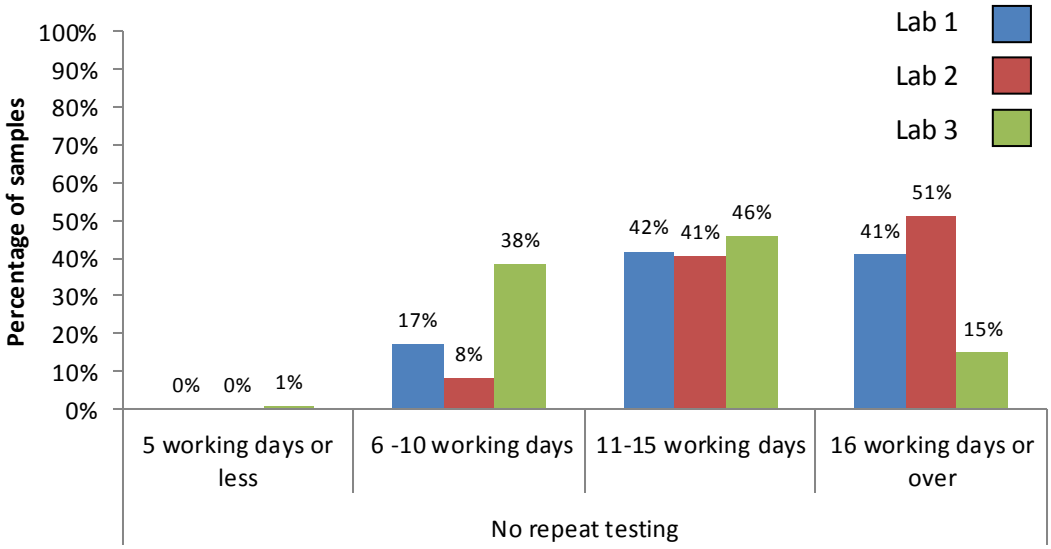

Figure 15. Turnaround time without repeat testing

Turnaround time for each sample, from receipt of sample at the TH to return of results for all genes in the panel to the CH, during a two month period where repeat testing of failed samples was not performed by any of the three THs.

3.5 Failure rates

Failure rates were closely scrutinised as part of SMP1 and test failures were categorised as either whole or partial gene failures for all genes tested within the

SMP1 panel. For the scope of SMP1 analysis, the proportion of samples failing all gene tests was low for each tumour type at 0.9-5.3% (table 29). Partial gene failure rates by gene and TH are shown in figure 16. Partial failure rates also varied by tumour type, with breast, ovary and prostate demonstrating the highest failure rates (figure 17). Contributory factors include differences in the gene panel and scope of each gene test (how many exons/codons were analysed), the number of gene tests required and variations in sample quality and quantity from different tumour types and originating clinical sites.

The proportion of partial failures varied by gene between the three THs. Several genes had almost no recorded partial gene failures, for example *BRAF* since this was a hotspot test covering a small number of codons, and *EML4-ALK* and *TMPRSS2-ERG* because these were both analysed by FISH. In contrast, screens of multiple exons of the tumour suppressor genes *TP53* and *PTEN* generally showed higher partial failure rates, due to challenges associated with analysis of FFPE material, and the larger size of the genetic region analysed increasing the likelihood of one of the fragments failing (figure 18). Due to the range of sample performance incorporated in the 'partial fail' category, it was still possible for a partially failed sample to yield a clinically relevant mutation result. Variations in failure rates in SMP1 are attributed to a number of different factors including variations in sample quality, case mix of tumour types, the use of different techniques and fluctuation in test performance. There were also differing approaches to repeat testing of failed samples and the designation of a failed test between different laboratories, though attempts were made to standardize these through collaborative working and expert consensus as the initiative progressed. Failure rates by tumour type and clinical hub are presented in chapter 4 for comparison to cellular pathology department sample handling data.

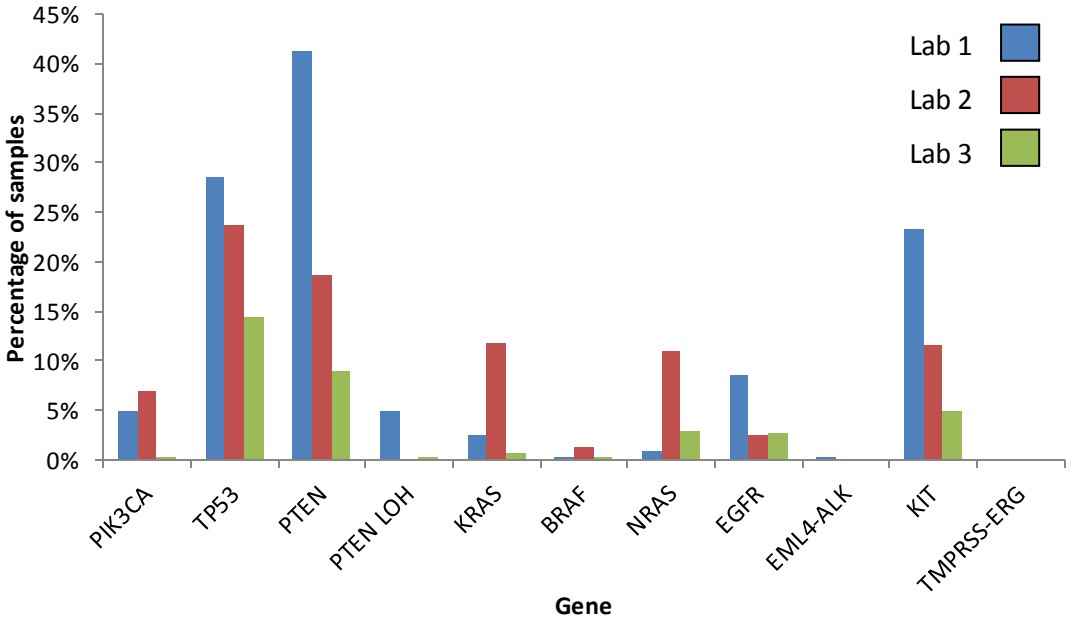

Figure 16. Partial failure rate by gene

The percentage of gene tests for each gene classified as a partial test failure, including anything between failure of a single exon/ codon/ fragment and all but one exon/ codon/ fragment. This analysis does not include *DDR2* since it was not analysed throughout the entire duration of the programme.

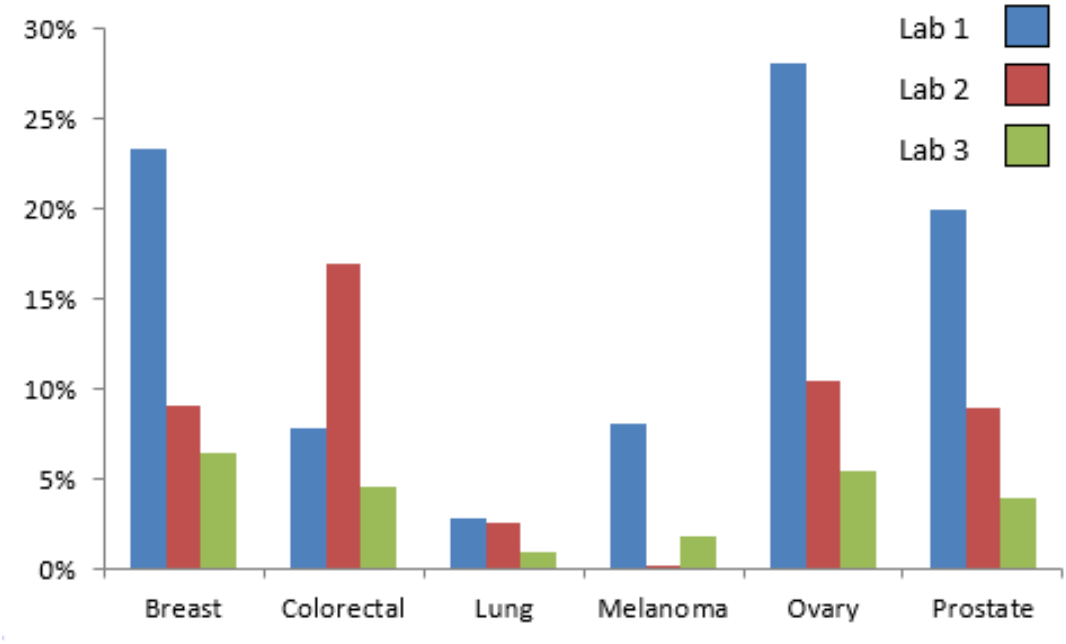

Figure 17. Partial failure rate by laboratory and tumour type

The percentage of gene tests for each tumour type classified as a partial test failure, including anything between failure of a single exon/ codon/ fragment and all but one exon/ codon/ fragment.

3.5.1 EQA scheme

In order to assess quality and reproducibility of genotyping, reporting and interpretation of SMP1 results, EQA schemes for each of the six tumour types were delivered by UK NEQAS for Molecular Genetics. Each tumour-specific EQA scheme involved the distribution of three samples per EQA round, with each sample requiring analysis for a specific gene panel. Three rounds of EQA were performed in SMP1 (54 samples per laboratory involving a total of 162 samples, table 39). For each round, one lead TH was responsible for providing a list of previously reported CRUK samples to UK NEQAS for selection and sourcing of material from suitable cases directly from the relevant CH. As the selected cases had already been analysed in the lead TH, results for that EQA round could be submitted without the need for repeat analysis within the lead TH. Each of the THs already participated in existing UK NEQAS schemes for *KRAS, EGFR, KIT* and *BRAF* molecular testing, therefore these genes were excluded from analysis in the SMP1 EQA. The lung tumour analysis was only included in EQA round 1, as all laboratories thereafter participated in the newly available EQA scheme for the *EML4-ALK* fusion, meaning that all genes in the SMP1 lung panel were covered by existing schemes. All results were submitted to UK NEQAS within six weeks and were scored by the EQA provider. The results were returned to each TH, with any discrepancies highlighted for review. Following investigation of any non-concordant results, reports were issued by UK NEQAS.

Table 39. UK NEQAS laboratory sample exchange and analysis results

Laboratory	TH 1			TH 2			TH 3		
EQA round	1	2	3	1	2	3	1	2	3
Breast	16/16 (100%)	14/14 (100%)	24/24 (100%)	16/16 (100%)	14/14 (100%)	24/24 (100%)	16/16 (100%)	14/14 (100%)	22/22 (100%)
Colorectal	18/18 (100%)	18/18 (100%)	18/18 (100%)	16/18 (89%)	18/18 (100%)	18/18 (100%)	16/18 (89%)	12/14 (86%)	18/18 (100%)
Lung	6/6 (100%)	-	-	6/6 (100%)	-	-	6/6 (100%)	-	-
Melanoma	12/12 (100%)	12/12 (100%)	12/12 (100%)	12/12 (100%)	12/12 (100%)	12/12 (100%)	12/12 (100%)	12/12 (100%)	12/12 (100%)
Ovarian	20/20 (100%)	24/24 (100%)	22/24 (92%)	22/22 (100%)	24/24 (100%)	22/24 (92%)	22/22 (100%)	22/24 (92%)	24/24 (100%)
Prostate	18/18 (100%)	10/10 (100%)	18/18 (100%)	18/18 (100%)	8/8 (100%)	18/18 (100%)	12/12 (100%)	12/12 (100%)	16/16 (100%)
TOTAL	100%	100%	98%	98%	100%	98%	98%	96%	100%

Summary of scores obtained for each EQA round. Individual scores for each tumour type EQA round are given along with percentages. The total scores for each laboratory may differ owing to samples that failed analysis at a particular laboratory being excluded from the scoring rather than points being deducted.

3.6 Discussion

The Stratified Medicine Programme phase one pilot study has demonstrated that the approach used is feasible and highly acceptable to patients. In addition to establishing the infrastructure for centralised molecular testing of tumour samples in support of clinical decision-making in cancer care, the programme activities demonstrated that embedding routine generic consent for sample donation at an appropriate point in the clinical care pathway is an effective way of advancing patient participation in research. This approach has been taken by several Clinical Hubs during SMP1, with teams at both sites incorporating information about research use of surplus tissue in the main hospital surgical procedure consent form. These documents, and the process of consenting the patient for tissue donation at the same time as consent is taken for surgery, were approved by an NHS REC.

A 96% consent rate shows that patients find tissue donation for research an acceptable process. Comments from our patients suggest that, in some cases, people are surprised that this does not happen automatically. Diversity of practice exists across the health system and, in order to optimise the process, sharing of good practice is required. Biobanking of multiple different sample types is a widespread and routine activity in the modern NHS, and standardization of processes related to the provision of information to patients about this activity, as well as recent efforts by the Confederation of Cancer Biobanks Harmonization project (http://ccb.ncri.org.uk/) to establish sample and operating standards in the area, are essential. Such endeavours will help ensure that high quality samples continue to be made available for research into diseases and their treatments for the ultimate future benefit of patients.

The molecular results of phase one of the Stratified Medicine Programme provide an insight into the mutational epidemiology of tumours occurring in a large cohort of patients in the United Kingdom. Since this is a relatively unselected population, it seems reasonable to interpret the results as representative of the wider population and thus they can be used to provide a baseline estimate of the prevalence of clinically actionable genetic findings for the UK population, invaluable information to help with commissioning molecular diagnostic service provision for cancer patients. The majority of samples tested in this cohort (52%) had an aberration in one gene, despite the limited scope of genetic analysis in SMP1. The proportion of samples failing all gene tests was low at 3%, indicating that this type of analysis is feasible in formalin-fixed paraffin-embedded tissue. Phase Two of the Stratified Medicine Programme is now underway and is providing molecular analysis of lung tumour specimens as pre-screening for determining patient eligibility for a multi-arm trial of novel therapeutics, the National Lung Matrix Trial.

In addition the SMP1 data has been used to illustrate the practical challenges, different factors and trade-offs inherent in delivering high-quality molecular analysis with acceptable reporting turnaround times and failure rates.

Chapter 4: Cross-sectional analysis of cellular pathology department specimen handling

4.1 Results for general sample handling

Responses were received from 15 out of 20 cellular pathology departments participating in the programme, a response rate of 75%. Not all departments dealt with all specimen types covered and all the molecular analysis for the programme was performed by the three central laboratories (THs) in Birmingham, Cardiff and Sutton.

4.1.1 Fixation

All laboratories used formalin as the standard tissue fixative, although there were minor differences in the chemical composition, mainly in the use of buffers (figure 18). One laboratory reported routine use of formal saline, which is formalin with the addition of 0.9% normal saline to create an isotonic solution. Of note, this fixative does not contain phosphate buffer and therefore allows formation of formic acid which is likely to accelerate nucleic acid degradation in the tissue. This laboratory has subsequently switched to using ready-made phosphate buffered formalin (personal communication). pH checking of formalin was not routinely performed in the laboratories surveyed, although an isotonic solution buffered to pH 7.2-7.4 is recommended in order to avoid cell shrinkage and maintain tissue ultrastructure and regular checking of formalin pH is now stipulated as a requirement under ISO15189 accreditation of diagnostic laboratories.

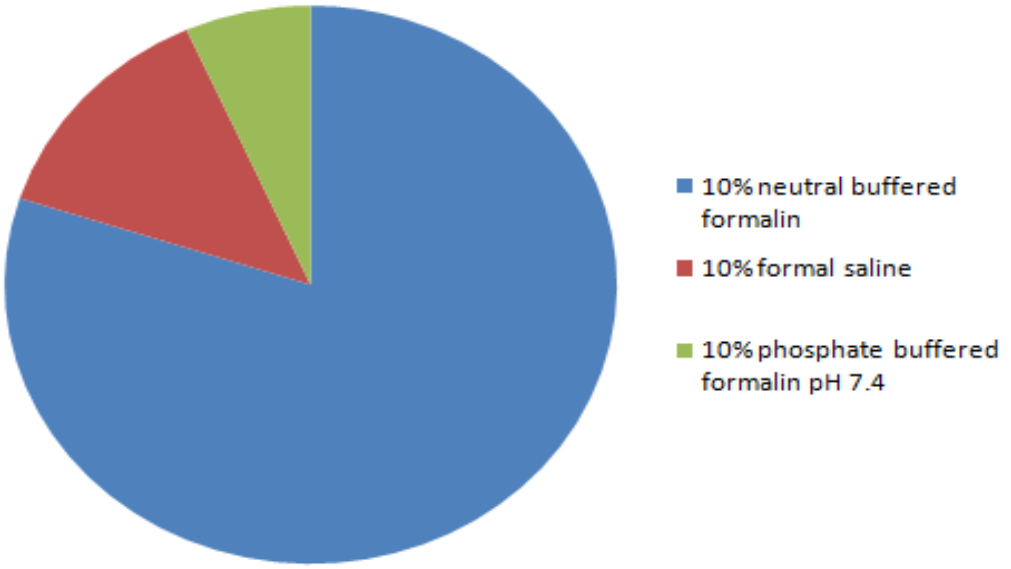

Figure 18. Fixative used by each cellular pathology laboratory

4.1.2 Tissue processing

A number of different automated tissue processing machines were in use and there were between 1-6 machines per laboratory. 14 out of 15 centres (93%) used xylene as a clearing agent. One also used isopropyl alcohol (IPA) just for prostate megablocks, one used IPA for all processing and a further laboratory had xylene or xylene-free processing available if required for particular situations.

The question asking about processing programmes was designed to take into account the requirements of different specimen types, but even allowing for this processing times showed marked variation as indicated in table 40.

Table 40. Range and average processing times for different specimen types across laboratories

Time (hours)	Biopsies	Routine processing	Large/ mega blocks	Fatty tissue
Range	1-15	9-20	14-48	19-63
Mean	8.5	12.5	32	36

4.1.3 Tissue economy: biopsies

Small biopsy samples are routinely examined at different tissue levels, with multiple sections taken several hundred micrometres apart in order to try and ensure that sufficient tissue content of the biopsy material present has been represented to the reporting pathologist. In diagnostic samples where tumour material may be limited, there is potential for tissue wastage at this stage if the intervening sections are discarded. Conversely, reflex cutting and saving of sections in anticipation of molecular analyses can be used to maximise the available tissue and avoid the need to 're-face' the block on the microtome again, which is widely regarded as being a major source of wastage. A further reason for saving intervening sections on glass slides is in anticipation of the requirement for immunohistochemistry, for example in breast cancer for evaluation of hormone receptor and HER2 status, or diagnostic prostate biopsies where basal cell immunohistochemistry may aid assessment of small foci of atypical glands showing appearances suspicious for carcinoma. There are cost and physical space implications to saving spare sections on glass slides, since these might never be required for use, and each glass slide costs approximately 76 pence (SuperFrost™ Plus slides, Thermo Scientific Gerhard Menzel, Ulm, Germany) with the cost increasing to £1.39 if sections are kept on coated and charged slides suitable for tissue adherence during the antigen retrieval and washing steps required for immunohistochemistry (e.g. SuperFrost™ Plus Adhesion slides, Thermo Scientific Gerhard Menzel, Ulm, Germany). Laboratories varied in whether they saved all intervening tissue in different tumour types (table 41).

Table 41. Percentage of respondents stating that intervening sections between levels were saved when cutting biopsy specimens from different tissue types

Tissue type (number of responses)	Spare sections cut and stored (% sites)	All intervening sections between levels kept (% sites)
Breast (10)	70	30
Colorectal (11)	27	27
Lung (14)	65	50
Ovary (11)	36	18
Prostate (12)	83	33

4.1.4 Microtomy

There is increasing recognition of the risk of cross-contamination of DNA in the cellular pathology department involving specimens that may subsequently be sent for molecular analysis. The analysis revealed that some laboratories have started to develop specific processes for microtomy when cutting tissue sections for molecular work as follows:

- blocks cut first thing in the morning (n=1)
- new microtome blade for each case (n=5)
- microtome blade cleaned between every case or new part of blade used (n=3)
- dedicated microtome for preparing sections (n=6)
- sending block away to specialist centre for cutting (n=1)
- specialist biomedical scientist trained to prepare sections (n=1)

There was variation in blade cleaning practices. It was discovered that one laboratory was using a chlorine/bleach containing substance to clean the microtome blade between cases, a process which is likely to be detrimental since bleach destroys DNA. All respondents reported that it was routine

practice in their department to store blocks and slides at ambient/room temperature.

4.1.5 Correlation to failure rates

Failure rates by tumour type and clinical hub are presented in figure 19 and show a high degree of variability. Apart from the possible contribution from differences in routine sample handling uncovered through the data presented above, there are number of other potentially confounding factors underlying these findings. These include characteristics of the originating specimen, such as whether it was derived from a biopsy or resection, the time elapsed since processing and whether it was processed within that laboratory or was a referred block from a different pathology department (this was a particular issue for clinical hub 8, CH8, which is a centre with a large referral practice). In addition, the fact that these samples were tested in three different molecular genetics laboratories adds another variable into the mix, especially in view of the different methods and approach to repeat analysis used in these laboratories. Given all these variables it is not surprising that it is difficult to identify trends in this data to correlate to sample handling practices.

