Evaluating the clinical utility of Tumour Mutational Burden assessment in combination with PD-L1 expression analysis in guiding immunotherapy treatment stratification in a Welsh lung cancer patient group

H M Roberts

PhD 2022

Evaluating the clinical utility of Tumour Mutational Burden assessment in combination with PD-L1 expression analysis in guiding immunotherapy treatment stratification in a Welsh lung cancer patient group

A thesis submitted to
Manchester Metropolitan University
for the degree of Doctor of Clinical Science
in the Faculty of Life Sciences

2022

Helen Marie Roberts

School of Biological Sciences

Contents

Word count: 39438

List of figures

List of tables

Appendix list

Abbreviations

AWMGS	All Wales Medical Genomics Service
ctDNA	Circulating tumour DNA
DNA	Deoxyribonucleic acid
d.p.	Decimal place
EQA	External Quality Assurance
EMQN	European Molecular Genetics Quality Network
FDA	U.S. Food and Drug Administration

FFPE	Formalin-fixed paraffin embedded
FN	False negative
FP	False positive
IHC	Immunohistochemistry
IGV	Integrated Genomics Viewer
Indels	Insertion-deletions
mAbs	Monoclonal antibodies
Mb	Megabase
NEQAS	National External Quality Assessment Service
NGS	Next Generation Sequencing
NICE	National Institute for Health and Care Excellence
NHSE	National Health Service England
NSCLC	Non-Small Cell Lung Cancer
ORR	Objective response rate
PD-1	Programmed cell death-1
PD-L1	Programmed cell death-ligand 1
PFS	Progression free survival
RECIST	Response Evaluation Criteria in Solid Tumours
ROC	Receiver operating characteristic
s.f.	Significant figure
SNP	Single Nucleotide Polymorphism
TMB	Tumour Mutational Burden
TN	True negative
TP	True positive
WCB	Welsh Cancer Bank
WES	Whole Exome Sequencing
WGS	Whole Genome Sequencing
+artefacts	Artefacts not removed
-artefacts	Artefacts removed
+synonymous	Synonymous variants included
-synonymous	Synonymous variants not included

Abstract

The immunotherapy, pembrolizumab, is not effective in the treatment of all lung cancer patients. Stratification of the use of this drug in non-small cell lung cancer (NSCLC) patient care is currently performed within the NHS using immunohistochemistry (IHC)-based PD-L1 expression analysis, but patient response rates following stratification remain low at 45%. A more accurate predictor of immunotherapy response is desirable to minimise the use of ineffective and costly therapy. Tumour Mutational Burden (TMB), defined as the number of somatic mutations[1] found within a tumour, has been identified in numerous research studies as a potential new biomarker for immunotherapy stratification in lung cancer patients either alone or in combination with PD-L1 expression analysis. Publications show that high TMB, quantified using either Whole Exome Sequencing (WES), or targeted Next Generation Sequencing (NGS), is associated with immunotherapy response. Following the recent United States Food and Drug Administration (FDA) approval of the TMB-stratified use of pembrolizumab for solid tumours including lung cancers, UK approval for TMB-based immunotherapy stratification via National Institute for Health and Care Excellence (NICE) may be granted in the coming years, replacing or supplementing the suboptimal PD-L1 expression analysis. With such approval would come the requirement for NHS Genomics laboratories to deliver TMB services.