The majority of the samples submitted for SMP1 analysis were derived from surgical resection specimens, meaning that tissue availability should not be the limiting factor, but marked differences were found in the success rates of genetic analysis on different types of specimen as well as the same tumour types from different NHS sites. Given that the quantity of DNA should not have been the limiting factor in surgical resection specimens, fixation is the factor most likely to account for test failures. Adequate fixation can be problematic in large resection specimens that are received in the pathology laboratory intact and not incised soon after immersion in formalin, to allow the formalin to penetrate the deeper tissues. This is particularly important for adequate tumour fixation since in order to achieve clear margins the tumour is invariably central within the specimen, and surrounded by normal uninvolved tissue. In most organs excised whole there is also likely to be a capsule around the exterior surface which will provide a further barrier to formalin penetration.

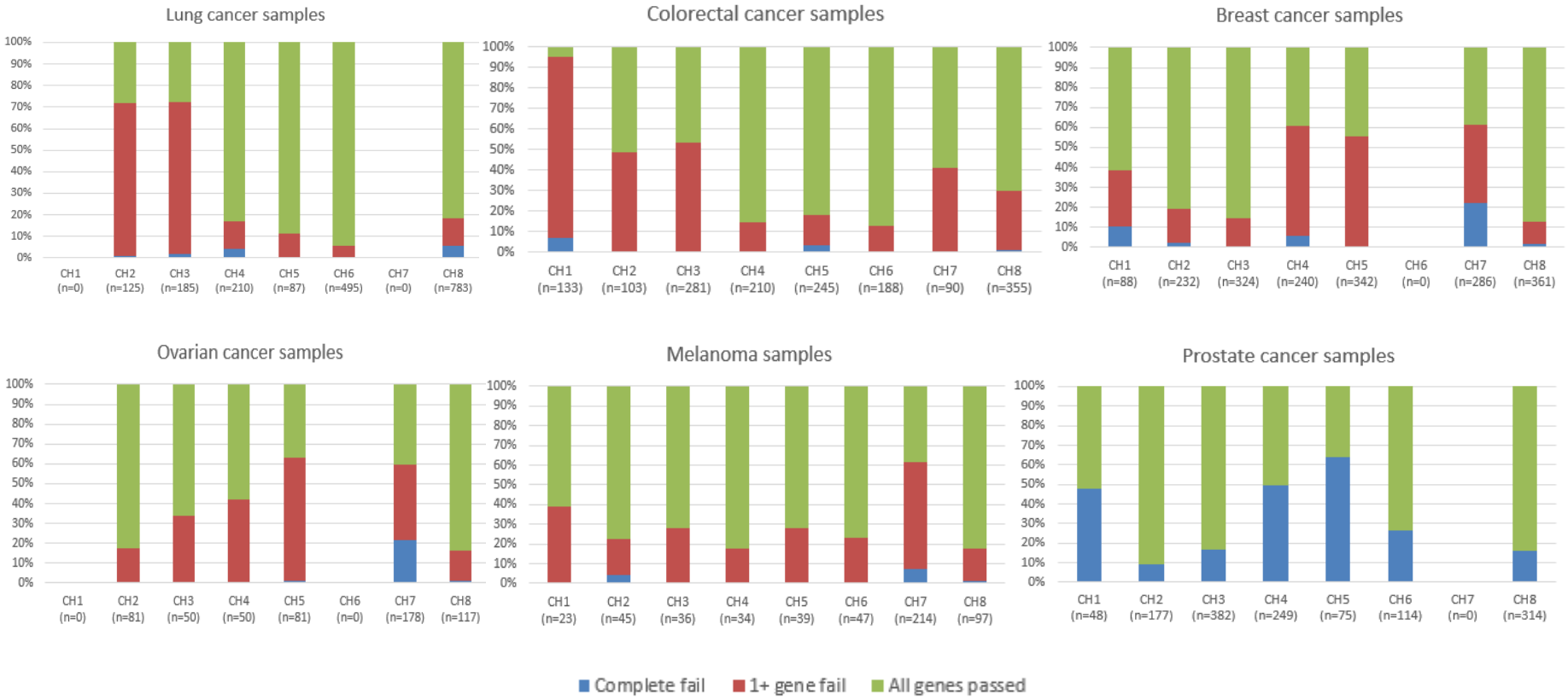

Figure 19. Sample test performance on sequencing-based assays by tumour type and clinical hub

No 1+ gene fails were seen for prostate cancer since only one sequencing-based test was applied to this tumour type so failure of this gene test would be classified as a complete fail.

4.1.6 Discussion

This pilot cross-sectional analysis of a small number of NHS pathology departments has demonstrated variation between sites in a number of different aspects of specimen handling. These laboratories are already involved in preparing and submitting a wide range of tumour samples for mutation analysis as part of routine clinical care and also for the CRUK-SMP. Differences in success rates for subsequent molecular analysis between different genes, centres and tumour types have also been demonstrated. The underlying reasons for these variations have been explored through the collaborative multidisciplinary network of researchers involved in the programme. There is a need to establish the baseline position of specimen handling processes in UK cellular pathology departments and optimise this, in order to prepare for more widespread use of predictive molecular analysis as part of the delivery of stratified cancer medicine. Multidisciplinary and inter-departmental collaboration is required to establish and implement optimised protocols for specimen and nucleic acid preservation.

4.2 Cross-sectional analysis of handling of endobronchial ultrasound-guided lung cancer samples

4.2.1 Results for EBUS sample handling

A request for information about sample handling was circulated to cellular pathology laboratory contacts at all twelve sites participating in SMP2 at the time and responses were received from all (100% response rate). The numbers of samples received by each department per week was variable but the majority of the sites received more than 5 samples per week (figure 20). The self-reported estimated rate of obtaining diagnostic material at the eight sites with available data was in excess of 90% (table 42). All sites prepared cell blocks from every sample, nine after direct cytology preparations and three stated that they had recently started using all material for the cell block. At two sites there was a current arrangement for assistance with preparation of the

slides by a biomedical scientist and rapid on-site examination of the samples by a pathologist, and at a further six this service had been available in the past but had since been stopped (table 43). Cytology processing systems, preservatives and fixatives used for handling residual fluid samples varied by site (table 44). The two most popular liquid-based cytology (LBC) systems use a different fixative base in their proprietary solutions. For ThinPrep (Hologic, Massachusetts, USA) this is methanol-based (PreservCyt) and for SurePath (BD, Oxford, UK) this is ethanol-based (CytoRich™ Red preservative fluid).

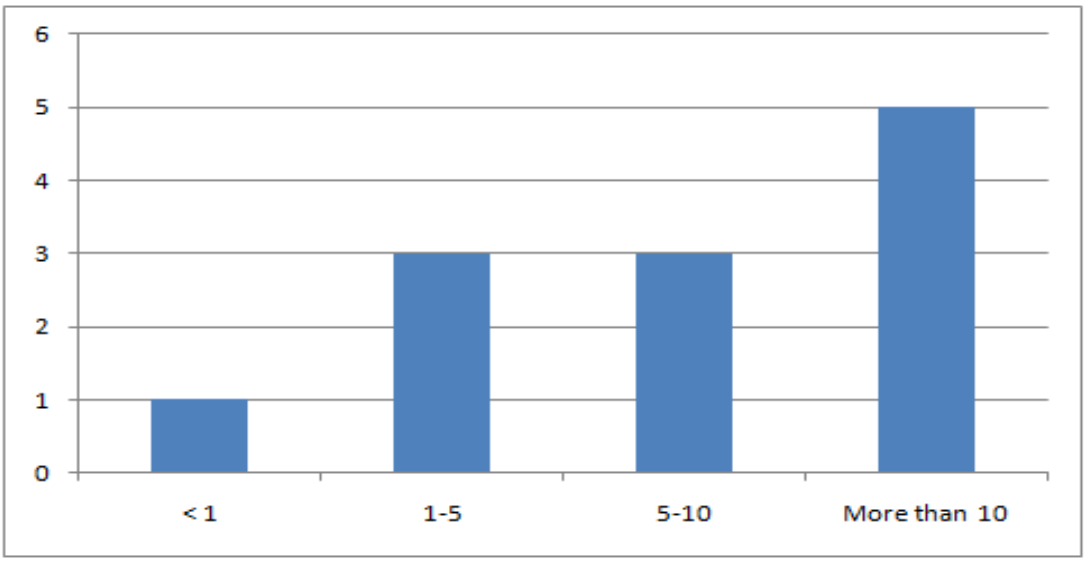

Figure 20. Estimated number of EBUS/EUS samples received by each laboratory per week for suspected lung cancer

Table 42. Reported sample adequacy rates

% of samples containing sufficient material for cytological assessment	Count of responses
95% or more	6
90% or more	2
Audit in progress	2
Not known	2

Table 43. Pathology staff support

Staff in attendance	Current	Past
Medical and scientific/technical staff	2	3
Medical staff only	0	1
Scientific/technical staff only	0	3
None	3	

In six departments a Papanicolou-stained preparation was prepared on a fixed sample only, and in three departments both DiffQuik staining on air-dried preparations and Papanicolou staining were performed.

For the two services where pathology input was currently provided at the time of the procedure, in one site two small smears were prepared for DiffQuik and Papanicolou stains for each pass of the needle until the material obtained permitted a provisional diagnosis and for the other one DiffQuik preparation was made only, to confirm technical adequacy of the sample.

Table 44. Sample handling of the residual fluid sample in different laboratories

Handling of residual fluid sample	Response count
LBC –ThinPrep	2
LBC - Cytosed	1
LBC – SurePath	0
Conventional/cytocentrifugation	4
Immediate formalin processing	3

LBC = liquid-based cytology

Respondents from all laboratories stated that they attempted a cell block and H&E on every EBUS specimen. All were in agreement that less than 5% of samples contained insufficient material to form a clot or cell pellet, i.e. cell block preparation was possible in more than 95% of specimens. 2 laboratories

used agar embedding to form the cell block and one has trialled marking edges with India ink to delineate the sample on the slide. One laboratory was trialling an automated cell block processing system.

4.2.2 Discussion

There is variation in practice and sample handling between EBUS services provided at different sites, reflecting the fact that this diagnostic service has developed in an opportunistic and piecemeal manner, often driven by individual enthusiasts keen to develop a new service. The different ways in which the services have developed is also representative of the extent to which pathologists have been involved and also their particular areas of sub-speciality interest. For example, sites at which a pathologist with subspecialist expertise in cytopathology has been instrumental in setting it up may rely more on direct preparations made immediately in the procedure room and this may also be driven by the bronchoscopist wanting a provisional diagnostic opinion and/or confirmation that diagnostic material has been obtained. The cross-sectional analysis data suggests that there is an emerging trend towards foregoing cytology preparations from EBUS samples. This may facilitate preservation of more material available in a cell block for morphological, immunohistochemical and molecular analysis.

4.3 Overall conclusions from cross-sectional analysis data

The findings of both cross-sectional analyses, carried out in different geographical and practice areas of a diagnostic pathology service, indicate that supposedly 'standard' operating procedures for a given laboratory process can hide a multitude of variables and differences in practice. This need not be an issue unless it impacts on the output of the process, but due to the multitude of variables impacting on the preparation of each specimen it has been difficult to tease out strong associations or prove the downstream impact. Laboratory quality assessment schemes in the United Kingdom have evolved to assess the concordance of results or outcome of a procedure, however with the adoption of the ISO15189 standard there is more of a focus on uniformity of process and standardisation of approach.

Chapter 5: Results of Experimental Work on Pathologist Assessment of Tumour Content and PAXgene® Tissue Fixation System

5.1 Results of Experimental Work on Pathologist Assessment of Tumour Content

An invitation to participate was sent out to all 24 members of the SMP2 pathology working group. Ten pathologists participated in the scoring of the scanned cases, a response rate of 42%, though not all pathologists submitted a score for all cases through the online system (table 45). Pathologist 1 was at the time a senior specialty registrar in cellular pathology and pathologist 10 is a molecular geneticist by background who has received training and developed expertise in tumour content assessment. The remainder of the assessors were consultant histopathologists, six of whom have sub-specialist expertise in thoracic pathology. The characteristics of the cases and number of scorers are shown in table 45, with accompanying images in figure 21.

Table 45. Characteristics of cases and number of assessors for each case included in the online slide scoring assessment

Case	Sample Type	Tissue	Histological subtype	Number of pathologist scores for whole section	Number of pathologist scores for annotated section	Annotated area present	Leeds TCD data available
1	Cytology cell block	Pleural effusion	ADC	9	N/A	No	Yes
2	Bronchoscopic biopsies	Bronchus/ lung	SCC	10	8	Yes	Yes
3	Bronchoscopic biopsies	Lung	SCC	9	8	Yes	Yes
4	Percutaneous core biopsies	Lung	SCC	10	9	Yes	Yes
5	Surgical thoracoscopic biopsies	Pleura	ADC	9	8	Yes	No
6	Bronchoscopic biopsies	Bronchus/ lung	SCC	10	9	Yes	Yes
7	Bronchoscopic biopsies	Bronchus/ lung	SCC	10	9	Yes	Yes
8	Percutaneous core biopsies	Lung	SCC	9	N/A	Yes	Yes
9	Bronchoscopic biopsies	Bronchus/ lung	SCC	10	8	No	Yes
10	Surgical thoracoscopic biopsies	Pleura	ADC	10	10	Yes	No
11	Bronchoscopic biopsies	Bronchus/ lung	SCC	Excluded from analysis due to poor quality of scanned image		Yes	No

ADC = adenocarcinoma; SCC = squamous cell carcinoma; TCD = digital tumour content determination.

Figure 22 displays the individual pathologist responses against the mean average of all pathologists and the numerical value generated by the Leeds digital tumour content determination. In every case, the 'gold standard' digital tumour content determination value falls within the range of pathologist responses, although it is usually at the lower end in assessment of both whole sections and annotated areas, suggesting that pathologists tend to over-estimate tumour percentage content in tissue sections. Table 46 displays data for whole slide assessment by individual pathologists in each case and table 47 displays the data by pathologist for assessment of the annotated area only.

Inter-observer agreement has been assessed using the kappa statistic, k (described in table 24). This is a method of expressing the level of agreement taking into account that expected purely by chance. This is commonly used in the pathology literature to determine consistency of assessment of morphological features and diagnosis between pathologists, particularly as part of external quality assessment schemes [132]. Unweighted kappa scores between multiple observers rely on exact agreement, giving no credit for being close (e.g. one category out versus three categories out) and therefore the chance of agreement reduces with a higher number of categories.

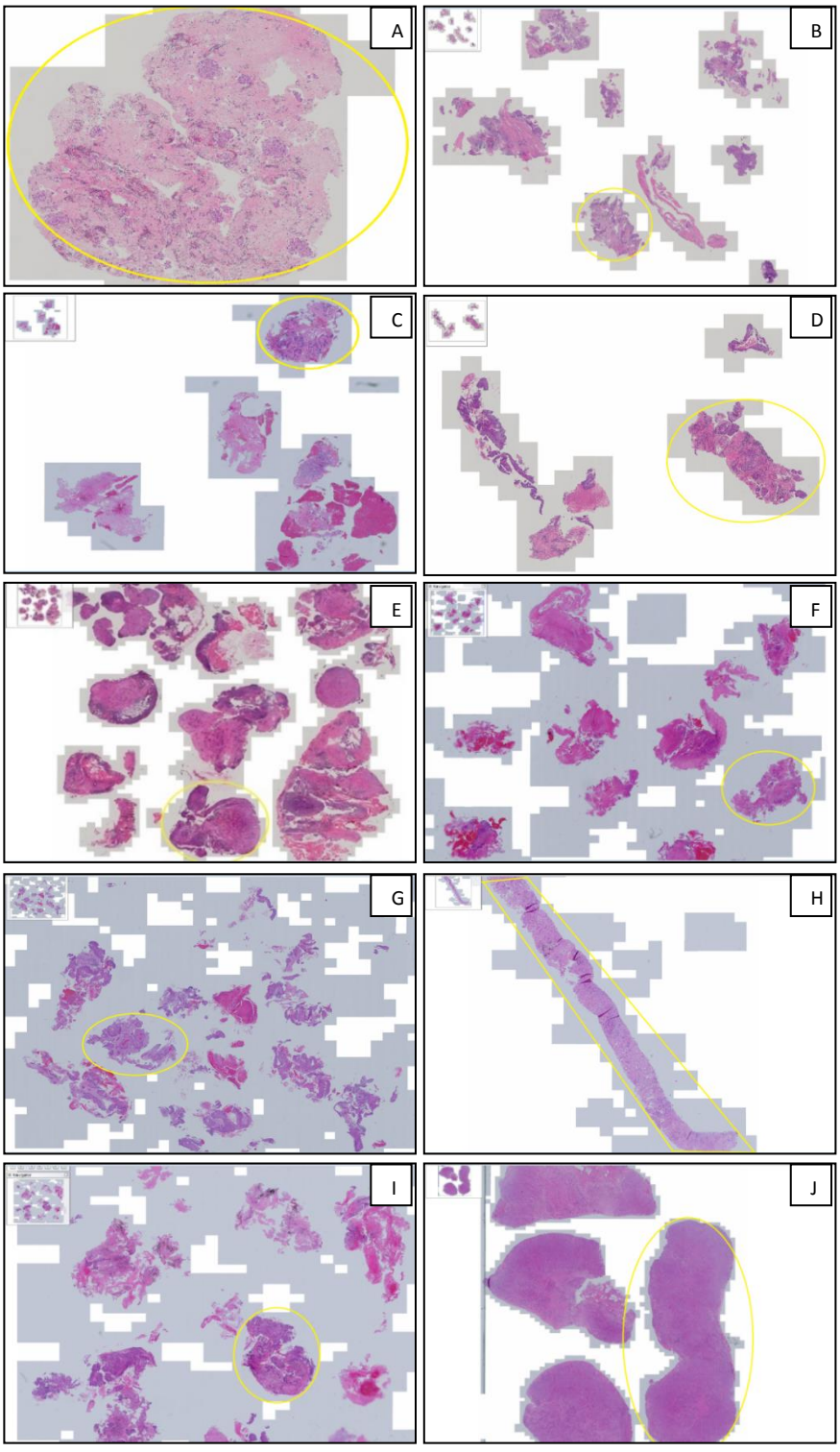

Figure 21 Images of scanned slides

A-J are cases 1-10 respectively. The annotated area to be assessed for each case is outlined in yellow. In case 1 (image A) and case 8 (image H) the yellow outline encloses the whole sample. No second value for the annotated area was requested in these cases since the distribution of the tumour meant that in routine practice it is likely that DNA would be extracted from the whole section and macrodissection for tumour enrichment would not be performed.

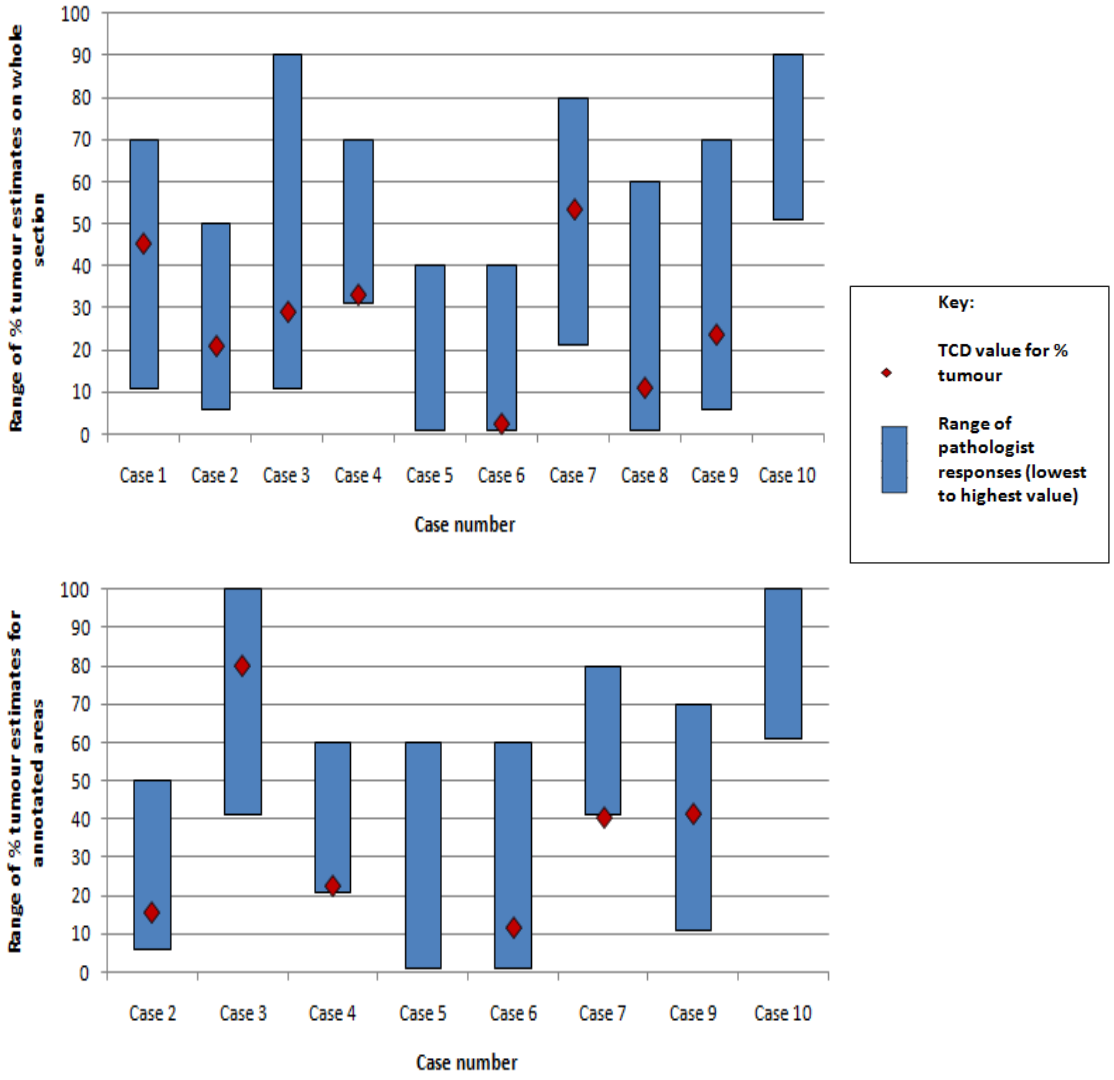

Figure 22. Range of pathologist responses for tumour content compared to digitally determined tumour content value

The upper figure is for assessment of whole slides and the lower figure is for assessment of the annotated area only.

The results for the individual cases have all been plotted on graphs with a y-axis from 0-100 to facilitate direct comparison between cases. There is considerable variation in the pathologists' opinions on tumour content assessment, with scores for whole sections varying by up to 90 percentage points and scores for annotated parts of the section showing less variability but still showing pathologist assessments differing by up to 60 percentage points.

Averaging the performance across all the pathologists gives a result that is closer to the result for digital tumour content assessment, taken as the 'gold standard', but this is as expected according to the general rule of repeatability, where taking the average of several responses averages out the measurement error, with the effect of approaching the true value. The implication of these results is that showing the same case to ten different pathologists and taking the average of their responses will give a result that approximates fairly closely to digital tumour content determination, however this is unlikely to be a feasible approach or efficient use of pathologists' time in practice.

The kappa statistics for the pathologists are low across the board. Negative kappa scores are particularly concerning since they are indicative of performance that is worse than that expected by chance with systematic disagreement in some cases.

The intraclass correlation coefficient (ICC) scores for assessment of whole slides (table 46) are generally low but two assessors stand out with a higher score. Pathologist 3 has a marginally higher score but a broad 95% confidence interval whereas assessments by pathologist 8 appear more consistent with a narrower 95% confidence interval. On questioning pathologist 8, there does not seem to be anything unusual or different about the method they use for tumour content assessment or any specific aspect of their technique that could be shared or replicated for routine practice by other pathologists. Pathologist 8 is an experienced consultant pathologist with sub-specialty expertise in lung pathology. It is not possible to determine from this data whether this pathologist's apparent aptitude for the tumour content assessment is a result of their experience and expertise or a difference in their cognitive approach to the exercise. These findings were not replicated in the data for pathologists

assessing the annotated area of the slide only, in which different pathologists showed marginally higher kappa scores and ICC values (table 47).

Table 46. Kappa and intraclass correlation coefficient scores with 95% confidence intervals for each assessor versus digital tumour content determination on whole slide images

Assessor	Non-weighted kappa	Intraclass correlation coefficient with 95% confidence intervals					
		Single measures			Average measures		
		ICC	95% CI lower	95% CI upper	ICC	95% CI lower	95% CI upper
1	0.127	0.404	-0.360	0.843	0.575	-1.123	0.915
2	0.111	0.504	-0.245	0.876	0.670	-0.648	0.934
3	0.286	0.722	0.106	0.937	0.838	0.192	0.968
4	-0.136	0.660	-0.326	0.958	0.795	-0.967	0.979
5	0.034	0.692	0.048	0.930	0.818	0.092	0.964
6	0.127	0.580	-0.140	0.899	0.734	-0.327	0.947
7	0.051	0.566	-0.162	0.895	0.723	-0.386	0.944
8	0.418	0.925	0.672	0.984	0.961	0.804	0.992
9	-0.067	0.706	0.075	0.933	0.828	0.139	0.965
10	0.077	0.662	-0.009	0.922	0.796	-0.017	0.959

Table 47. Kappa and intraclass correlation coefficient scores with 95% confidence intervals for each assessor versus digital tumour content determination on the annotated area of each slide only

Assessor	Non-weighted kappa	Intraclass coefficient values with 95% confidence intervals					
		Single measures			Average measures		
		ICC	95% CI lower	95% CI upper	ICC	95% CI lower	95% CI upper
1	0.143	0.956	0.724	0.994	0.978	0.840	0.997
2	-0.0296	0.833	0.211	0.975	0.909	0.348	0.987
3	-0.0296	0.898	0.447	0.985	0.946	0.617	0.993
4	-0.0425	0.783	-0.078	0.975	0.878	-0.170	0.987
5	0.0005	0.862	0.167	0.985	0.926	0.286	0.992
6	0.1305	0.956	0.647	0.995	0.978	0.786	0.998
7	-0.0296	0.914	0.514	0.987	0.955	0.679	0.994
8	0.0006	0.207	-0.649	0.831	0.342	-3.699	0.908
9	0.0006	0.924	0.562	0.989	0.961	0.719	0.995

The standard deviation and upper and lower limits of the range for 95% limit of agreement have been calculated for pathologist values compared to the 'gold standard' digital tumour content determination (TCD) value and show a tendency of pathologists to over-estimate tumour content, as indicated by the standard deviation lying towards the upper limit in most cases (table 48).

Table 48. Standard deviation and upper and lower limits of the range for the 95% limit of agreement, calculated for pathologist values compared to the 'gold standard' digital TCD value

Case	Whole slide			Annotated area only		
	Standard deviation	Lower limit of 95% LA	Upper limit of 95% LA	Standard deviation	Lower limit of 95% LA	Upper limit of 95% LA
1	19	-47	27	Not applicable		
2	11	-14	29	13	-14	34
3	28	-29	80	16	-35	27
4	12	-0.6	45	12	-7	41
6	8	-11	19	12	-17	30
7	14	-23	33	10	-2	37
8	17	-21	45	Not applicable		
9	21	-31	52	17	-28	38

LA = limit of agreement

Implementation of these findings in practice requires understanding how accurate tumour content determination really needs to be, based on how the result will be used and affect clinical decision-making. One consequence could be deciding to test or not test a patient's sample, depending on whether it is above or below a tumour percentage content threshold set according to the limit of detection of the technology that will be used for mutation analysis. In SMP2, precise tumour DNA content measurement helps with interpretation of sequencing data to determine whether adequate read depth/ coverage of a gene has been achieved to confirm 'wild-type' status. This is important since it is a key molecular eligibility criterion of genes required for different arms of the National Lung Matrix Trial, for example wild-type *RB1* (retinoblastoma) tumour suppressor gene status and a functional retinoblastoma protein is required for the trial arm investigating the cyclin-dependent kinase 4/6 inhibitor palbociclib.

5.1.1 Summary

The findings of this work and specifically the ability of cellular pathologists to give an accurate value for tumour content in a sample being considered for molecular analysis should be interpreted in the context of a lack of clear consensus or guidelines about how this should be determined. As described above, not only is there a lack of formal guidance but there are also several different indications for performing this type of assessment, requiring differing levels of accuracy and with different implications for patient treatment. UK NEQAS, working in conjunction with NHS England, Genomics England and the Belfast-based digital pathology company PathXL (Belfast, Northern Ireland), have recently established an online pilot external quality assessment scheme for tumour content assessment. This is primarily being run through genomic medicine centres and the results of this will hopefully be used to contribute to standard setting in this area. As a result of the tumour content assessment work carried out through SMP2, the following text has been added to the study sample and patient eligibility criteria document (version 9, implemented on 1st March 2016) to try and provide some guidance for pathologists and help to standardise the approach:

"Tumour content should be assessed in an H&E stained section as follows:

- Percentage of tumour cell nuclei present expressed as a proportion of all cell nuclei present (including admixed inflammatory and stromal cells) to the nearest 5% or 10%
- It should be noted that the assessment should be based on nuclear size/volume as a surrogate marker of DNA content rather than surface area of tumour on the slide, such that a small cluster of lymphocytes will yield more DNA than an equivalent sized nest of tumour cells, in which each cell will be larger.
- This model does not take into account the 3-dimensional nature of the tissue in the block (and how this is represented in serial sections) or tumour cell hyperdiploidy/aneuploidy
- Within the marked area only if macrodissection is to be performed and an area has been marked for macrodissection
- Viable tumour only, excluding necrotic areas or apoptotic cells"

5.2 Results of Experimental Work on PAXgene® Tissue Fixation System

5.2.1 Sample details

36 paired tumour or normal tissue samples were obtained by a pathologist from 18 fresh surgical resection specimens as described in the methods in chapter 2. Characteristics of the samples are shown in table 49. The same two pathologists assessed the tissue sections for morphology, histochemistry and immunohistochemistry as described in the following sections.

Table 49. Sample characteristics

Case number	Tissue type	Histological subtype
1	Lung	Squamous cell carcinoma
2	Lung	Squamous cell carcinoma
3	Lung	Squamous cell carcinoma
4	Pleura	Reactive/ inflammatory
5	Lymph node	Reactive
6	Lung	Adenocarcinoma
7	Lung	Typical carcinoid
8	Lung	Adenocarcinoma
9	Spleen	Normal tissue
10	Lung	Metastatic melanoma
11	Lung	Adenocarcinoma
12	Lung	Adenocarcinoma
13	Thymus	Thymoma
14	Lung	Squamous cell carcinoma
15	Lymph node	Hodgkin lymphoma
16	Lymph node	Metastatic melanoma
17	Pancreas	Normal tissue
18	Lymph node	Reactive

5.2.2 Morphology assessment

The results of blinded independent assessment of cellular and tissue preservation and suitability for diagnosis in H&E-stained sections by two pathologists are shown in table 50, with nuclear, cytoplasmic and other tissue components assessed to give a maximum possible score of 12 for each case. Representative images of lung and lymph node tissue are shown in figures 23 and 24. The pathologists showed 94% concordance (17/18 cases) in predicting which was the PAXgene® Tissue-fixed and paraffin-embedded (PFPE) tissue sample of each pair. The pathologists reported that these could be easily identified due to a generalised increased intensity of eosin staining in the section and also swelling and central clearing of erythrocytes, both recognised artefacts in other studies using PAXgene® Tissue system[133-135]. There were signs of increased tissue fragility in PFPE tissue compared to FFPE tissue, particularly in necrotic areas where tearing of sections was more commonly seen. Average scores for PFPE (10.7 for pathologist 1 and 9.6 for pathologist 2) were slightly lower than those for FFPE tissue (11.8 for both pathologists). Lymph nodes showed noticeably inferior preservation in PAXgene® Tissue, with cell shrinkage and tissue disaggregation as well as slightly less crisp nuclear features. Overall both pathologists assessed all FFPE (100%) and 89% of PFPE sections (16/18) as suitable for diagnosis. A Bland-Altman plot of difference versus average for morphology scores for the series of FFPE and PFPE tissues is displayed in figure 24. This shows reasonable levels of equivalence of the two different types of material for morphological analysis, with the trend towards positive values indicating higher scores and a general preference for FFPE material by the two pathologists. The main exception to this is case 15, a lymph node with Hodgkin lymphoma, where the pathologists differed in their view on whether FFPE or PFPE gave superior morphology (table 50, figure 25). This was the only case in which taking the mean average of the two pathologist's scores gave a higher score for PFPE than FFPE tissue.

Table 50. Pathologist assessment of morphology and suitability of FFPE and PFPE-derived sections for diagnosis

Case number	Formalin fixed and paraffin embedded tissue				PAXgene® Tissue fixed and paraffin embedded tissue			
	Pathologist 1		Pathologist 2		Pathologist 1		Pathologist 2	
	Score	Suitable for diagnosis	Score	Suitable for diagnosis	Score	Suitable for diagnosis	Score	Suitable for diagnosis
Lung								
1	12	Yes	12	Yes	12	Yes	11	Yes
2	12	Yes	12	Yes	10	Yes	10	Yes
3	12	Yes	11	Yes	12	Yes	10	Yes
6	12	Yes	12	Yes	10	Yes	9	Yes
7	9	Yes	12	Yes	9	Yes	10	Yes
8	12	Yes	12	Yes	12	Yes	9	Yes
10	12	Yes	12	Yes	8	Borderline	8	Yes
11	12	Yes	12	Yes	11	Yes	10	Yes
12	12	Yes	12	Yes	11	Yes	12	Yes
14	12	Yes	12	Yes	10	Yes	9	Yes
Lymphoid tissue								
5	12	Yes	12	Yes	10	Borderline	9	Yes, just acceptable
15	11	Yes	9	Yes	10	Yes	12	Yes
16	12	Yes	12	Yes	10	Yes	9	Yes
18	12	Yes	12	Yes	10	Yes	6	No
Other tissues								
4	12	Yes	12	Yes	12	Yes	12	Yes
9	12	Yes	12	Yes	12	Yes	9	No
13	12	Yes	12	Yes	12	Yes	9	Yes
17	12	Yes	12	Yes	12	Yes	9	Yes

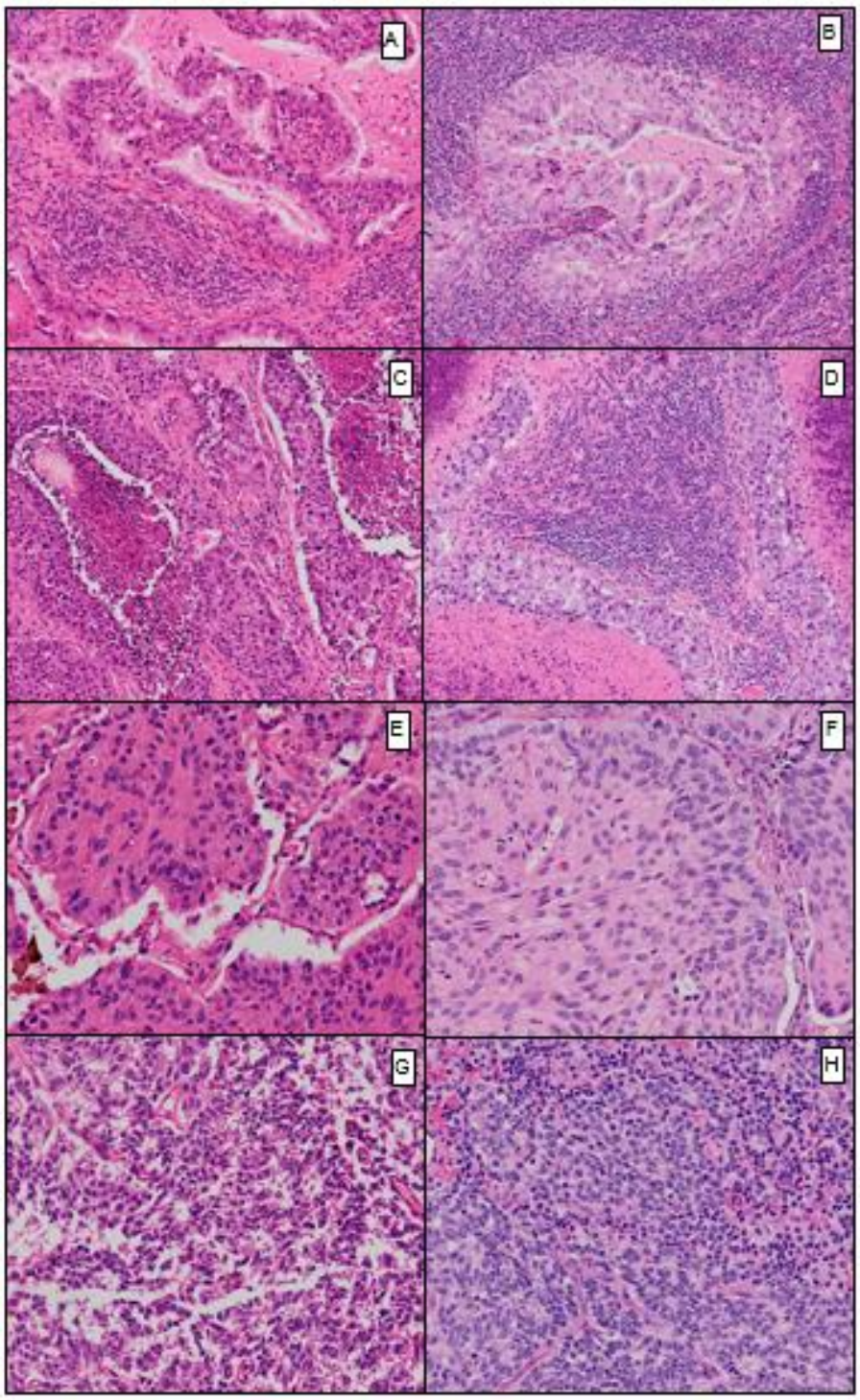

Figure 23. Representative images of haematoxylin and eosin-stained sections of lung tumour for formalin (FFPE)- and PAXgene® Tissue (PFPE) - fixed samples

A-D show adenocarcinoma (cases 11 and 6, 100x overall magnification); E-F show squamous cell carcinoma (case 3, 100x overall magnification in E and 200x overall magnification in F) and G-H show typical carcinoid (case 7, 200x overall magnification). Images A, C, E and G are PFPE sections and images B, D, F and H are FFPE sections.

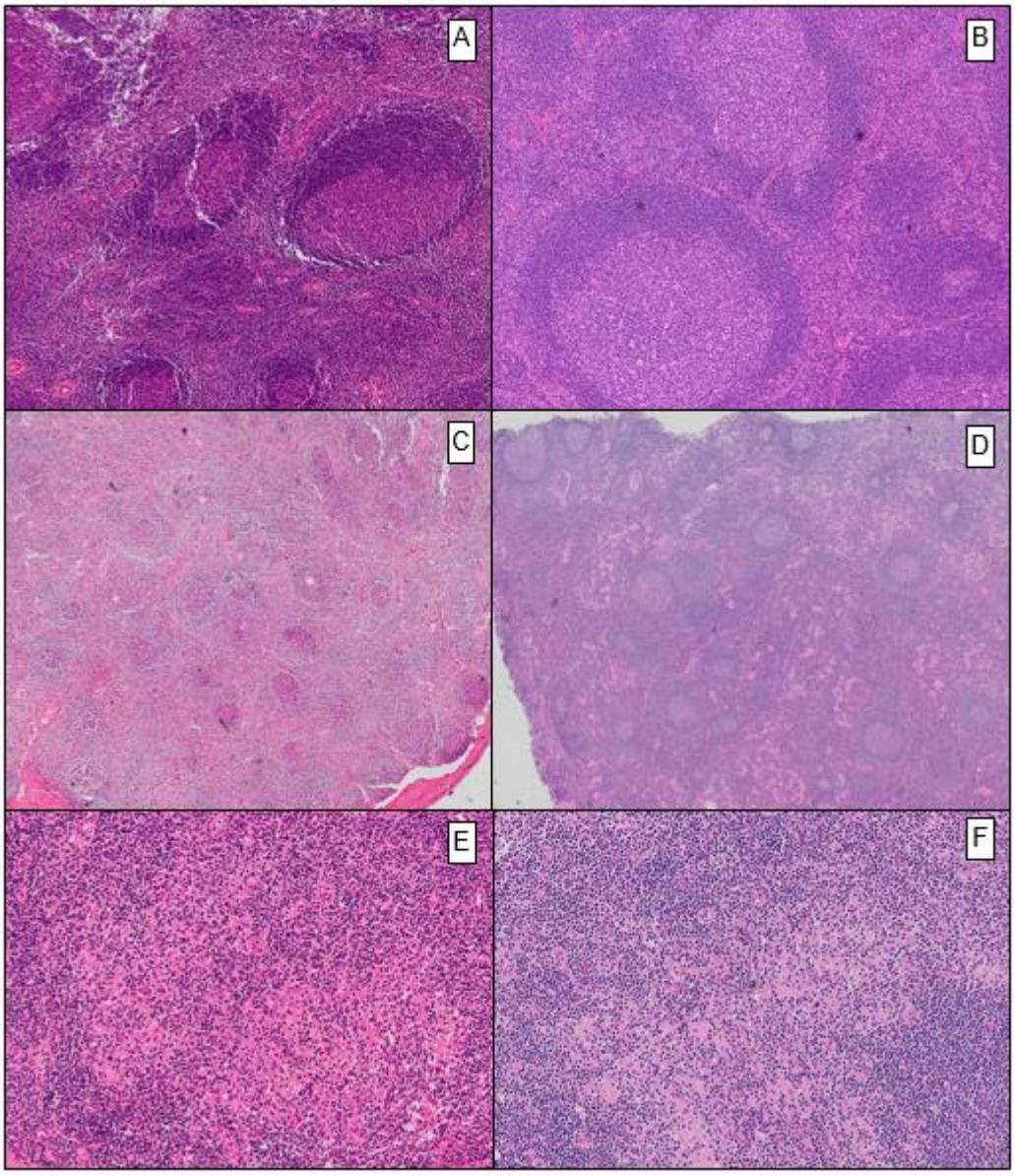

Figure 24. Representative images of haematoxylin and eosin-stained sections of lymph node for formalin-and PAXgene® Tissue- fixed samples

A-D show reactive lymph node (cases 5 and 18, 40x overall magnification in A and B and 20x overall magnification in C and D) and E-F show a Hodgkin lymphoma of mixed cellularity subtype (case 15, 100x overall magnification). Images A, C and E are PFPE sections and images B, D and F are FFPE sections.