Despite the considerable international interest in TMB as a biomarker, there remains a lack of consensus in how TMB is calculated. TMB-focused studies to date show differences in the NGS panels used to determine variant number, the variants included within the TMB assessment, and the definition of 'high TMB' via the use of different TMB thresholds to separate likely responders from non-responders. The clinical impact of these variables has to be understood and controlled prior to service implementation within the NHS to ensure high quality services are provided to patients. This research study aimed to produce novel data regarding the impact of these variables on TMB score and TMB status. This study has provided increased understanding in this area by demonstrating the impact on TMB estimation and TMB high status when three NGS panels (Illumina TruSight™ Oncology 500 panel, Agilent SureSelect Community Design Glasgow Cancer Core panel, and Nonacus Cell3™ Target: Pan Cancer panel) targeting varying proportions (1.58-1.94Mb) of the genome, were used to determine TMB using different TMB quantification methods on the same cohort of Welsh NSCLC patients with high PD-L1 expression status and known pembrolizumab response status. TMB quantification for all three NGS panels was performed using the Institut Curie TMB tool. TMB values were generated following the application of different variant filtering parameters based on the inclusion/exclusion of sequencing artefacts, which is an area not well-researched currently in terms of impact on TMB, and the inclusion/exclusion of synonymous variants, which is an area of difference within TMB publications. The utility of ROC curve generated TMB high thresholds for immunotherapy response prediction were evaluated alongside a 10 variants/Mb threshold, which is a threshold used in a number of TMB publications. This evaluation enabled the primary research question to be answered by demonstrating the potential clinical utility of a combined TMB and PD-L1 biomarker for immunotherapy response. Sequencing data from the Illumina and Nonacus panels highlighted an increase in sensitivity for the separation of

[1] From this point on within this thesis, mutations in the tumour will be referred to as 'variants' in line with existing practices within the All Wales Medical Genomics Service (AWMGS).

responders and non-responders when a combined TMB and PD-L1 biomarker was used compared to the use of PD-L1 expression analysis alone. The Agilent NGS panel failed to produce any sequencing data above the minimum coverage level. The study identified elements of analysis providing optimal TMB quantification, and generated suggestions for minimising the clinical impact of panel- and analysis-dependent TMB variation to improve the clinical utility of TMB as a biomarker.

Given the small size of the cohort (n=17), limited by the cost of NGS and the financial constraints of this research, this thesis represents a pilot study. The findings could be used to shape the design of future larger scale research studies evaluating the utility of different panel/analysis combinations, or to drive further research into the clinical utility of TMB in a larger Welsh cohort, which would be more representative of the Welsh population as a whole and would provide more weighting to the findings of this small pilot study.

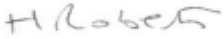
The study makes recommendations that could guide NHS Genomic laboratories in how to progress TMB service validations, and which could contribute to future best practice guidelines for TMB service delivery. These recommendations support the use of: NGS panels >1.6Mb in size, the Institut Curie TMB tool, and ROC curves in TMB evaluation, whilst the need for artefact removal prior to TMB calculation is not favoured. The feasibility of TMB service implementation within the NHS environment was highlighted by the potential cost neutral status of a TMB service and the recent launch of an External Quality Assurance (EQA) pilot scheme for TMB quantification (Abate 2020); recommendations for future EQA schemes are provided.

Declaration

DNA extractions were performed by the AWMGS extraction team. Amy Powell, AWMGS trainee Clinical Scientist, supported in NGS kit and sample acquisition; NGS library preparations were supported by AWMGS Genetic Technologists, Jenny Waizeneker, Arfhan Rafiq, and Pablo Reviriego. Assistance with bioinformatic analysis of NGS data was provided by Christian Files, AWMGS Bioinformatician.

Helen Roberts was solely responsible for all other aspects of the study including the literature review, study design, project cost analysis, WCB application for sample acquisition, statistical analysis of data, and preparation of the thesis.

With the exception of any statements to the contrary, all the data presented in this report are the result of my own efforts. In addition, no parts of this report have been copied from other sources. I understand that any evidence of plagiarism and/or the use of unacknowledged third-party data will be dealt with as a very serious matter.



Signed:

Date: 30th November 2022

Copyright statement

Acknowledgements

Thanks go to my academic supervisor, Jane Sarginson, and workplace supervisor and AWMGS Head of Cancer, Rhian White, for their advice, support and guidance. I would also like to express my thanks to the clinical, technical and bioinformatics staff at AWMGS for supporting this project, in particular to Amy Powell, Jenny Waizeneker, Arfhan Rafiq, Pablo Reviriego, and Christian Files. Thanks also to the Welsh Cancer Bank (WCB) for the use of the samples, and to the cancer patients of Wales who have contributed tumour samples to the WCB making this research possible.