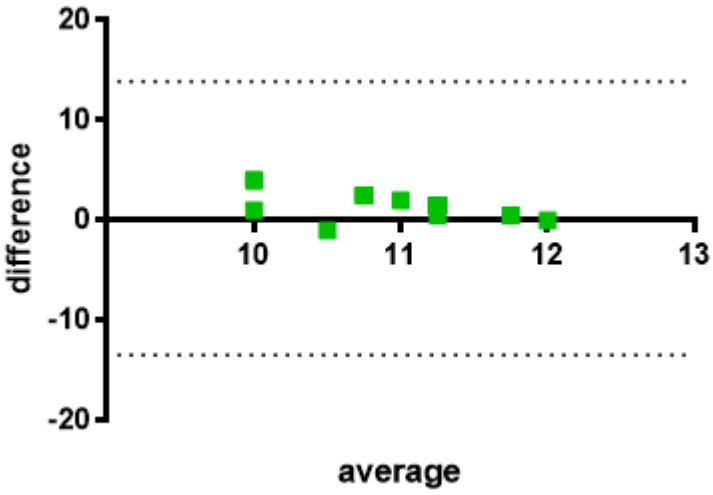

Figure 25. Bland-Altman plot of mean average morphology scores for FFPE against PFPE tissue

The dotted lines indicate 95% limits of agreement.

These results suggest that PAXgene® Tissue-fixed samples can be recognised easily in H&E-stained sections by pathologists and overall show comparable morphology to formalin-fixed samples, which is generally suitable for diagnostic purposes. In this series of cases, lymphoid tissue showed slightly inferior preservation which may affect the diagnostic process. This slight inferiority of morphology may be tolerable to the pathologist in certain cases, but is likely to be problematic in situations where the nuclear chromatin pattern is particularly important, such as in differentiating reactive conditions from malignant conditions, pyknotic nuclei from mitotic figures (for example in uterine smooth muscle tumours treated with hormone modulation therapy) and also in specific situations such as detecting plasmacytic differentiation, Hodgkin Reed-Sternberg cells, centrocytes and centroblasts in lymphoreticular pathology; identifying and grading neuroendocrine tumours and grading of many tumour types, including breast and renal carcinoma.

5.2.3 Histochemistry assessment

The results of blinded independent assessment by two pathologists of tissue sections stained for several different histochemical stains routinely used for diagnostic purposes are shown in table 51. These show reasonable levels of agreement in scoring between the two pathologists with no definite inferiority

of the PAXgene®-fixed tissues, indicating that both the usual manual and automated staining protocols in use in the laboratory are applicable to tissue prepared using the PAXgene® Tissue system, and based on this sample set do not require modification. The one exception to this is that both pathologists found PAS staining to be of lower quality in PAXgene®-fixed pancreatic tissue than formalin-fixed pancreatic tissue. These findings are not conclusive on the basis of a single sample and require further work. It may be relevant that this was pancreatic tissue, which is a tissue type known to undergo rapid autolysis due to its inherent high content of proteolytic digestive enzymes, though the tissue did not show morphological evidence of autolysis. Apart from the pleural sample stained with EVG, which was assessed by both pathologists as sub-optimal tissue for staining, all the histochemical-stained slides were assessed as a pass or better with appropriate demonstration of the tissue components of interest.

Table 51. Assessment of histochemical staining quality by two pathologists

Case number	Formalin fixed and paraffin embedded tissue		PAXgene® Tissue fixed and paraffin embedded tissue	
	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2
Lung – EVG				
5	4	5	4	4
12	4	4	4	5
14	3	4	4	5
22	4	5	5	3
Pleura - EVG*				
16	3	2	3	2
Spleen – Reticulin				
9	4	3	3	4
Pancreas – PAS				
17	4	5	3	3
Pancreas – DPAS				
17	4	4	4	3

*Both pathologists commented that this was sub-optimal material for EVG staining due to the amount of necrosis and loss of tissue architecture, making it likely that there was no elastin in the section to stain and accounting for low scores on both types of preparation.

5.2.4 Immunohistochemistry assessment

The results of blinded independent assessment by the two study pathologists of tissue sections incubated with different immunohistochemical antibodies varying by tissue and tumour type are shown in tables 52-57 with representative images displayed in figures 26 and 27. There was a high level of concordance in scoring between the two pathologists. Although PAXgene® Tissue is a non-cross linking fixative and the hypothesis was that the usual heat-induced epitope retrieval step would not be required, the scores for nuclear antibodies are unacceptably low for PFPE without heat-induced epitope retrieval. This indicates that the antigen retrieval step is required at least for antibodies targeting nuclear proteins, and may be necessary to ensure nuclear permeability during the immunohistochemical reactions. Recent work by the SPIDIA consortium has led to the identification of a similar issue with FISH on PFPE tissue, where the probes need to bind nuclear DNA, but the investigators found that this could be overcome by a period of post-fixation of the tissue section in formalin[136].

The scores for nuclear proliferation marker Ki67 on PFPE were inferior to those for FFPE, even with the use of heat-induced epitope retrieval. Having discussed this with scientists at Qiagen and based on their previous experience, they advise that performance of the Ki67 antibody is dependent on the pH of the buffer used for epitope retrieval and works best with the high pH (pH9.0) target retrieval solution. This was the only antibody in our series in which the departmental protocol recommends use of low pH (6.1) target retrieval solution and repeat is in progress using the high pH solution to see if this gives better results for PFPE tissue sections.

In contrast, the expression of cytoplasmic and membranous antibodies was better and in most cases acceptable to the pathologists. MNF116 in PFPE lung cancer samples with no heat-induced epitope retrieval was even preferred to the corresponding FFPE sections by pathologists in the eight cases in which it was used.

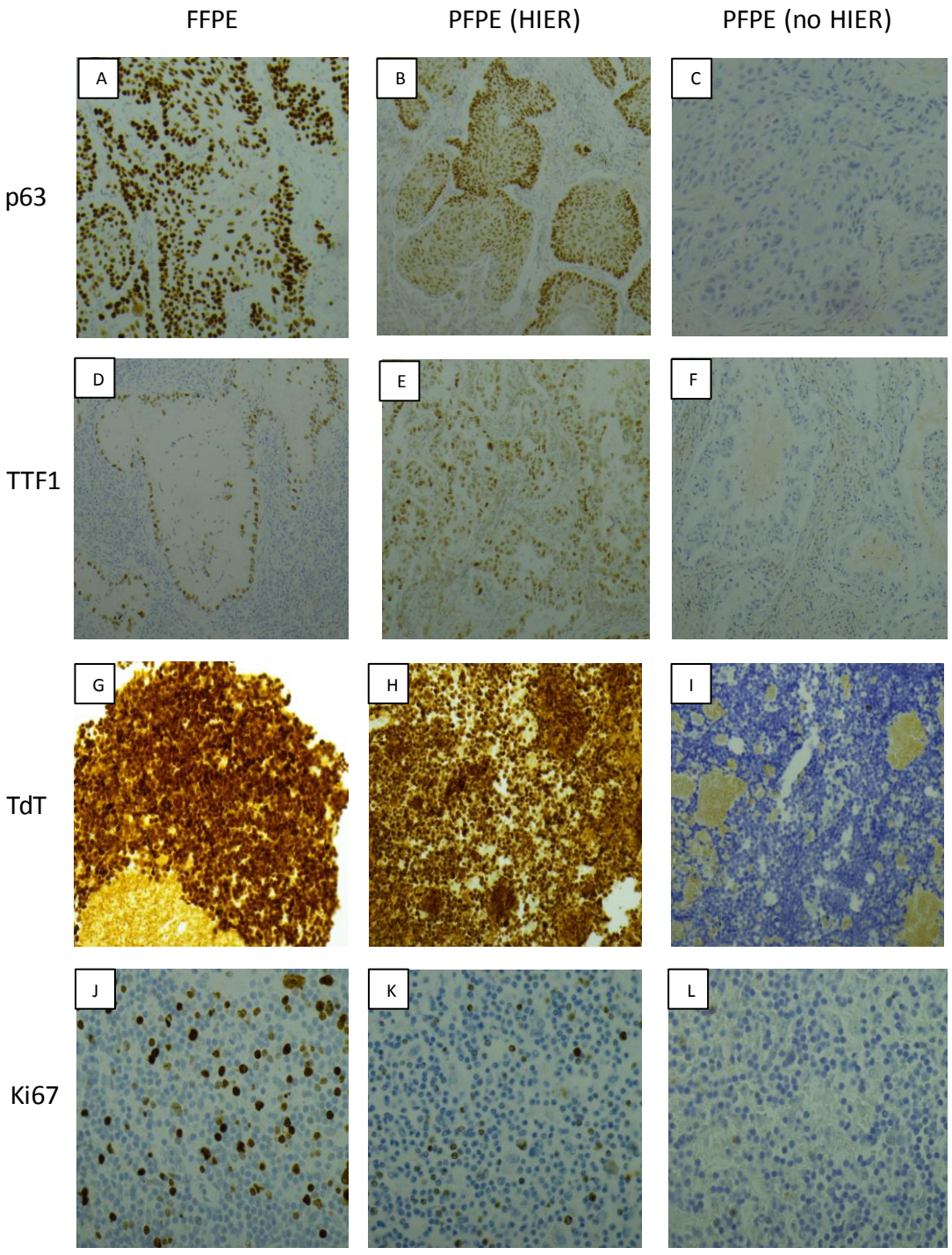

Figure 26. Comparison of immunohistochemistry for antibodies with nuclear expression

A-C, squamous cell carcinoma of lung, case 3, 100x magnification; D-F, adenocarcinoma of lung, case 11, 100x magnification; G-I, thymoma, case 13, 200x magnification; J-L, Hodgkin lymphoma, case 15, 400x magnification.

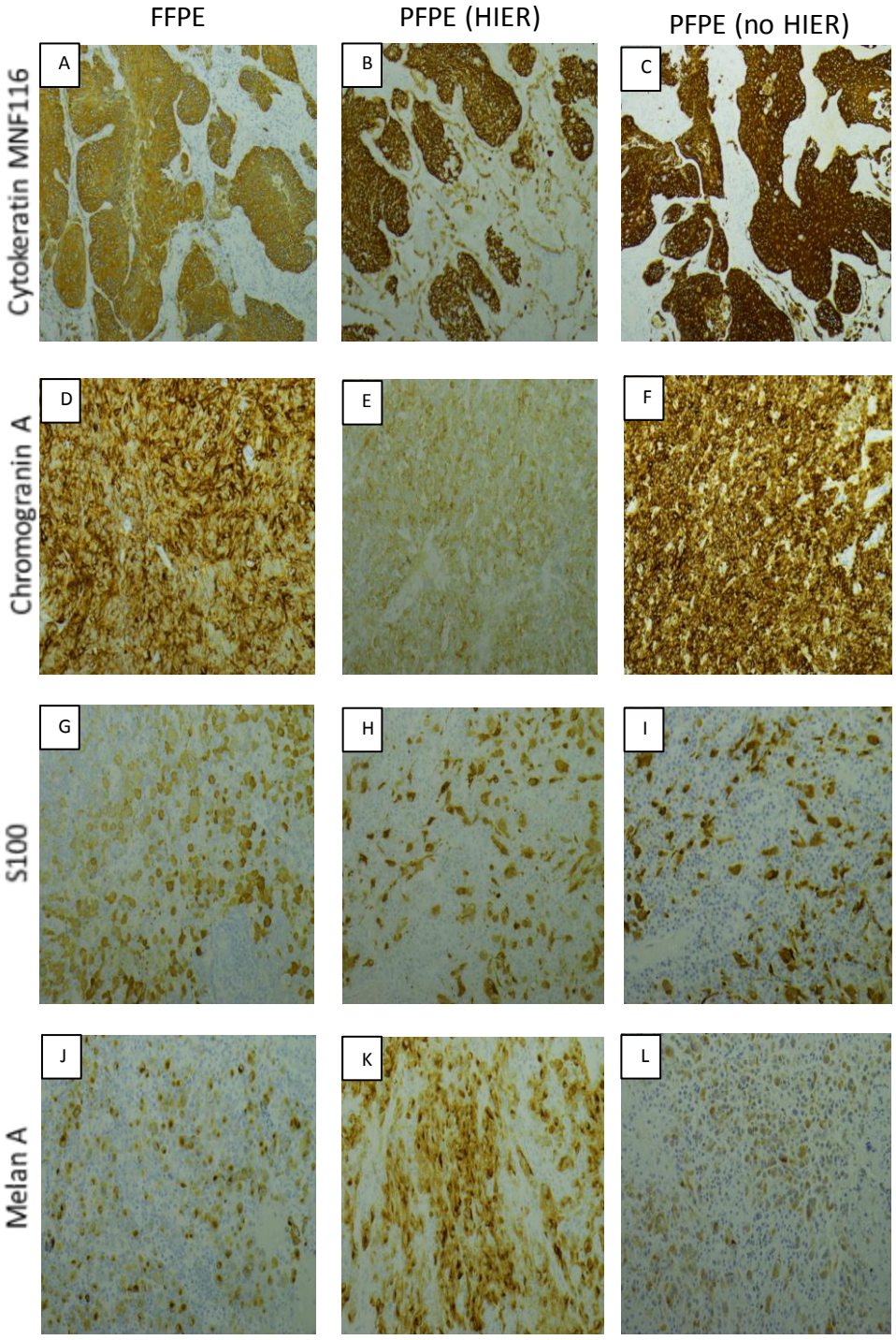

Figure 27. Comparison of immunohistochemistry for antibodies with cytoplasmic and/or membrane expression

A-C, squamous cell carcinoma of lung, case 3, 100x magnification; D-F, typical carcinoid of lung, case 7, 200x magnification; G-L, melanoma, case 16, 200x magnification.

Overall the results indicate that usual laboratory immunohistochemistry protocols for these antibodies cannot be successfully applied to tissues prepared using the PAXgene® Tissue system without modification, particularly for antibodies to proteins located in the cell nucleus. There is evidence of a difference in permeability of the formalin- and PAXgene® Tissue- fixed samples. Work is ongoing to modify and optimise the protocols for PFPE tissue, including trials of post-fixation in formalin and immersion in detergent NP-40 to increase tissue permeability to nuclear antibodies.

Table 52. Scores for pathologist assessment of immunohistochemistry with selected antibodies in lung cancer sections

Scoring system described in table 19 (page 71). ADC = adenocarcinoma; FFPE = formalin-fixed paraffin embedded; HIER = heat-induced epitope retrieval; PFPE = PAXgene® Tissue-fixed paraffin embedded; SCC = squamous cell carcinoma.

Case	Histological subtype	p63						TTF1						MNF116						Comments
		FFPE		PFPE				FFPE		PFPE				FFPE		PFPE				
				HIER		No HIER				HIER		No HIER				HIER		No HIER		
		Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	
1	SCC	4	5	3	1	2	1	4	5	3	3	2	1	3	3	3	3	4	5	
2	SCC	4	5	3	3	1	1	1	1	4	3	1	1	3	3	3	4	4	5	TTF1 genuinely negative
3	SCC	4	4	3	3	1	1	4	4	3	2	3	1	4	4	5	5	5	5	
6	ADC	5	4	3	3	2	1	4	4	3	3	2	1	3	3	4	4	4	5	
8	ADC	1	1	1	1	1	1	4	3	3	3	2	1	4	4	3	3	5	5	p63 genuinely negative
11	ADC	4	2	3	1	2	1	4	4	2	1	3	3	4	4	4	4	4	5	
12	ADC	1	1	1	1	1	1	4	4	3	3	1	1	3	4	3	5	4	5	p63 genuinely negative
14	SCC	5	5	3	3	2	1	5	5	3	3	2	1	4	5	5	4	4	5	

Table 53. Scores for pathologist assessment of immunohistochemistry with selected antibodies in tissue sections containing metastatic melanoma

Scoring system described in table 19 (page 71). FFPE = formalin-fixed paraffin embedded; HIER = heat-induced epitope retrieval; PFPE = PAXgene® Tissue-fixed paraffin embedded.

Case	Tissue type	Melan A						HMB45						S100					
		FFPE		PFPE				FFPE		PFPE				FFPE		PFPE			
				HIER		No HIER				HIER		No HIER				HIER		No HIER	
		Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2
10	Lung	4	5	4	5	3	3	3	4	2	2	3	3	3	4	4	4	3	4
16	Lymph node	4	4	3	4	2	3	3	4	3	3	4	5	3	3	4	5	4	5

Table 54. Scores for pathologist assessment of immunohistochemistry with selected antibodies in reactive lymphoid tissue

Scoring system described in table 19 (page 71). FFPE = formalin-fixed paraffin embedded; HIER = heat-induced epitope retrieval; PFPE = PAXgene® Tissue-fixed paraffin embedded.

Case	CD20						CD3						CD10						Ki67					
	FFPE		PFPE				FFPE		PFPE				FFPE		PFPE				FFPE		PFPE			
			HIER		No HIER				HIER		No HIER				HIER		No HIER				HIER		No HIER	
	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2
5	5	5	3	5	3	5	5	5	3	3	4	4	4	4	3	5	2	3	5	5	3	3	2	2
18	3	3	3	3	4	4	4	5	4	4	2	1	3	5	3	4	2	3	5	5	5	4	2	2

Table 55. Scores for pathologist assessment of immunohistochemistry with selected antibodies in a lymph node showing involvement by Hodgkin lymphoma

Scoring system described in table 19 (page 71). FFPE = formalin-fixed paraffin embedded; HIER = heat-induced epitope retrieval; PFPE = PAXgene® Tissue-fixed paraffin embedded.

Case	CD30						CD15						Ki67					
	FFPE		PFPE				FFPE		PFPE				FFPE		PFPE			
			HIER		No HIER				HIER		No HIER				HIER		No HIER	
	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2
15	4	5	3	4	3	5	3	5	4	4	4	3	5	4	2	1	3	2

Table 56. Scores for pathologist assessment of immunohistochemistry with selected antibodies in a pulmonary typical carcinoid tumour

Scoring system described in table 19 (page 71). FFPE = formalin-fixed paraffin embedded; HIER = heat-induced epitope retrieval; PFPE = PAXgene® Tissue-fixed paraffin embedded.

Case	CD56						Chromogranin						Synaptophysin						Ki67					
	FFPE		PFPE				FFPE		PFPE				FFPE		PFPE				FFPE		PFPE			
			HIER		No HIER				HIER		No HIER				HIER		No HIER				HIER		No HIER	
	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2
7	5	5	4	5	5	5	4	5	2	1	4	5	5	5	5	5	5	5	4	4	4	4	1	1

Table 57. Scores for pathologist assessment of immunohistochemistry with selected antibodies in a thymoma

Scoring system described in table 19 (page 71). FFPE = formalin-fixed paraffin embedded; HIER = heat-induced epitope retrieval; PFPE = PAXgene® Tissue-fixed paraffin embedded.