This research project was funded by AWMGS. Thanks to Sian Morgan, Head of AWMGS, for allowing me to undertake this project within the AWMGS laboratory.

Finally, thanks go to my husband Joseph and children, Albie and George, for all the laughs, hugs, patience and support whilst I have completed this research.

Preface

This professional doctorate was completed as part of the National School of Healthcare Science Higher Specialist Scientist Training (HSST) programme in the specialty of Molecular Pathology of Acquired Disease within the Life Sciences Division. This is a bespoke five-year workplace-based training programme supported by a doctoral-level academic award and Royal College qualifications.

The work-based training component of the HSST programme is overseen by the National School of Healthcare Science. The academic component of HSST, supported by Manchester Academy for Healthcare Scientist Education (MAHSE), is the Doctor of Clinical Science (DClinSci), which is divided into three sections (A-C).

This thesis fulfils the requirements of the Research, Development and Innovation section of the Doctor of Clinical Science (DClinSci) Life Sciences programme (section C2). Section C2 (200 credits) research project aims to improve health and health outcomes and may include scientific, clinical, service transformation, innovation, leadership, policy, education or educational research. The other element of the Research, Development and Innovation section is section C1 (70 credits), for which an innovation proposal and business case were submitted (appendices 1, 2 and 3).

Section A (120 credits) of the DClinSci focusses on Leadership and Professional Development. Section A is supported by the Alliance Manchester Business School (AMBS) and is composed of five modules (appendix 1). Section B (150 credits) of the DClinSci is covered by the Royal College of Pathologists (FRCPath) part 1 examination, which I obtained in 2015 in the speciality of Molecular Genetics (appendix 4). In October 2022, I sat the FRCPath part 2 examination (practical and oral) in the speciality of Molecular Pathology. This thesis will be used to satisfy the written component of the FRCPath part 2 to obtain Fellowship of the Royal College of Pathologists (FRCPath).

Previously I have attained a BSc in Genetics from the University of Sheffield. I currently work full time as a Health and Care Professions Council (HCPC) registered Principal Clinical Scientist at the AWMGS based in Cardiff as Head of Solid Tumour Services. In this Clinical Scientist role, and alongside this professional doctorate, I have had the opportunity to be involved in two national clinical trials for solid tumour patients as well as having led NHS service developments, from which I have contributed to publications or have presented findings:

- Richman, S.D., Adams, R., Quirke, P., Butler, R., Hemmings G., Chambers, P., **Roberts, H**., James, M.D., Wozniak, S., Bathia, R., Pugh, C., Maughan, T., Jasani, B. (2016) Pre-trial inter-laboratory analytical validation of the FOCUS4 personalised therapy trial. *J Clin Pathol* 69:35-41.
- **Roberts, H.,** Cobreros-Ugidos, M., White, R., Morgan, S. (2020) *The complexities of the delivery of an RNA sequencing service for FFPE solid tumour samples within an NHS setting. MAP conference* (Molecular Analysis for Precision Oncology).

Chapter 1: Introduction

Lung cancer is one of the most common cancers, with over 43,000 new lung cancers diagnosed each year in the UK (NICE, 2021). The most common form of lung cancer is non-small cell lung cancer (NSCLC), making up 87% of cases (NHS, 2022). As well as surgery, chemotherapy and radiotherapy, there are a number of targeted therapies available to specific lung cancer patient groups, the use of which is dependent on specific tumour attributes; this could be the presence of specific gene variants in the tumour (e.g. *EGFR* gene variants) or expression of a particular protein on the tumour cell surface (e.g. PD-L1 expression). These targeted approaches are beneficial over traditional chemotherapy or radiotherapy as they reduce the risk of adverse patient side effects by specifically targeting cancer cells. The All Wales Medical Genomics Service (AWMGS) has been providing targeted *EGFR* gene testing for NSCLC patients since 2010, whereby tumour samples are analysed to detect specific *EGFR* variants that dictate either a sensitivity or a resistance to *EGFR*-targeted therapy. Other genomic tests, aligned to the use of specific targeted treatments, are also now available to NSCLC patients for the interrogation of the *KRAS*, *BRAF*, *ALK*, *ROS1*, *NTRK1/2/3*, *RET*, and *MET* genes. The expansion of genomic testing in lung cancer is a result of the identification of novel biomarkers in clinical trials, and the ability of some NHS laboratories, including AWMGS, to use NGS technology to interrogate multiple gene targets was driven by the Cancer Research UK Stratified Medicine Programme. This programme pioneered the use of NGS for the evaluation of lung tumours within the NHS, with the NGS results from this study resulting in stratification of patients into relevant arms of the Lung Matrix trial. Despite the increasing number of targeted therapies available to lung cancer patients, the 5-year survival for lung cancer remains below 10% (NICE, 2021); therefore, there is a continued drive to improve lung cancer survival rates in the UK and, as such, lung cancer was selected as the focus of this thesis performed within the AWMGS.

The use of anti-PD-1 monoclonal antibodies (mAbs) is a targeted immunotherapy approach used in the treatment of NSCLC patients (Topalian *et al.* 2012). The use of one such immunotherapy, pembrolizumab, is stratified so that only patients whose tumours are shown to express PD-L1 protein are offered this treatment in accordance with National institute for Health and Care Excellence (NICE) guidance (NICE, 2016a). Unfortunately, this method of stratification is imperfect, with a large proportion of patients not responding to treatment (Sul *et al.* 2016). An alternative or complimentary stratification approach is therefore being sought at an international level with the aim to better target the appropriate patient population, so improving patient outcome by avoiding unnecessary side effects in NSCLC patients who will not benefit from this anti-PD-1 mAb and allowing these patients faster access to other treatments that may be of more clinical benefit.

Tumour Mutational Burden (TMB), defined as the number of somatic variants found within a tumour which can be estimated using Whole Genome Sequencing (WGS), Whole Exome Sequencing (WES) or targeted NGS panels (Campesato *et al.* 2015; Johnson *et al.* 2016; Kowanetz *et al.* 2017; Hellmann *et al.* 2018a, Pestinger *et al.* 2020), has emerged as a potential new biomarker for prediction of immunotherapy response in cancer patients (Rizvi *et al.* 2015; Hugo *et al.* 2016; Carbone *et al.* 2017; Hellmann *et al.* 2018b). Despite a lack of definition of the targeted panels suitable for TMB assessment in terms of panel size and gene content, a lack of guidance regarding

how TMB should be calculated in relation to the different types of genetic variants counted within the TMB estimate, and a lack of consensus regarding how 'TMB high' should be defined, in June 2020, the United States Food and Drug Administration (FDA) approved pembrolizumab for the treatment of patients with unresectable or metastatic TMB-high solid tumours, including lung cancers (Marcus *et al.* 2021). This approval was based on the results of the KEYNOTE-158 trial showing improved survival of patients with high TMB tumours when this immunotherapy was used (Marabelle *et al.* 2020). The approval specified that TMB had to be calculated using an FDA-approved targeted NGS panel with a TMB-high value of >10variants/Mb (Marcus *et al.* 2021). Treatment stratification using TMB is yet to be NICE-approved and NHS Genomic services, including AWMGS, do not currently deliver TMB testing. There is therefore a need for TMB to be investigated within an NHS setting to ensure laboratories have the necessary expertise to deliver this testing in preparation for any NICE-approved TMB-based services that may emerge. Prior to the delivery of TMB services within the NHS, the outstanding questions regarding how TMB should be calculated would have to be answered to ensure that appropriate panel and analysis choices were made within NHS Genomics laboratories, ensuring the same high-quality service, providing the same level of clinical utility, was being delivered to patients across the UK by different NHS laboratories. It may not be feasible to dictate to NHS Genomic laboratories a specific NGS panel by which TMB should be calculated owing to laboratories already having established NGS services for the delivery of existing solid tumour genetic testing and the prohibitive cost implications of implementation of another NGS pathway for TMB detection; therefore, assessment of the utility of different panels for TMB estimation is important within the NHS setting.