Case	MNF116						CD3						CD20						TdT					
	FFPE		PFPE				FFPE		PFPE				FFPE		PFPE				FFPE		PFPE			
			HIER		No HIER				HIER		No HIER				HIER		No HIER				HIER		No HIER	
	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2	Pathologist 1	Pathologist 2
13	3	4	5	4	4	4	5	4	4	3	3	4	4	5	4	3	4	4	5	5	5	5	1	1

5.2.5 Assessment of extracted DNA quantity and quality

Figures 28-32 display the NanoDrop quantification and purity measures and Qubit dsDNA quantification from the matched tissue pairs. For sample purity assessed using NanoDrop, the optimal range for the 260/280 ratio lies between 1.8 (pure DNA) and 2.0 (pure RNA), with values outside this range indicative of co-existing contaminants. The FFPE values tended to be higher than the corresponding FFPE ratios, with the PFPE ratios generally closer to 1.8 and within the 1.8-2.0 range implying higher purity of DNA in these samples. Only one sample in this series has a 260/280 value less than 1.8. This is the PFPE derived DNA from sample 17, pancreatic tissue, which also showed an unusually high 260/230 ratio compared to the rest of the samples and one of the lowest yields of DNA on both NanoDrop and Qubit quantification. As well as the presence of contaminants such as protein or reagents used in nucleic acid extraction, low 260/280 ratios may be due to low DNA concentrations. Pancreatic tissue is rich in proteolytic digestive enzymes and the low DNA concentration may represent DNA destruction in the tissue as a result of autolysis prior to fixation, though this was not a prominent morphological feature in the H&E-stained sections. This brings into question how reliable morphological assessment of autolysis is and it is possible that DNA quality is a more sensitive marker of tissue degradation, with the classic morphological features appearing late in the process and after DNA damage has already occurred.

260/230 ratios are typically higher than 260/280 and usually between 2.0-2.2, though samples with 260/230 ratios greater than 1.8 are generally considered suitable for analysis. A low 260/230 value may be due to salt or solvent contamination of the sample and may be resolved through re-purification of the sample. In this series, two FFPE lung samples (2 and 8) had 260/230 values less than 1.8 and one PFPE sample (case 9, spleen) but the corresponding samples prepared in the other fixative had values within the normal range indicating that this is a fixative rather than tissue-related finding.

Figure 32 illustrates the higher concentrations of DNA measured with NanoDrop than the dsDNA-specific Qubit technique. The literature suggests that yields on NanoDrop are typically 3-4 times higher than on Qubit[95-97], and

this is borne out in the PFPE data where the mean average concentration measured using NanoDrop is 2.93 (range 1.73 - 4.79) times the Qubit value. The situation is different for the FFPE samples, in which the difference between the two measurements shows much more variation, with the mean average concentration measured using NanoDrop being 11 (range 2.49 – 53) times higher than the Qubit value. This may be due to the NanoDrop quantification including contaminants such as fragmented single-stranded DNA due to formalin damage in these samples.

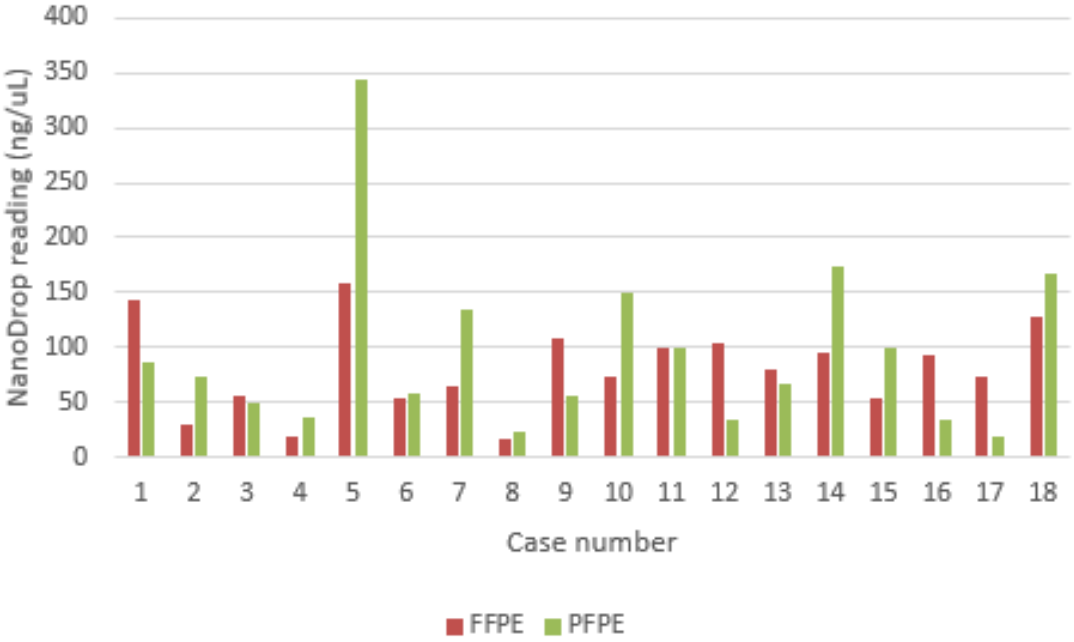

Figure 28. NanoDrop quantification for DNA extracted from matched tissue samples prepared in PAXgene® Tissue and formalin fixative

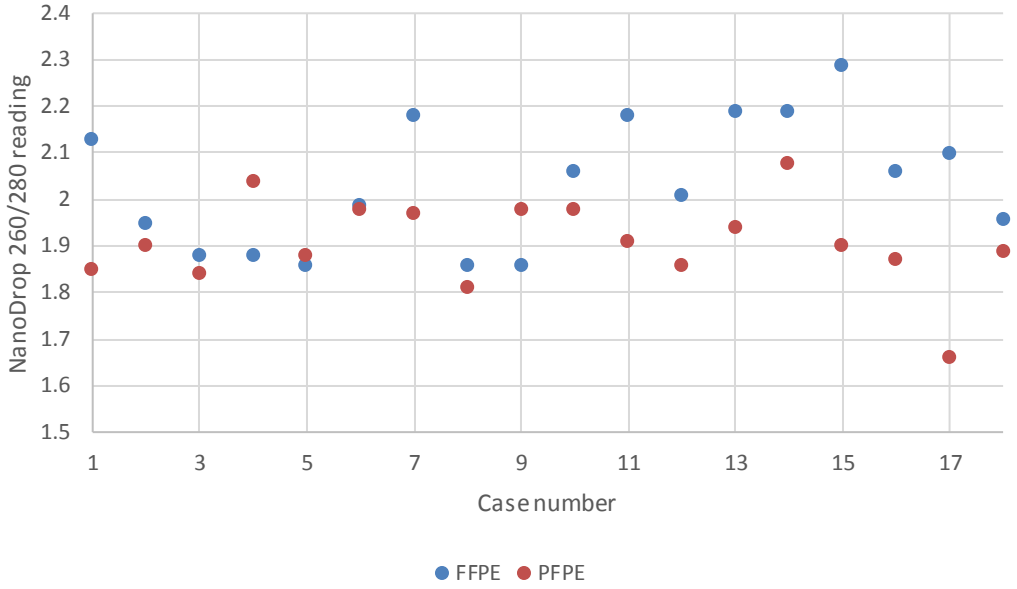

Figure 29. NanoDrop 260/280 purity assessment for DNA extracted from matched tissue samples prepared in PAXgene® Tissue and formalin fixative

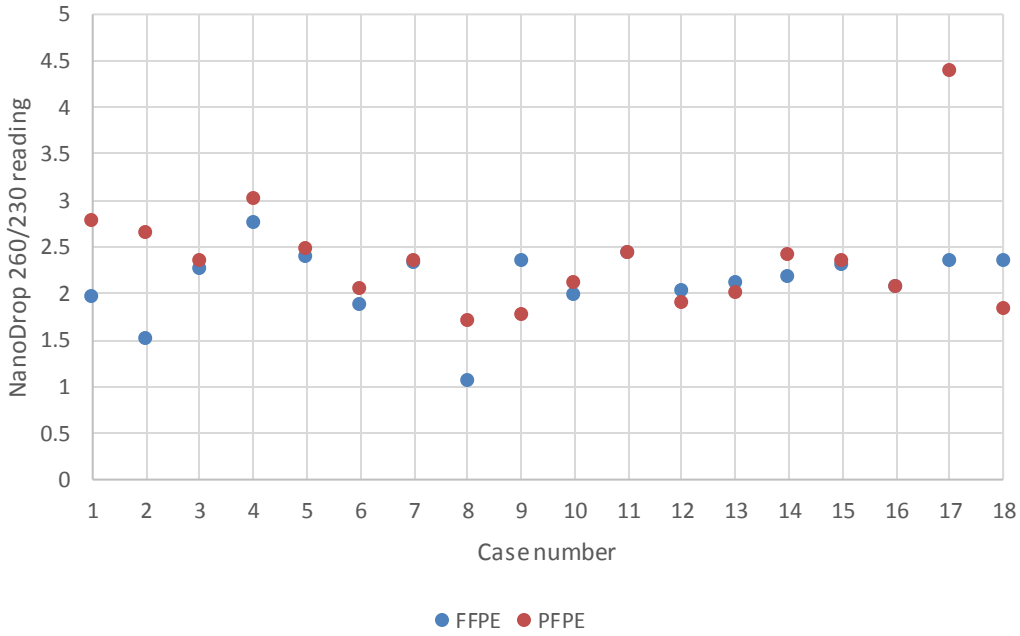

Figure 30. NanoDrop 260/230 purity assessment for DNA extracted from matched tissue samples prepared in PAXgene® Tissue and formalin fixative

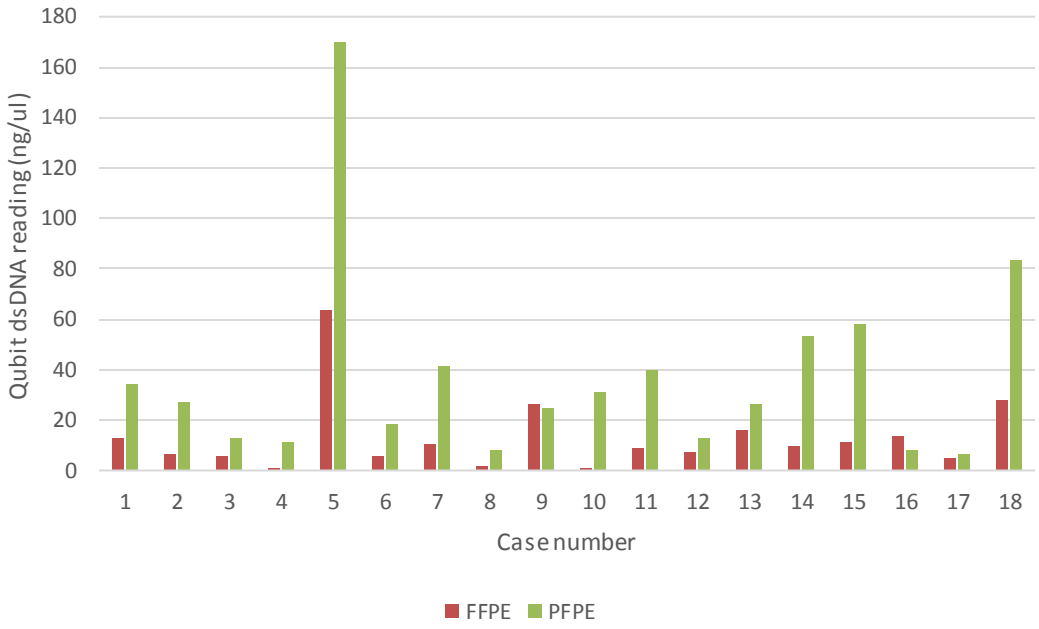

Figure 31. Qubit dsDNA quantification for DNA extracted from matched tissue samples prepared using PAXgene® Tissue and formalin fixative

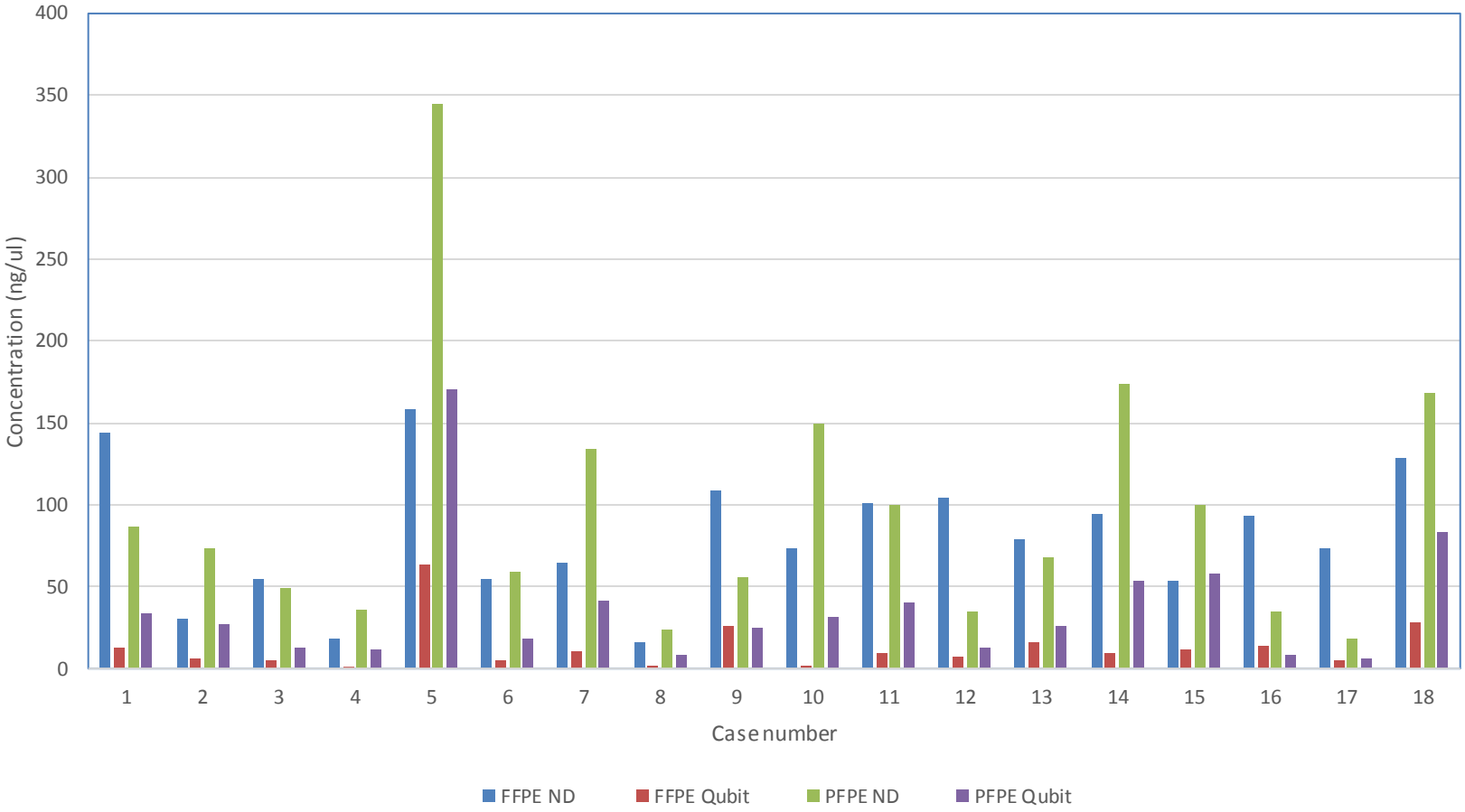

Figure 32. Comparison of DNA yields from FFPE and PFPE assessed using NanoDrop (ND) or Qubit dsDNA-specific assays

Whilst every effort was made to sample identical sized pieces of tissue for the FFPE and PFPE mirror blocks and DNA was extracted from the same number of sections, it was not possible to control for differences in cellularity between the two samples due to for example different stromal and inflammatory cell content or the presence of acellular tissue such as in necrotic areas. These factors are likely to have impacted on the DNA yield from the different tissue sections, limiting direct comparisons of DNA yield.

Results of fragmentation analysis using the BIOMED2 control PCR primers for different tissue and tumour types are shown in figures 33-37. All of the PFPE samples show a fluorescence signal of varying amplitude at both 600bp and 400bp. In contrast, none of the FFPE samples have anything more than a barely discernable 600bp peak and many also lack a 400bp peak, indicating a greater degree of fragmentation in DNA extracted from these samples compared to PFPE tissue. This does appear to be a statistically significant difference, with a p value of <0.0001 on a paired t-test with a standard deviation of 78, though the sample size in this set was only 18 and did not meet the 20 sample criteria on which the power calculation was based.

The BIOMED2 control PCR primer sets are designed to measure the relative abundance of DNA fragments of different lengths in a given sample. Although in the original protocol primers were designed generating DNA fragments up to 1000bp in length, for routine diagnostic practice use of primer sets up to 400bp in length are sufficient to establish the presence of DNA of suitable quality for immunoglobulin and T cell receptor clonality assessment. When interpreting the findings of BIOMED2 analysis it is important to be aware that various sequencing techniques require DNA fragments of different lengths, though dsDNA is essential starting material. Technical protocols tend to stipulate absolute input amounts of DNA rather than fragment lengths required.

5.2.6 Discussion

This work involving a series of samples representing a broad range of tissue types demonstrate that PAXgene® Tissue system provides similar morphological preservation and superior DNA preservation compared to

fixation in neutral buffered formalin. Whilst histochemistry methods for tinctorial stains appear to be transferrable without modification, this is not the case for immunohistochemistry and in particular nuclear antigens. The routine adoption of PAXgene® Tissue in a cellular pathology service would therefore require further work on protocol optimisation. From a practical perspective, routine use of the PAXgene® Tissue system would require a dedicated formalin-free tissue processor. This resource implication would have to be justified due to the requirement to validate, maintain and include in UKAS/ISO accreditation scope any laboratory equipment that is involved in delivery of the clinical diagnostic service. A further major workflow difference is the need to remember to switch the tissue from sample chamber 1 (fixative) to sample chamber 2 (stabiliser) of the dual chamber pot after between 3 and 24 hours. This was in part due to the need to batch infrequently acquired tissue samples for processing, and might be simplified by having more regular tissue processing runs. The stabilisation step cannot be completely omitted and Qiagen scientists advise a minimum of 2 hours per sample in the stabiliser solution following fixation and prior to processing. It may be possible to simplify the workflow by adding samples to stabiliser solution in the tissue processor following fixation, to be held at ambient temperature whilst waiting for a suitable sized batch to process. On the basis of this work there appears to be a role for PAXgene® Tissue system in preparing tissue samples for combined morphological and molecular analysis. Initially this is justifiable for preparation of a 'genomic block' from a fresh resection specimen, with the advantage over fresh frozen tissue of superior morphological preservation when compared to a frozen section, facilitating accurate assessment of tumour content and necrosis. The remainder of the specimen could then be fixed in formalin and dissected as normal to generate the diagnostic report with further molecular analysis guide by the content of the PAXgene® fixed tissue, final staging and treatment options under consideration for the patient.