This research will deliver new insights into the controversies surrounding TMB assessment by investigating the impact on TMB quantification of: targeted NGS panels of varying size (in Mb) and gene content; different TMB calculations based on the inclusion/exclusion of specific variant types; varying TMB high thresholds (including the 10 variants/Mb FDA-approved threshold). This project is novel as, at the time of the study proposal in 2020, there were no publications of such a comparative analysis of TMB assessment within the same patient cohort, and to date this is the first evaluation of TMB technology performed within a diagnostic environment utilising diagnostic-grade samples. There are no comparisons in the literature of the TMB detection capabilities of the targeted panels evaluated within this thesis.

The evaluation of different methods of TMB estimation will identify an optimal set of conditions that provide the greatest utility of TMB as a biomarker within this small cohort. These conditions would require confirmation of utility in research studies involving larger cohorts, but could provide preliminary guidance regarding the panel/analysis choices for TMB service validations within NHS Genomics laboratories in preparation for future NICE-approval of TMB assessment. This information could be used to develop future Association for Clinical Genomic Science (ACGS) best practice guidelines for TMB stratification services within the NHS.

This research will answer the primary research question, establishing whether a combined TMB and PD-L1 biomarker has clinical utility in a Welsh NSCLC patient cohort, encompassing if this combined biomarker predicts pembrolizumab response in a more effective way than PD-L1 expression status alone, in accordance with other published datasets for non-Welsh populations (Carbone *et al.* 2017; Peters *et al.* 2017;

Seiwert *et al.* 2018). Such population-specific utility would be important to demonstrate prior to establishing a TMB service within AWMGS. This study represents the first TMB evaluation of a Welsh cohort, therefore is producing novel data within the field, although, given the small cohort size, this work can be seen as a pilot study from which further research studies could emerge interrogating the utility of TMB assessment within a wider Welsh population. This study provides AWMGS with valuable experience in the evaluation of TMB; this will ensure that, in the event that TMB is NICE approved for use as an immunotherapy response prediction biomarker in the future, Welsh cancer patients will continue to receive the best possible care by AWMGS having the technical knowledge to deliver a TMB service.

The literature review described at the start of this thesis discusses the existing published research regarding the utility of TMB as a biomarker to predict immunotherapy response in NSCLC patients and describes the elements of TMB assessment lacking standardisation, so provides the context to the focus of this thesis. The method chapter primarily focuses on describing the process employed to evaluate the clinical utility of TMB within this Welsh cohort in terms of its effectiveness as a biomarker of immunotherapy response prediction. The feasibility of the implementation of a TMB service within the NHS is also considered by assessing the cost effectiveness of a targeted NGS-based TMB service, and investigating the availability of External Quality Assurance (EQA) schemes for providing quality assurance in TMB assessment, which would be essential to the implementation of a TMB service in an NHS Genomics laboratory requiring ISO15189 accreditation.

The method chapter describes the target enrichment performed using three capture-based NGS gene panels from Illumina, Nonacus and Agilent, and details the bioinformatic data analysis methods used and the TMB calculations performed, including the process of generating optimal ROC curve TMB high thresholds. The statistical analysis performed is described in the methods chapter. The results of the statistical analysis of the data using the paired t-test, Spearman correlation coefficient, Kaplan Meier curves and log rank test, and Kruskal-Wallis test are documented to: compare TMB scores from different panel/analysis combinations; identify associations between immunotherapy response and TMB scores generated from different panel/analysis combinations; describe any statistical significant difference in survival between TMB high + PD-L1 high expressor and TMB low + PD-L1 high expressor patient groups, as well as between PD-L1 high expressor and TMB high + PD-L1 high expressor patient groups.

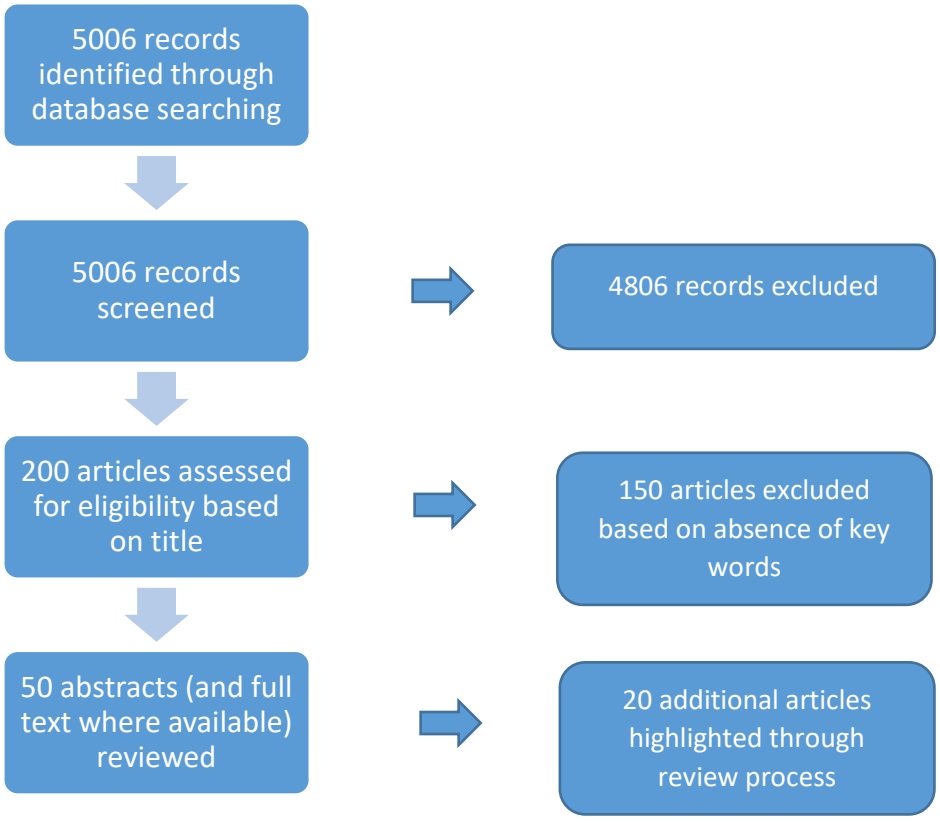
The discussion chapter evaluates the data in relation to the aims of this study, discussing how the TMB panel, calculations, and thresholds impact on the TMB (high/low) status of samples within this Welsh cohort, and discussing the optimal panel/analysis combination within this thesis. The utility of a combined TMB + PD-L1 biomarker for immunotherapy response prediction in this cohort is considered in relation to both the improved sensitivity of patient stratification using this combined biomarker, as well as the feasibility of a TMB service within the NHS. Conclusions surrounding the utility of a combined TMB + PD-L1 biomarker for immunotherapy response prediction in the clinical setting are made including highlighting elements of analysis that could improve the utility of TMB as a biomarker, and making recommendations for TMB quantification relevant to both future research studies and

clinical service validations. Suggestions of future work are highlighted, building on the findings of this pilot study.

1.1 Literature review

A systematic review of available literature was performed, which identified a wealth of both research publications and clinical trial data regarding the use of TMB as a potential immunotherapy treatment stratification biomarker (figure 1). This literature review was submitted in 2020 as part of section C1 of this Professional Doctorate (Roberts 2020) but has been updated to include relevant more recent publications where appropriate. The literature review that follows describes the utility of immunotherapy within cancer care pathways, the benefits of the stratified use of this therapy, the evolvement of TMB as an immunotherapy stratification biomarker in lung cancer, and the variation in TMB assessment methodologies currently in use.