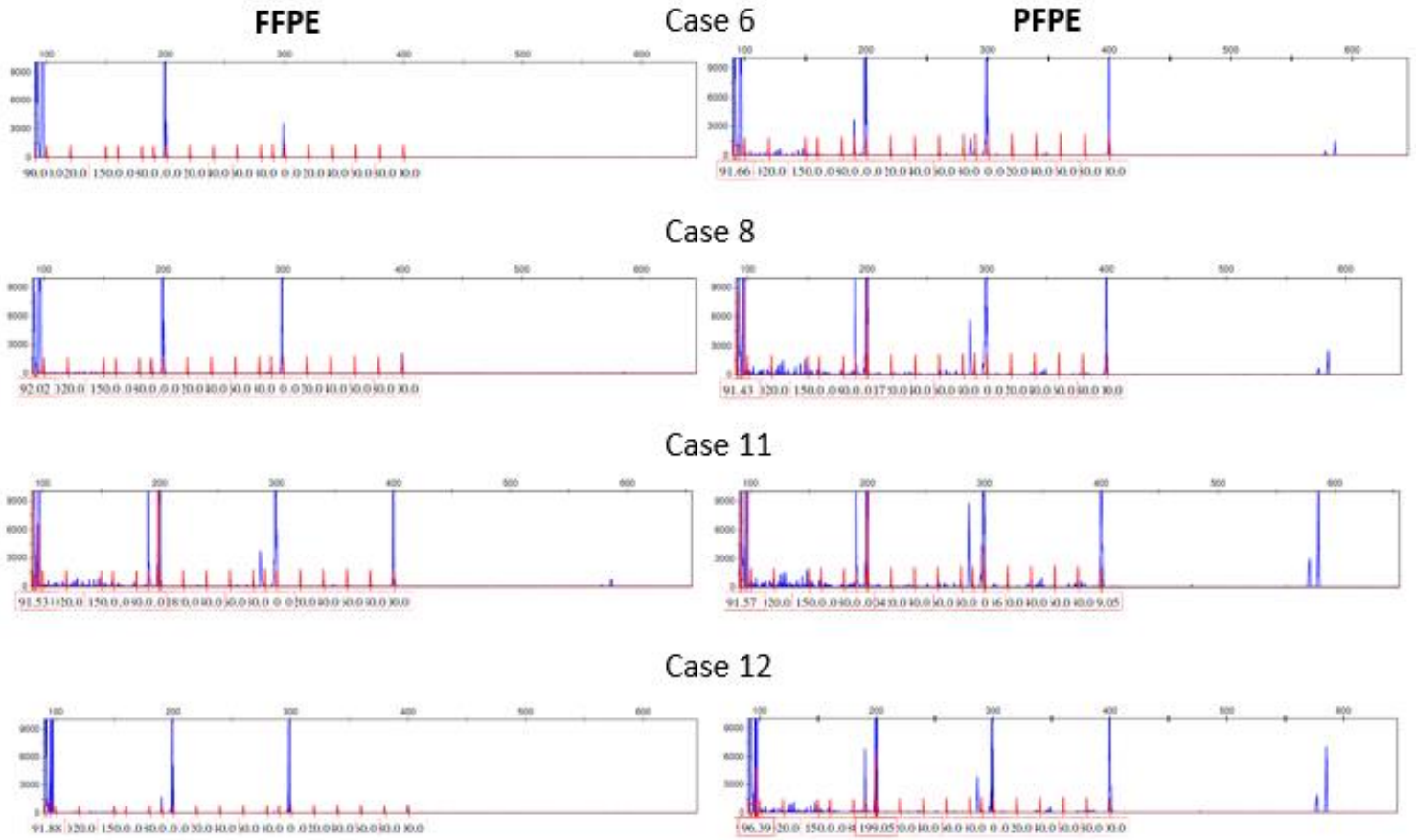

Figure 33. BIOMED2 PCR results for DNA extracted from matched formalin and PAXgene® Tissue fixed pairs of lung adenocarcinoma tissue

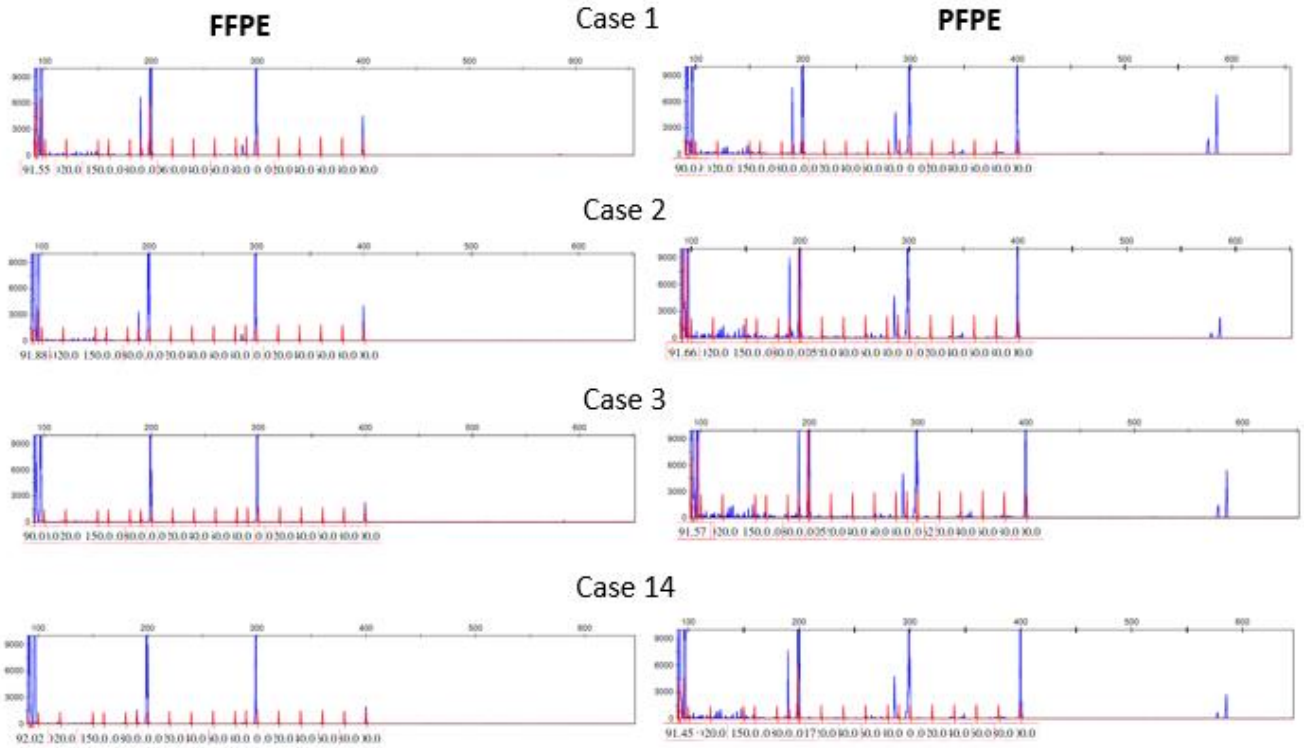

Figure 34. BIOMED2 PCR results for DNA extracted from matched formalin and PAXgene® Tissue fixed pairs of lung squamous cell carcinoma tissue

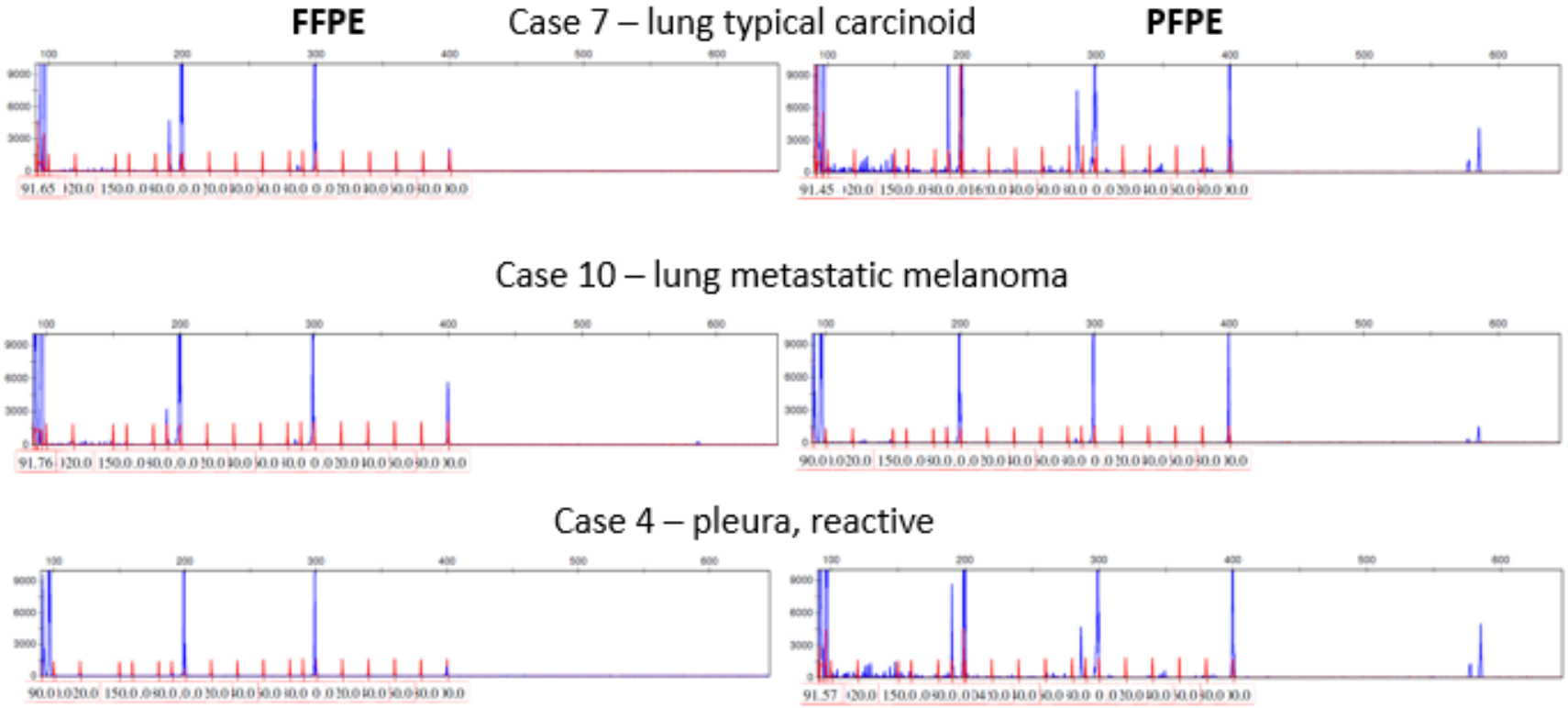

Figure 35. BIOMED2 PCR results for DNA extracted from matched formalin and PAXgene® Tissue fixed pairs of lung and pleural tissue of varying histology

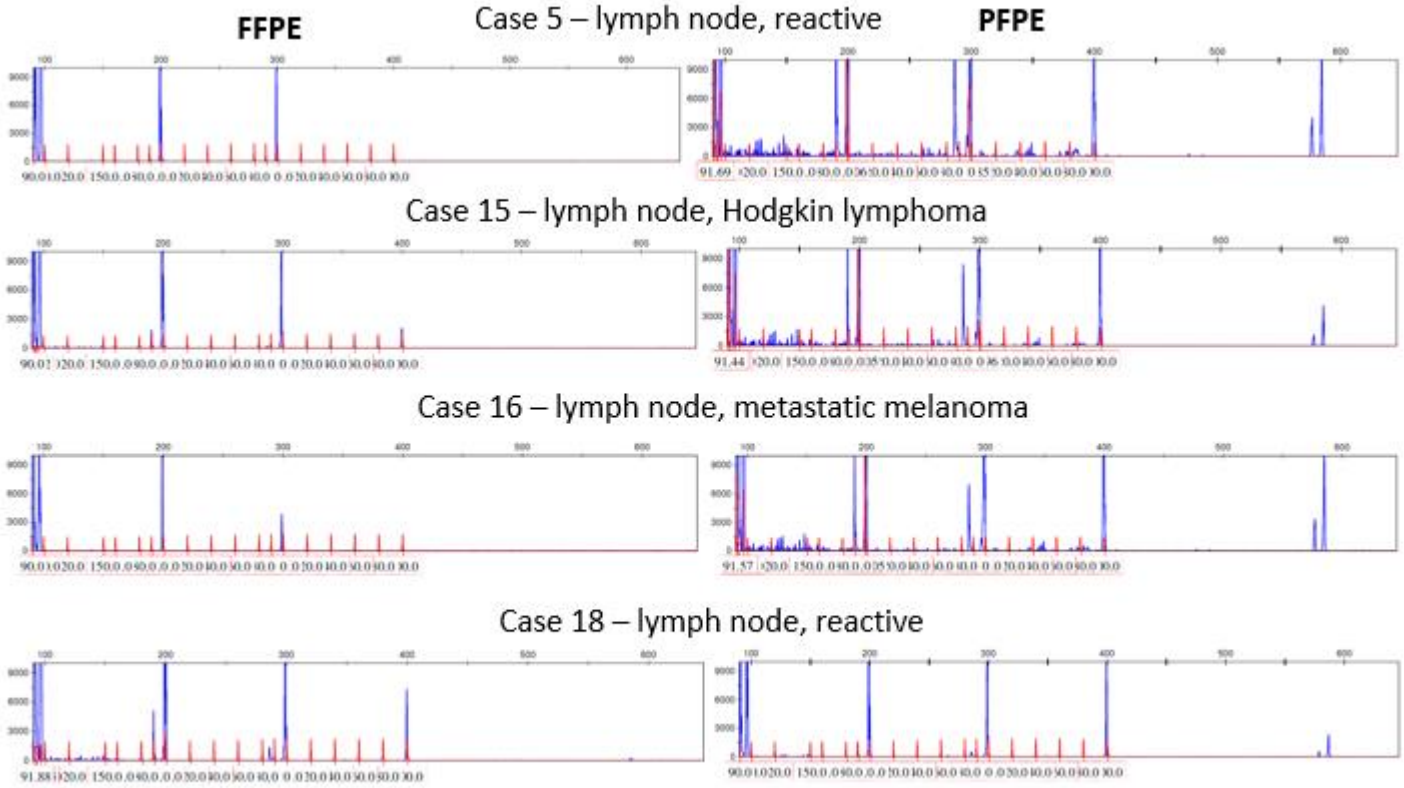

Figure 36. BIOMED2 PCR results for DNA extracted from matched formalin and PAXgene® Tissue fixed pairs of lymph node of varying histology

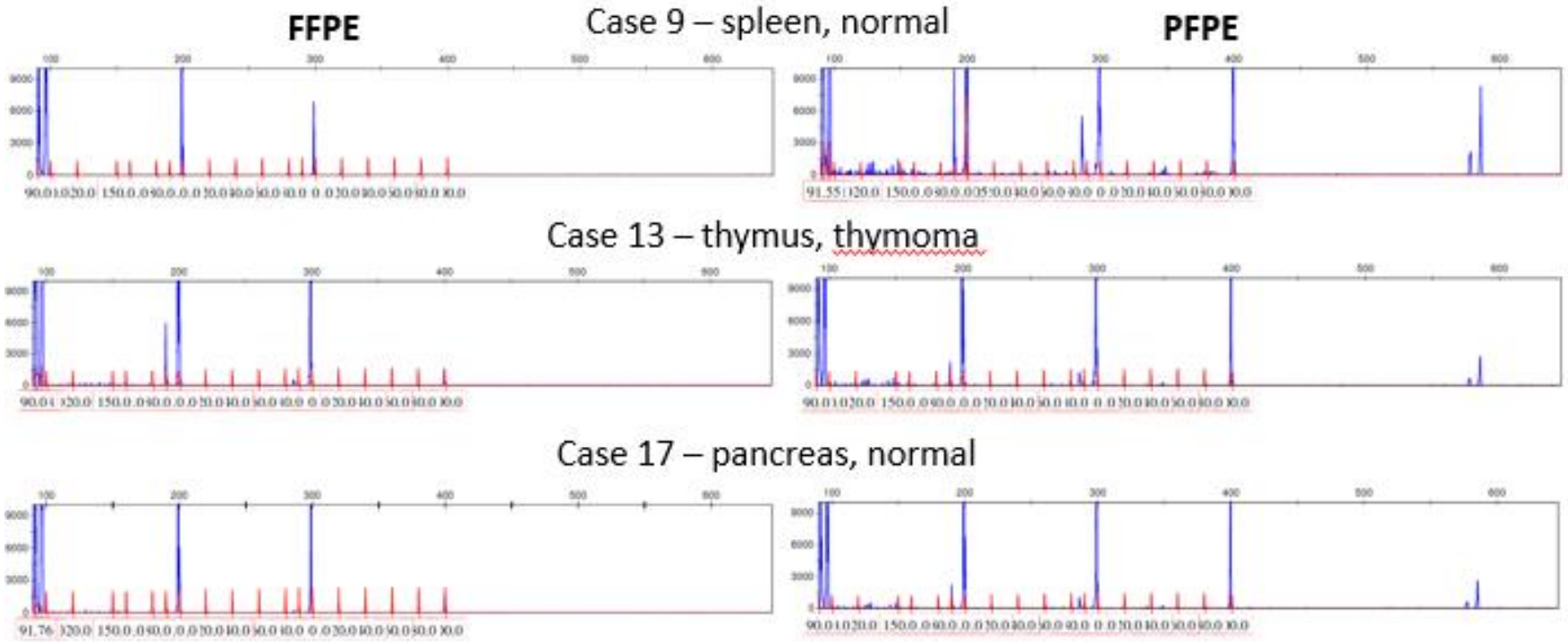

Figure 37. BIOMED2 PCR results for DNA extracted from matched formalin and PAXgene® Tissue fixed pairs of different tissue types of varying histology

Chapter 6: **Conclusions**

6.1 Rising to the challenge in cellular pathology

This is a time of great opportunity for cellular pathologists to make the most of recent developments in knowledge of the molecular basis of cancer, in order to enhance patient care and reinforce the central role of pathology in this process. Through application of new diagnostic modalities to tissue samples coming through laboratories on a daily basis, these discoveries can be translated directly into patient benefit by providing access not only to more treatment options, but particularly to those more accurately tailored to the characteristics of an individual's tumour. For this to happen, there are a number of areas to be mastered. Firstly, the knowledge and skills required to deliver stratified medicine, addressing the education need in genomics, technology and the molecular pathology of cancer among existing and future pathologists. Secondly, addressing laboratory processes to optimise sample preparation for the demands of molecular analysis, in addition to morphology assessment and immunophenotyping. There is also a need to reconfigure laboratory workflow to maximise efficiency of the process and minimize the risk of cross-contamination between samples – 'molecular hygiene'. Finally there is a need to consider the attitude of histopathologists towards tissue, specifically any feelings of ownership, and the increasing need to preserve tissue for downstream genomic applications – 'tissue economy'.

The following case study exemplifies the crucial role of the histopathologist in interpreting the results of molecular analysis and putting them into the appropriate clinical context (with thanks to Dr Darren Fowler, Consultant Histopathologist, Oxford University Hospitals NHS Trust). Rhabdoid tumour is a rare paediatric soft tissue tumour that has recently been found to be characterised by mutations in the *INI1* gene. The INI1 protein is expressed by normal muscle cells, but in rhabdoid tumour the inactivating mutation implicated in tumorigenesis leads to loss of INI1 protein expression. This can be detected in formalin-fixed, paraffin embedded tissue sections using immunohistochemistry, with the adjacent normal muscle cells acting as a positive internal control. In one particular case, the pathologist reported the

morphology and electron microscopy findings as those of rhabdoid tumour, with the diagnosis supported by immunohistochemical demonstration of loss of INI1 expression. Confirmation using a second technique was attempted and DNA was extracted from a tissue section assessed as containing greater than 90% tumour cell nuclei. The report received back from a sequencing-based investigation reported no evidence of *INI1* gene deletion in the tumour. On discussion at the paediatric oncology meeting some doubt was expressed by the oncologist about the diagnosis in view of these contradictory findings. The pathologist stood by their morphological diagnosis but agreed to attempt to provide further evidence using a different technique. Multiplex ligation-dependent probe amplification (MLPA) was performed on a further tissue sample from the paraffin block and this confirmed bi-allelic deletion of INI1 within the tumour. A possible explanation for the negative result in the sequencing is that DNA was also amplified from contaminating stromal cells admixed with the tumour, which would not have carried the *INI1* mutation and produced a wild-type result. For this reason, deletions can be difficult to detect using sequencing methods. This example exemplifies the need for clinical teams receiving results to have adequate understanding of the technical aspects, limitations and possible pitfalls of analyses requested on clinical samples, particularly an understanding of the possible explanations for false positive or negatives. Histopathologists are ideally and uniquely placed to understand and interpret the results of these investigations, with the morphological context of the originating sample being of paramount importance.

The results in chapter 4 illustrate that despite our increasingly comprehensive documentation of presumed 'standard' operating procedures and sophisticated quality management systems, sample handling in cellular pathology departments is highly variable and there is an urgent need to establish standards for optimised sample preparation in support of molecular analysis of tumour tissue for diagnostic, prognostic and predictive markers. This can be achieved through systematic experimentation into each and every one of the variables that might impact on sample quality, as detailed in this thesis, using output measures that accurately reflect the requirements of specific downstream applications. In the current era this involves confirming utility for high-throughput sequencing technologies as well as targeted methods that are

currently in use. However this would be a time-consuming and expensive exercise with endless possible combinations of different sample handling processes to evaluate and compare. A further option for collating real-world data would involve comprehensive data collection from laboratories already involved in sample preparation for molecular analysis in the NHS. The current implementation of ISO15189 standards for clinical laboratory accreditation and regular cycle of inspections provides an opportunity to collect this data, and then compare it to the outcome of molecular analysis for samples provided by each laboratory to try and identify trends and possible examples of above or below average performance which might then be linked to particular sample handling variables. It is likely to be more realistic to strive towards harmonisation rather than full standardisation of the tissue handling practices across NHS cellular pathology departments, an approach exemplified by the Confederation of Cancer Biobanks harmonisation project[137].

6.2 Contribution of the CRUK SMP

Through the activities of SMP1, it has been possible to develop an infrastructure for a national testing network in support of the delivery of molecularly stratified cancer therapeutics and clinical trials. Central to this has been establishing a collaborative, multidisciplinary network of expertise, with sharing of technical protocols and knowledge and the gradual development of consensus-based reporting and interpretation of somatic genetic aberrations in tumour samples. The flexible approach of the programme has made it possible for sites to adapt their existing NHS processes over time. Despite the use of different technologies at the outset of the programme, participating laboratories were able to demonstrate an acceptable degree of reproducibility of the results of analysis (96-100%) through participation in a bespoke EQA scheme designed and implemented in conjunction with UK NEQAS for Molecular Genetics. The use of standardised electronic messaging for test requesting and reporting was successfully implemented and favourably received at clinical and laboratory sites, with the major benefits being an electronic audit trail, reduction in duplication of data entry and also reduction of the associated risk of transcription errors.

Importantly, SMP1 demonstrated that the stratified medicine approach and genetic interrogation of tissue samples is acceptable to patients, with a consistently high consent rate among those approached. Experience in attempting to collate and use routine clinical data captured as part of usual NHS care clearly demonstrated the benefits of using existing NHS information standards such as COSD, SNOMED and TNM with data element attributes defined in the NHS data dictionary, but also demonstrated the limited capability of current systems to organise and store data in a structured format for automated retrieval at a later date. Experience shared by the laboratory staff of creating and handling the molecular data highlighted the importance of shared access databases that could be continually updated to aid the interpretation of variants, in order to classify as many as possible as either known polymorphisms or potentially actionable variants for cancer therapy. The value of using a standardised language for reporting genetic aberrations was demonstrated and facilitated aggregated data analysis.

Relatively low complete failure rates, generally less than 5%, were seen in SMP1, despite analysing DNA extracted from FFPE material for different tumour types and from many contributing pathology laboratories. It proved possible to add additional genetic markers once the programme workflow was established. Turnaround times, analytical sensitivity and the requirement for repeat testing are inextricably linked, so it was not possible to improve one parameter without detriment to another. The pursuit of detecting all mutations in a sample had to be balanced with efforts to achieve clinically meaningful turnaround times. The use of multiple testing modalities (e.g. sequencing, microsatellite analysis, FISH) to detect different types of aberration and even multiple PCR reactions to detect mutations within a single gene, proved particularly time-consuming and led to the development and adoption of a single NGS panel to cover as many of the genetic targets as possible. This technology continues to evolve during the current second phase of the Stratified Medicine Programme, as molecular pre-screening for the National Lung Matrix Trial.

The challenges experienced during SMP2 include sample size and DNA quantity and quality, due to the move to use of tissue remaining from small biopsy and cytology cell block samples following standard of care tests

(including *EGFR*, *ALK* and PD-L1). The move to NGS technology brought use of a standard technology across all three participating laboratories but work was required to design appropriate quality control steps to predict sample performance on the 28-gene panel, to establish bioinformatics approaches for reporting the new genetic aberrations and determine confidence of variant/ wild-type calling different genes, and to cross-validate the panel for reporting translocations and copy number aberrations. All these steps have been performed with a focus on delivering acceptable turnaround times for clinical trial enrolment.

Work carried out through SMP2 for chapter 5 of this thesis demonstrated that pathologists are not accustomed to counting cells or accurately quantifying proportions of different tissue components and that our assessment is semi-quantitative at best. Tumour content assessment is increasingly used to inform decisions about sample adequacy for predictive molecular and immunohistochemical analysis, the need for macrodissection and even confidence of NGS variant calling (for example to determine wild type status for certain arms of the National Lung Matrix Trial). It remains to be seen whether pathologists or other laboratory professionals can be trained to assess tumour content more reproducibly (for example ongoing work by UK NEQAS in conjunction with the 100,000 genomes project), or alternatively whether digital image analysis algorithms are the most accurate and cost-efficient method to employ.

6.3 The future direction and implementation of stratified medicine

Cellular pathologists are uniquely placed to facilitate implementation of stratified medicine into routine patient care, due to our position at the interface between clinical and laboratory medicine. This work has attempted to highlight the importance of the expertise in tissue handling that resides in cellular pathology workforce, which is the central theme of this thesis. Cellular pathologists are experts in the morphological, immunohistochemical and increasingly molecular characterisation of disease and the process of establishing of a new taxonomy of cancer and embedding the technologies to

deliver it should be embraced and led by cellular pathologists, who are already experienced integrators of multiple inputs of clinical and laboratory data and have a central role in patient management for different cancer types through clinical reporting and membership of multiple different multidisciplinary teams. That said, experience in CRUK SMP1/2 illustrate that delivering stratified medicine for cancer care is a truly multidisciplinary endeavour, requiring each speciality involved to understand just enough of the others to work collaboratively and synergistically to contribute to the end result. Stratified medicine approaches also have applications beyond malignant disease which remain to be further explored and realised.

Key roles of the cellular pathologist in stratified cancer medicine include histological subtyping of the tumour based on morphology and any necessary immunohistochemistry, to confirm that tumour tissue being analysed and provide estimate of percentage tumour cell vs. non-tumour cell nuclei. There is also a responsibility to employ protocols that optimise nucleic acid preservation of the tissue and to use the tissue sparingly, in order to leave sufficient material for immediate and future analyses. Finally the cellular pathologist should be responsible for interpretation of the tumour genotyping results, in consultation with other members of the laboratory team and issue of the final integrated histopathology report.

Further work to ensure the full realisation of stratified cancer medicine is ongoing through the CRUK SMP, 100,000 genomes and other initiatives. For CRUK SMP there is a focus on increasing the input quantity of DNA where safe and clinically appropriate to do so, either at the time of diagnostic sampling or through specific research protocol biopsies. The focus on cellular pathology involves promotion of tissue economy and establishing the evidence base for optimised tissue handling, including continuing to explore possible alternatives to formalin. A current collaboration between the 100,000 genomes project experimental pathway and STRATFix consortium involves submission of PAXgene® Tissue-fixed samples matched to formalin-fixed and/or fresh frozen samples already sequenced in the implementation initiation phase of the 100,000 genomes project to the Genomics England biorepository for whole genome sequencing. This will allow direct comparison between the different tissue preparation methods for the most demanding and comprehensive form

of cancer analysis. Through CRUK SMP2, multivariate analysis of a comprehensive tissue handing dataset for over 500 samples is underway and the results will undoubtedly provide insights into which of the many aspects of this process should be prioritised for standardisation. In parallel, Illumina are continuing to develop next generation sequencing (NGS) protocols for low input quantities of fragmented FFPE-derived nucleic acids and also work on the most appropriate QC assay to predict sample performance for sequencing applications.

It is challenging but not impossible to attempt to transform medical care in terms of both technology and approach in order to deliver stratified medicine at a time of such financial constraint in the NHS. The challenge is outweighed by the potential opportunities and benefits for patient care and the broader healthcare system and economy. This is also an ideal time to reinforce the central role of pathology - the science and study of the mechanisms of disease - and its practitioners, not only in delivering but also in advancing patient care.

As a result of the developments described in this work, it is envisaged that the management of patients with cancer will increasingly involve and be informed by broader genetic characterisation of their tumours, not only at diagnosis but also at key time points throughout the course of the disease such as relapse or recurrence to overcome issues of tumour evolution with time and treatment. This can be achieved from tumour resection or biopsy but also by analysis of cell-free circulating tumour-derived DNA, which is likely to become a more widely used technique as validation studies progress to trials and clinical implementation.

Appendices

Appendix A: Data collection spreadsheet for SMP1 clinical and technology hub key performance indicators

Clinical Hub	Month								
Tumour Type	Total number of patients approached	Total number of patients consented	Total number of patients with all specimens sent to TH		Total Number of patients with genetic results available to clinician	Number of patients with 15 working days or less between sample acquisition and results available to clinician	Number of patients with 16 - 24 working days between sample acquisition and results available to clinician	Number of patients with 25 or more working days between sample acquisition and results available to clinician	Number of patients with complete dataset in their clinical record
			Resection	Biopsy					
Colorectal									
Breast									
Prostate									
Lung									
Ovary									
Melanoma									

Appendix A

Technology Hub	**Month**

Laboratory specimen reference number	**Source Clinical Hub**	**Report date of failed sample**	**Type of sample**	**Tumour type**	**Biopsy/ Resection**	**Gene tests**	**Failure type**	**Reason for Failure (category)**	**Notes (please include details on individual assay failures)**	**Successfully re-tested? (Yes/No)**
Example 001	*Clinical Hub 1*	*1.11.11*	*Tumour*	*Breast*	*Biopsy*	*TP53*	*Individual assay*	*Insufficient quality DNA for testing*	*Exon 4*	
Example 002	*Clinical Hub 1*	*2.11.11*	*Tumour*	*Melanoma*	*Resection*	*KIT*	*Whole gene test*	*Insufficient quantity DNA for testing*		
						PIK3CA	*Individual assay*	*Insufficient quantity DNA for testing*	*Exon 9*	
Example003	*Clinical Hub 2*	*3.11.11*	*Tumour*	*Colorectal*	*Resection*	*All*	*All gene tests*	*Necrotic tissue*		

Appendix B: Cellular Pathology Department Sample Handling Cross-sectional analysis

Organization responses apply to:

Fixation

1. What type/concentration of formalin is used in the department?
2. Is microwave fixation used and if so for what type(s) of specimen?
3. Does the department supply formalin to other clinical areas such as operating theatres or do they order in a separate supply?
4. Are cancer resection specimens generally received fresh or fixed? Are there any major exceptions to this e.g. specific research tissue collections?

Breast	
Colorectal	
Lung	
Ovarian	
Prostate	

Biopsy specimen handling

5. When examining biopsy samples at multiple levels:

Sample type:	Are spare sections routinely cut?	Are unused spares stored in the archive or thrown out?	Are intervening sections between levels discarded or kept?
Breast			
Colorectal			
Lung			
Ovarian			
Prostate			

Resection specimen handling

6. Are luminal GI cancer resection specimens received intact or opened by the surgeon?

7. How long are rectal cancer resections left to fix before cutting?

8. If your department receives partial liver resections for metastatic carcinoma:

    a. Are they received fresh?

    b. Are they incised to fix?

    c. How long are they left before sampling?

    d. Is a standard operating protocol followed in the department or does practice vary between pathologists?

9. How long are radical prostatectomy specimens left to fix before cutting?

10. Are mastectomy specimens sliced upon receipt? How long are they generally left to fix further before cutting?

11. Are lung lobectomy specimens insufflated with formalin on receipt?

12. Are any tumour types routinely left for further fixation in cassettes after blocks have been taken?

Tissue processing

13. What clearing agent is used?

14. What tissue processors are available and are they used for specific specimen types?

15. Briefly what are the different programs/runs for different types of specimens (e.g. rapid for biopsies, 9h/overnight for most and any special programs for breast/megablocks etc.)?

Tissue microtomy

16. How often are microtome blades changed? What are the blades cleaned with?

17. Are there any special procedures for preparing sections for molecular work (e.g. cut first thing in the morning, dedicated microtome or new microtome blade)?

Block storage

18. Are the paraffin blocks stored at ambient/room temperature in the archive? If not, what temperature are they stored at?

Appendix C: Endobronchial ultrasound-guided lung cancer sample handling cross-sectional analysis

1. From approximately how many patients does your department receive EBUS/EUS samples in an average week? *Patients rather than samples to allow for the fact that each patient may have multiple samples taken (e.g. from different lymph node stations).*

Less than 1	
1-5	
5-10	
More than 10	
Not known	

2. How many passes does the operator make on each lymph node/area sampled in order to obtain the sample?

3. Are you able to provide an estimate or audit data for the overall percentage of EBUS/EUS samples that contain sufficient material for cytological assessment?

4. What type of samples does your department receive from EBUS/EUS procedures? *We are aware that there is some overlap in the categories below and that multiple may apply, so please select all relevant according to the terminology used in your centre and estimate the percentage of the total represented by each.*

Sample type	**Estimated percentage of total**
Sample specifically described as 'biopsy'	
Sample specifically described as 'needle washings'	
Cells in fluid suspension not otherwise specified, including FNA	
Clot	
Other – please state:	

5. Has one or more members of cellular pathology staff attended EBUS/EUS procedure lists in person to provide input or advice on technical, sample adequacy or diagnostic/interpretative issues, either in the past or currently?

Yes, currently – technical/scientific staff	
Yes, currently – medical staff	
Yes, in the past – technical/scientific staff	
Yes, in the past – medical staff	
Other – please state	

6. How long does it generally take for specimens to arrive in the cytology laboratory after the list/procedure has finished?

Same day (up to 12 hours)	
Next day (up to 24 hours)	
More than 24 hours	

Please comment on any specific contributory factors you are aware of such as need to transfer specimens between different buildings, off-site laboratory facilities etc.:

7. In what medium do you receive EBUS/EUS specimens? *If more than one please indicate and estimate percentage of total for each.*

Medium	**Estimated percentage of total**
Saline	
10% neutral buffered formalin solution	
CytoLyt	
PreservCyt	
SurePath preservative fluid	
CytoRich Red	
Other liquid based cytology (LBC) proprietary medium – please state:	
Other fixative/preservative/stabilisation solution – please state:	
Transport medium not known	

8. Which cytology processing system does your laboratory use?

LBC – SurePath	
LBC – ThinPrep	
Non-LBC conventional cytology processing system	
Other – please state:	

9. Do you make either direct smear, spread or spun cytology preparation from EBUS fluid samples?

Yes	Always		How many slides are prepared and what stain(s)?	
	Sometimes - please describe determinants below under 'comments'			
No, sample processed straight to cell block				
Comments				

10. Are any additional substances added to the sample during processing?

CytoRich Red	
Other – please state:	

11. What percentage of samples would you estimate contain sufficient material to attempt a cell block following cytology processing?

12. Is it routine practice in the laboratory to attempt a paraffin-embedded cell block on every EBUS/EUS specimen?

13. Are sections always cut from the cell block and examined at least using H&E staining?

Some laboratories process a cell block as a method of storing the residual sample but do not routinely cut and examine sections, and will only do this

dependent on the contents of the cytology preparations and in the appropriate clinical context.

Yes	
No	

14. Which of the following techniques are used in your laboratory in making a cell block?

Cytospin for pellet	
Induced clot	
Agar embedding	
Other – please state:	

15. Are there any other sample handling factors not covered above that you consider might show variability between laboratories and could be usefully explored as part of this work?

Appendix D: STRATFix project sample handling data collection pro-forma

Data item	PAXgene® sample	Formalin sample
Sample study ID		
Sample lab ID		
Tissue type		
Histological subtype		
Sampling method		
Date acquired (dd/mm/yyyy)		
Time sample obtained in theatre (hh:mm 24hr format)		
Date sample received in lab (dd/mm/yyyy)		
Time received in lab (hh:mm 24hr format)		
Date sample into PAXgene® fixative (dd/mm/yyyy)		
Time sample into PAXgene® fixative (hh:mm 24hr format)		
PAXgene® fixative batch number		
Sample size (mmxmmxmm)		
Container type		
Additional material in cassette		
Matched sample taken into formalin		
Size of matched formalin-fixed sample (mmxmmxmm)		

2+2 formalin fixation		
Date sample into PAXgene® stabiliser (dd/mm/yyyy)		
Time sample into PAXgene® stabiliser (hh:mm 24hr format)		
PAXgene® stabiliser batch number		
Storage temperature between fixation and processing (˚C)		
Temperature during sample transfer (˚C)		
Date sample onto processor (dd/mm/yyyy)		
Time sample onto processor (hh:mm 24hr format)		
Processor type		
Processor regimen		
Embedding paraffin type		
Block storage location		
Block storage temperature (˚C)		
Date sent away (dd/mm/yyyy)		
Sample type sent away		
Sendaway destination		
Date received back (dd/mm/yyyy)		

Bibliography

Campbell, MJ and Machin, D. *Medical Statistics: A commonsense approach* (2nd edition). Chichester, John Wiley & Sons, 1995.

Cook, DJ. *Cellular Pathology: An Introduction to techniques and applications* (2nd edition). Oxfordshire, Scion Publishing Limited, 2007.

Kalinithi, P. *When Breath Becomes Air.* London, Vintage, 2017.

Copson ER, Hall P, Board RE, Cook G and Selby P (Eds). Problem Solving Through Precision Oncology: A Case Study Based Reference and Learning Resource. Oxfordshire, EBN Health, 2016.

List of References

1. Davies H, Bignell GR, Cox C, et al. Mutations of the BRAF gene in human cancer. *Nature* 2002;417(6892):949-54.

2. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. *Nature* 2009;458(7239):719-24.

3. Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med* 2012;366(10):883-92.

4. Puzanov I, Amaravadi RK, McArthur GA, et al. Long-term outcome in BRAF(V600E) melanoma patients treated with vemurafenib: Patterns of disease progression and clinical management of limited progression. *Eur J Cancer* 2015;51(11):1435-43.

5. Wagle N, Emery C, Berger MF, et al. Dissecting therapeutic resistance to RAF inhibition in melanoma by tumor genomic profiling. *J Clin Oncol* 2011;29(22):3085-96.

6. Dawson SJ, Tsui DW, Murtaza M, et al. Analysis of circulating tumor DNA to monitor metastatic breast cancer. *N Engl J Med* 2013;368(13):1199-209.

7. Rakha EA, Pinder SE, Bartlett JM, et al. Updated UK Recommendations for HER2 assessment in breast cancer. *J Clin Pathol* 2015;68(2):93-9.

8. Calvo F. *From « precision medicine » towards « personalized medicine » for cancer patients. Keynote presentation at the All Ireland Cancer Consortium Conference 2015*; 2015.

9. Nowak F. Europe Does It Better: Molecular Testing across a National Health Care System—The French Example. 2013. http://meetinglibrary.asco.org/content/121-132.

10. Lonchamp EaN, F. Molecular genetic testing for equal access to targeted therapies in France in 2011: Institut National du Cancer, 2012.

11. Lonchamp EaN, F. Molecular genetic tests for access to targeted therapies in France in 2012, 2012.

12. CRUK. Cancer incidence in the UK in 2011, 2014.

13. National Cancer Screening Programmes. All breast cancer report: A UK analysis of all symptomatic and screen-detected breast cancers diagnosed in the UK in 2006, 2009.

14. Sparano JA, Gray RJ, Makower DF, et al. Prospective Validation of a 21-Gene Expression Assay in Breast Cancer. *N Engl J Med* 2015;373(21):2005-14.

15. Cardoso F, van't Veer LJ, Bogaerts J, et al. 70-Gene Signature as an Aid to Treatment Decisions in Early-Stage Breast Cancer. *N Engl J Med* 2016;375(8):717-29.

16. NICE DG10 Oncotype DX, Sept 2013.

17. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature* 2000;406(6797):747-52.

18. Curtis C, Shah SP, Chin SF, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* 2012;486(7403):346-52.

19. Robson M, Im SA, Senkus E, et al. Olaparib for Metastatic Breast Cancer in Patients with a Germline BRCA Mutation. *N Engl J Med* 2017;377(6):523-33.

20. Health and Social Care Information Centre. National Lung Cancer Audit, 2014.

21. NICE. The diagnosis and treatment of lung cancer (update). Full guideline. In: Cancer NCCF, Velindre NHS Trust, Cardiff, Wales, 2011.

22. Li T, Kung HJ, Mack PC, et al. Genotyping and genomic profiling of non-small-cell lung cancer: implications for current and future therapies. *J Clin Oncol* 2013;31(8):1039-49.

23. Bergethon K, Shaw AT, Ou SH, et al. ROS1 rearrangements define a unique molecular class of lung cancers. *J Clin Oncol* 2012;30(8):863-70.

24. Miyamae Y, Shimizu K, Hirato J, et al. Significance of epidermal growth factor receptor gene mutations in squamous cell lung carcinoma. *Oncol Rep* 2011;25(4):921-8.

25. Paik PK, Varghese AM, Sima CS, et al. Response to erlotinib in patients with EGFR mutant advanced non-small cell lung cancers with a squamous or squamous-like component. *Mol Cancer Ther* 2012;11(11):2535-40.

26. Park SH, Ha SY, Lee JI, et al. Epidermal growth factor receptor mutations and the clinical outcome in male smokers with squamous cell carcinoma of lung. *J Korean Med Sci* 2009;24(3):448-52.

27. Rekhtman N, Paik PK, Arcila ME, et al. Clarifying the spectrum of driver oncogene mutations in biomarker-verified squamous carcinoma of lung: lack of EGFR/KRAS and presence of PIK3CA/AKT1 mutations. *Clin Cancer Res* 2012;18(4):1167-76.

28. Mok TS, Wu Y-L, Thongprasert S, et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. *N Engl J Med* 2009;361(1533-4406 (Electronic)):947-57.

29. Sequist LV, Bell DW, Lynch TJ, et al. Molecular predictors of response to epidermal growth factor receptor antagonists in non-small-cell lung cancer. *J Cliln Oncol* 2007;25(1527-7755 (Electronic)):587-95.

30. Sharma SV, Bell DW, Settleman J, et al. Epidermal growth factor receptor mutations in lung cancer. *Nat Rev Cancer* 2007;7(3):169-81.

31. Zhang YL, Yuan JQ, Wang KF, et al. The prevalence of EGFR mutation in patients with non-small cell lung cancer: a systematic review and meta-analysis. *Oncotarget* 2016;7(48):78985-93.

32. Tsao MS, Sakurada A, Cutz JC, et al. Erlotinib in lung cancer - molecular and clinical predictors of outcome. *N Engl J Med* 2005;353(2):133-44.

33. Leary AF, Castro DG, Nicholson AG, et al. Establishing an EGFR mutation screening service for non-small cell lung cancer - sample quality criteria and candidate histological predictors. *Eur J Cancer* 2012;48(1):61-7.

34. Karapetis CS, Khambata-Ford S, Jonker DJ, et al. K-ras mutations and benefit from cetuximab in advanced colorectal cancer. *N Engl J Med* 2008;359(17):1757-65.

35. Vermorken JB, Stohlmacher-Williams J, Davidenko I, et al. Cisplatin and fluorouracil with or without panitumumab in patients with recurrent or metastatic squamous-cell carcinoma of the head and neck (SPECTRUM): an open-label phase 3 randomised trial. *Lancet Oncol* 2013;14(8):697-710.

36. Shaw AT, Yeap BY, Mino-Kenudson M, et al. Clinical features and outcome of patients with non-small-cell lung cancer who harbor EML4-ALK. *J Clin Oncol* 2009;27(26):4247-53.

37. Inamura K, Takeuchi K, Togashi Y, et al. EML4-ALK fusion is linked to histological characteristics in a subset of lung cancers. *J Thorac Oncol* 2008;3(1):13-7.

38. Kohno T, Ichikawa H, Totoki Y, et al. KIF5B-RET fusions in lung adenocarcinoma. *Nat Med* 2012;18(3):375-7.

39. Shaw AT, Yeap BY, Solomon BJ, et al. Effect of crizotinib on overall survival in patients with advanced non-small-cell lung cancer harbouring ALK gene rearrangement: a retrospective analysis. *Lancet Oncol* 2011;12(11):1004-12.

40. Kwak EL, Bang YJ, Camidge DR, et al. Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. *N Engl J Med* 2010;363(18):1693-703.

41. Katayama R, Shaw AT, Khan TM, et al. Mechanisms of acquired crizotinib resistance in ALK-rearranged lung Cancers. *Sci Transl Med* 2012;4(120):120ra17.

42. Carpenter EL, Mosse YP. Targeting ALK in neuroblastoma--preclinical and clinical advancements. *Nat Rev Clin Oncol* 2012;9(7):391-9.

43. Shaw AT, Kim D-W, Mehra R, et al. Ceritinib in ALK-rearranged non-small-cell lung cancer. *N Engl J Med* 2014;370(26):2537-9.

44. De Roock W, Jonker DJ, Di Nicolantonio F, et al. Association of KRAS p.G13D mutation with outcome in patients with chemotherapy-refractory metastatic colorectal cancer treated with cetuximab. *JAMA* 2010;304(16):1812-20.

45. Parkin DM, Mesher D, Sasieni P. 13. Cancers attributable to solar (ultraviolet) radiation exposure in the UK in 2010. *Br J Cancer* 2011;105 Suppl 2:S66-9.

46. Boniol M, Autier P, Boyle P, et al. Cutaneous melanoma attributable to sunbed use: systematic review and meta-analysis. *BMJ* 2012;345:e4757.

47. Kirkham N. Bordeline melanocytic lesions. In: M. NBJaG (ed.) *Melanoma: Critical debates.*: John Wiley & Sons; 2008.

48. Brose MS, Volpe P, Feldman M, et al. BRAF and RAS mutations in human lung cancer and melanoma. *Cancer Res* 2002;62(23):6997-7000.

49. Long GV, Menzies AM, Nagrial AM, et al. Prognostic and clinicopathologic associations of oncogenic BRAF in metastatic melanoma. *J Clin Oncol* 2011;29(10):1239-46.

50. Chapman PB, Hauschild A, Robert C, et al. Improved survival with vemurafenib in melanoma with BRAF V600E mutation. *N Engl J Med* 2011;364(26):2507-16.

51. Flaherty KT, Infante JR, Daud A, et al. Combined BRAF and MEK inhibition in melanoma with BRAF V600 mutations. *N Engl J Med* 2012;367(18):1694-703.

52. Flaherty KT, Robert C, Hersey P, et al. Improved survival with MEK inhibition in BRAF-mutated melanoma. *N Engl J Med* 2012;367(2):107-14.

53. Bos JL. ras oncogenes in human cancer: a review. *Cancer Res* 1989;49(17):4682-9.

54. Devitt B, Liu W, Salemi R, et al. Clinical outcome and pathological features associated with NRAS mutation in cutaneous melanoma. *Pigment Cell Melanoma Res* 2011;24(4):666-72.

55. Jakob JA, Bassett RL, Jr., Ng CS, et al. NRAS mutation status is an independent prognostic factor in metastatic melanoma. *Cancer* 2012;118(16):4014-23.

56. Ellerhorst JA, Greene VR, Ekmekcioglu S, et al. Clinical correlates of NRAS and BRAF mutations in primary human melanoma. *Clin Cancer Res* 2011;17(2):229-35.

57. Curtin JA, Busam K, Pinkel D, et al. Somatic activation of KIT in distinct subtypes of melanoma. *J Clin Oncol* 2006;24(26):4340-6.

58. Jin SA, Chun SM, Choi YD, et al. BRAF mutations and KIT aberrations and their clinicopathological correlation in 202 Korean melanomas. *J Invest Dermatol* 2013;133(2):579-82.

59. CRUK. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer.

60. Ledermann J, Harter P, Gourley C, et al. Olaparib maintenance therapy in patients with platinum-sensitive relapsed serous ovarian cancer: a

preplanned retrospective analysis of outcomes by BRCA status in a randomised phase 2 trial. *Lancet Oncol* 2014;15(8):852-61.

61. Sun X, Huang J, Homma T, et al. Genetic alterations in the PI3K pathway in prostate cancer. *Anticancer Res* 2009;29(5):1739-43.

62. Demichelis F, Fall K, Perner S, et al. TMPRSS2:ERG gene fusion associated with lethal prostate cancer in a watchful waiting cohort. *Oncogene* 2007;26(31):4596-9.

63. Mosquera JM, Mehra R, Regan MM, et al. Prevalence of TMPRSS2-ERG fusion prostate cancer among men undergoing prostate biopsy in the United States. *Clin Cancer Res* 2009;15(14):4706-11.

64. Brenner JC, Ateeq B, Li Y, et al. Mechanistic rationale for inhibition of poly(ADP-ribose) polymerase in ETS gene fusion-positive prostate cancer. *Cancer Cell* 2011;19(5):664-78.

65. Tsiatis AC, Norris-Kirby A, Rich RG, et al. Comparison of Sanger sequencing, pyrosequencing, and melting curve analysis for the detection of KRAS mutations: diagnostic and clinical implications. *J Mol Diagn* 2010;12(4):425-32.

66. Little S. Amplification-refractory mutation system (ARMS) analysis of point mutations. *Curr Protoc Hum Genet* 2001;Chapter 9:Unit 9 8.

67. Lih CJ, Harrington RD, Sims DJ, et al. Analytical Validation of the Next-Generation Sequencing Assay for a Nationwide Signal-Finding Clinical Trial: Molecular Analysis for Therapy Choice Clinical Trial. *J Mol Diagn* 2017;19(2):313-27.

68. Sheffield VC, Beck JS, Kwitek AE, et al. The sensitivity of single-strand conformation polymorphism analysis for the detection of single base substitutions. *Genomics* 1993;16(2):325-32.

69. Lang AH, Drexel H, Geller-Rhomberg S, et al. Optimized allele-specific real-time PCR assays for the detection of common mutations in KRAS and BRAF. *J Mol Diagn* 2011;13(1):23-8.

70. Cree IA, Deans Z, Ligtenberg MJ, et al. Guidance for laboratories performing molecular pathology for cancer patients. *J Clin Pathol* 2014;67(11):923-31.

71. Lopez-Rios F, Angulo B, Gomez B, et al. Comparison of molecular testing methods for the detection of EGFR mutations in formalin-fixed paraffin-embedded tissue specimens of non-small cell lung cancer. *J Clin Pathol* 2013;66(5):381-5.

72. Smits AJ, Kummer JA, de Bruin PC, et al. The estimation of tumor cell percentage for molecular testing by pathologists is not accurate. *Mod Pathol* 2014;27(2):168-74.

73. Hamilton PW, Wang Y, Boyd C, et al. Automated tumor analysis for molecular profiling in lung cancer. *Oncotarget* 2015.

74. Gulley ML, Braziel RM, Halling KC, et al. Clinical laboratory reports in molecular pathology. *Arch Pathol Lab Med* 2007;131(6):852-63.

75. Fox CH, Johnson FB, Whiting J, et al. Formaldehyde fixation. *J Histochem Cytochem* 1985;33(8):845-53.

76. Rait VK ZQ, Fabris D et al.. Conversions of formaldehyde-modified 2'-deoxyadenosine 5'-monophosphate in conditions modeling formalin-fixed tissue dehydration *Journal of Histochemistry & Cytochemistry* 2006;54(3):301-10.

77. Medawar PB. The rate of penetration of fixatives. *J Roy Micr Soc* 1941;61:46.

78. Underhill B. The rate of penetration of fixatives. *Journal of the Royal Microscopical Society* 1932:113-20.

79. Eckhart L, Bach J, Ban J, et al. Melanin binds reversibly to thermostable DNA polymerase and inhibits its activity. *Biochem Biophys Res Commun* 2000;271(3):726-30.

80. Williams C, Ponten F, Moberg C, et al. A high frequency of sequence alterations is due to formalin fixation of archival specimens. *Am J Pathol* 1999;155(5):1467-71.

81. Do H, Wong SQ, Li J, et al. Reducing sequence artifacts in amplicon-based massively parallel sequencing of formalin-fixed paraffin-embedded DNA by enzymatic depletion of uracil-containing templates. *Clin Chem* 2013;59(9):1376-83.

82. Srinivasan M, Sedmak D, Jewell S. Effect of fixatives and tissue processing on the content and integrity of nucleic acids. *Am J Pathol* 2002;161(6):1961-71.

83. Chung JY, Braunschweig T, Williams R, et al. Factors in tissue handling and processing that impact RNA obtained from formalin-fixed, paraffin-embedded tissue. *J Histochem Cytochem* 2008;56(11):1033-42.

84. RCPath Professional Support Unit. Review of cellular pathology governance, breast reporting and immunohistochemistry at Sherwood Forest Hospitals NHS Foundation Trust. 2013. http://www.rcpath.org/publications-media/news-and-press-releases/news-and-press-releases. (accessed 1st April 2015).

85. UK NEQAS. Run 103/32 Immunocytochemistry modules. 2013.

86. Baloglu G, Haholu A, Kucukodaci Z, et al. The effects of tissue fixation alternatives on DNA content: a study on normal colon tissue. *Appl Immunohistochem Mol Morphol* 2008;16(5):485-92.

87. Loughrey MB. Standards and datasets for reporting cancers: Dataset for colorectal cancer histopathology reports July 2014. Royal College of Pathologists, London <G049_ColorectalDataset_July14.pdf>. 2014. http://www.rcpath.org/publications-media/publications/datasets/colorectal-cancer.

88. Hammond ME, Hayes DF, Wolff AC. Clinical Notice for American Society of Clinical Oncology-College of American Pathologists guideline recommendations on ER/PgR and HER2 testing in breast cancer. *J Clin Oncol* 2011;29(15):e458.

89. Compton C, Fenoglio-Preiser CM, Pettigrew N, et al. American Joint Committee on Cancer Prognostic Factors Consensus Conference: Colorectal Working Group. *Cancer* 2000;88(7):1739-57.

90. Compton CC. Pathologic prognostic factors in the recurrence of rectal cancer. *Clin Colorectal Cancer* 2002;2(3):149-60.

91. Compton CC, Fielding LP, Burgart LJ, et al. Prognostic factors in colorectal cancer. College of American Pathologists Consensus Statement 1999. *Arch Pathol Lab Med* 2000;124(7):979-94.

92. Cserni G, Vinh-Hung V, Burzykowski T. Is there a minimum number of lymph nodes that should be histologically assessed for a reliable nodal staging of T3N0M0 colorectal carcinomas? *J Surg Oncol* 2002;81(2):63-9.

93. Fracasso T, Heinrich M, Hohoff C, et al. Ultrasound-accelerated formalin fixation improves the preservation of nucleic acids extraction in histological sections. *Int J Legal Med* 2009;123(6):521-5.

94. Bass BP, Engel KB, Greytak SR, et al. A review of preanalytical factors affecting molecular, protein, and morphological analysis of formalin-fixed, paraffin-embedded (FFPE) tissue: how well do you know your FFPE specimen? *Arch Pathol Lab Med* 2014;138(11):1520-30.

95. Kapp JR, Diss T, Spicer J, et al. Variation in pre-PCR processing of FFPE samples leads to discrepancies in BRAF and EGFR mutation detection: a diagnostic RING trial. *J Clin Pathol* 2015;68(2):111-8.

96. Simbolo M, Gottardi M, Corbo V, et al. DNA qualification workflow for next generation sequencing of histopathological samples. *PLoS One* 2013;8(6):e62692.

97. Sedlackova T, Repiska G, Celec P, et al. Fragmentation of DNA affects the accuracy of the DNA quantitation by the commonly used methods. *Biol Proced Online* 2013;15(1):5.

98. van Dongen JJ, Langerak AW, Bruggemann M, et al. Design and standardization of PCR primers and protocols for detection of clonal immunoglobulin and T-cell receptor gene recombinations in suspect lymphoproliferations: report of the BIOMED-2 Concerted Action BMH4-CT98-3936. *Leukemia* 2003;17(12):2257-317.

99. Sah S, Chen L, Houghton J, et al. Functional DNA quantification guides accurate next-generation sequencing mutation detection in formalin-fixed, paraffin-embedded tumor biopsies. *Genome Med* 2013;5(8):77.

100. Formaldehyde, 2-butoxyethanol and 1-tert-butoxypropan-2-ol. *IARC Monogr Eval Carcinog Risks Hum* 2006;88:1-478.

101. Chafin D, Theiss A, Roberts E, et al. Rapid two-temperature formalin fixation. *PLoS One* 2013;8(1):e54138.

102. Bussolati G, Chiusa L, Cimino A, et al. Tissue transfer to pathology labs: under vacuum is the safe alternative to formalin. *Virchows Arch* 2008;452(2):229-31.

103. (UICC) IUAC. *The TNM Classification of Malignant Tumours 7th edition*: Wiley-Blackwell; 2009.

104. Horlings HM, Flanagan AM, Huntsman DG. Categorization of cancer through genomic complexity could guide research and management strategies. *J Pathol* 2015.

105. Craft WF, Conway JA, Dark MJ. Comparison of histomorphology and DNA preservation produced by fixatives in the veterinary diagnostic laboratory setting. *PeerJ* 2014;2:e377.

106. Shaw EC, Hanby AM, Wheeler K, et al. Observer agreement comparing the use of virtual slides with glass slides in the pathology review component of the POSH breast cancer cohort study. *J Clin Pathol* 2012;65(5):403-8.

107. Wright AI, Grabsch HI, Treanor DE. RandomSpot: A web-based tool for systematic random sampling of virtual slides. *J Pathol Inform* 2015;6:8.

108. Landis JR and Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159-74.

109. The Cancer Genome Atlas Research Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature* 2014;511(7511):543-50.

110. Ding L, Getz G, Wheeler DA, et al. Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* 2008;455(7216):1069-75.

111. Pao W, Wang TY, Riely GJ, et al. KRAS mutations and primary resistance of lung adenocarcinomas to gefitinib or erlotinib. *PLoS Med* 2005;2(1):e17.

112. Cagle PT, Allen TC. Lung cancer genotype-based therapy and predictive biomarkers: present and future. *Arch Pathol Lab Med* 2012;136(12):1482-91.

113. Forbes SA, Beare D, Gunasekaran P, et al. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Research* 2015;43(Database issue):D805-11.

114. Lindeman NI, Cagle PT, Beasley MB, et al. Molecular testing guideline for selection of lung cancer patients for EGFR and ALK tyrosine kinase inhibitors: guideline from the College of American Pathologists, International Association for the Study of Lung Cancer, and Association for Molecular Pathology. *Arch Pathol Lab Med* 2013;137(6):828-60.

115. Benvenuti S, Frattini M, Arena S, et al. PIK3CA cancer mutations display gender and tissue specificity patterns. *Human Mutation* 2008;29(2):284-88.

116. Banck MS, Grothey A. Biomarkers of Resistance to Epidermal Growth Factor Receptor Monoclonal Antibodies in Patients with Metastatic Colorectal Cancer. *Clin Cancer Res* 2009;15(24):7492-501.

117. Mao C, Liao RY, Qiu LX, et al. BRAF V600E mutation and resistance to anti-EGFR monoclonal antibodies in patients with metastatic colorectal cancer: a meta-analysis. *Mol Biol Rep* 2011;38(4):2219-23.

118. Irahara N, Baba Y, Nosho K, et al. NRAS mutations are rare in colorectal cancer. *Diagn Mol Pathol* 2010;19(3):157-63.

119. Oden-Gangloff A, Di Fiore F, Bibeau F, et al. TP53 mutations predict disease control in metastatic colorectal cancer treated with cetuximab-based chemotherapy. *Br J Cancer* 2009;100(8):1330-5.

120. The Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature* 2011;474(7353):609-15.

121. Kurman RJ, Shih Ie M. Molecular pathogenesis and extraovarian origin of epithelial ovarian cancer--shifting the paradigm. *Hum Pathol* 2011;42(7):918-31.

122. Levine DA, Bogomolniy F, Yee CJ, et al. Frequent mutation of the PIK3CA gene in ovarian and breast cancers. *Clin Cancer Res* 2005;11(8):2875-8.

123. Yamamoto S, Tsuda H, Takano M, et al. PIK3CA mutation is an early event in the development of endometriosis-associated ovarian clear cell adenocarcinoma. *J Pathol* 2011;225(2):189-94.

124. Gounaris I, Charnock-Jones DS, Brenton JD. Ovarian clear cell carcinoma--bad endometriosis or bad endometrium? *J Pathol* 2011;225(2):157-60.

125. Martini M, Ciccarone M, Garganese G, et al. Possible involvement of hMLH1, p16(INK4a) and PTEN in the malignant transformation of endometriosis. *Int J Cancer* 2002;102(4):398-406.

126. Mosquera JM, Perner S, Demichelis F, et al. Morphological features of TMPRSS2-ERG gene fusion prostate cancer. *J Pathol* 2007;212(1):91-101.

127. Lee JH, Choi JW, Kim YS. Frequencies of BRAF and NRAS mutations are different in histological types and sites of origin of cutaneous melanoma: a meta-analysis. *Br J Dermatol* 2011;164(4):776-84.

128. Colombino M, Capone M, Lissia A, et al. BRAF/NRAS mutation frequencies among primary tumors and metastases in patients with melanoma. *J Clin Oncol* 2012;30(20):2522-9.

129. Beadling C, Jacobson-Dunlop E, Hodi FS, et al. KIT gene mutations and copy number in melanoma subtypes. *Clin Cancer Res* 2008;14(21):6821-8.

130. Curtin JA, Stark MS, Pinkel D, et al. PI3-kinase subunits are infrequent somatic targets in melanoma. *J Invest Dermatol* 2006;126(7):1660-3.

131. Omholt K, Krockel D, Ringborg U, et al. Mutations of PIK3CA are rare in cutaneous melanoma. *Melanoma Res* 2006;16(2):197-200.

132. Ellis IO, Coleman D, Wells C, et al. Impact of a national external quality assessment scheme for breast pathology in the UK. *J Clin Pathol* 2006;59(2):138-45.

133. Groelz D, Sobin L, Branton P, et al. Non-formalin fixative versus formalin-fixed tissue: a comparison of histology and RNA quality. *Exp Mol Pathol* 2013;94(1):188-94.

134. Kap M, Smedts F, Oosterhuis W, et al. Histological assessment of PAXgene® tissue fixation and stabilization reagents. *PLoS One* 2011;6(11):e27704.

135. Gillard M, Tom WR, Antic T, et al. Next-gen tissue: preservation of molecular and morphological fidelity in prostate tissue. *Am J Transl Res* 2015;7(7):1227-35.

136. Oberauner-Wappis L, Loibner M, Viertler C, et al. Protocol for HER2 FISH determination on PAXgene®-fixed and paraffin-embedded tissue in breast cancer. *Int J Exp Pathol* 2016;97(2):202-6.

137. Confederation of Cancer Biobanks. *Biobank Quality Standard: Collecting, storing and providing human biological material and data for research.*; 2014. http://ccb.ncri.org.uk/biobanking/harmonisation-and-benchmarking.