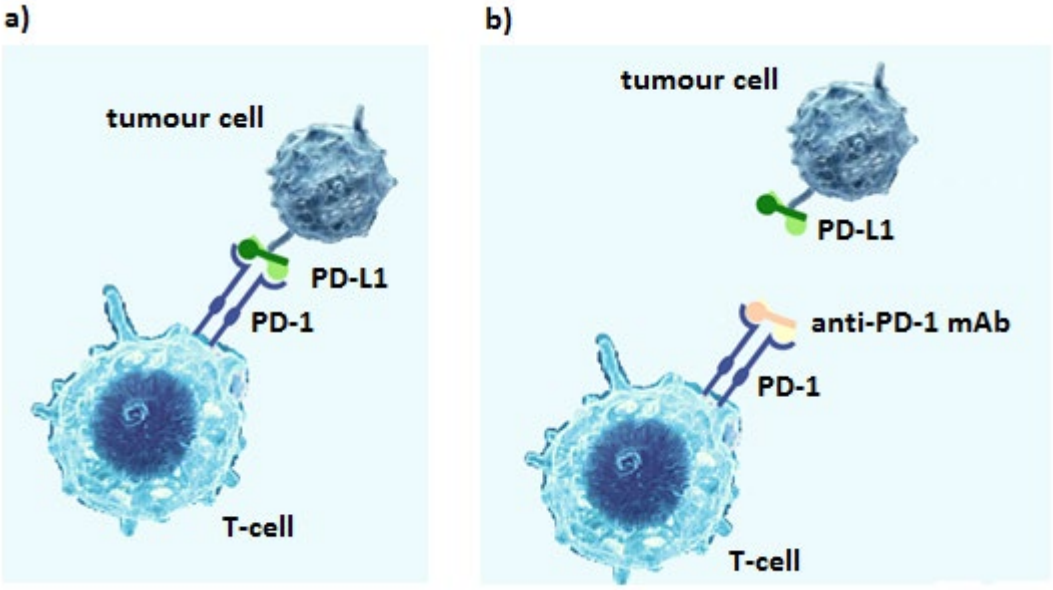
Figure 1: PRISMA diagram summarising the literature search. A search of PubMed (Ncbi.nlm.nih.gov, 2019) using the search term: "tumour mutational burden" identified 5006 results. The first 20 pages of papers listed (10 per page) were assessed based on the title of the paper; this was determined by key words such as "immunotherapy" or "NGS" or "solid tumour" within the article title. A total of 50 articles met the inclusion criteria and the abstracts and/or full texts were reviewed. Review of these 50 articles led to the identification of around 20 further articles of interest.



1.2 Immunotherapy as a treatment option in cancer patient care

Cancer immunotherapies represent a diverse range of treatments developed over the last decade that aim to restore the ability of the immune system to eliminate cancer cells. One such immunotherapy involves the use of anti-PD-1 mAbs for the treatment of melanoma, renal-cell cancer, and NSCLC (Topalian *et al.* 2012; figure 2).

Figure 2: Mode of action of immune checkpoint inhibitors, such as anti-PD-1 mAbs. a) PD-L1 expressing tumour cell binds to PD-1 receptor expressing T-cell and blocks the T-cell initiated immune response; b) Anti-PD-1 mAb binds to PD-1 receptor expressing T-cell, preventing the tumour cell from deactivating immune response.



Immunotherapy can be beneficial over the conventional treatment options of surgery, chemotherapy and radiation by targeting only specific cancer cells, e.g. anti-PD-1 mAbs inhibit only PD-L1 expressing tumour cells, and so potentially causes less toxicity to healthy patient cells. However, immunotherapy has its limitations, as it is not effective for all cancer patients. In a study of 495 patients with advanced NSCLC, Garon *et al.* (2015) determined the objective response rate (ORR) to the PD-1 mAb pembrolizumab to be 19%, with a median duration of response of 12 months. The reason for this variation in cancer patient response to immunotherapies is due to the heterogeneity of the cancer, the variation in the prevalence of immune cells within the tumour microenvironment of different cancers, and the ability of the tumour to evade detection by the immune system (Chiriva-Internati and Bot 2015; Incorvaia *et al.* 2019). The ability to predict who is most likely to respond to cancer immunotherapies has huge benefit to both patients and the NHS as a whole, as costly ineffective treatment and immune-related adverse effects can be avoided via a stratified approach.

1.2.1 Immunotherapy treatment stratification in lung cancer patients

Currently, the only NICE-approved immunotherapy treatment stratification is the determination of PD-L1 expression levels in tumours through the use of IHC to predict response to the anti-PD-1 mAb pembrolizumab in NSCLC patients (NICE, 2016a). The

premise is that, tumour cells expressing high levels of PD-L1 will have the best response to pembrolizumab (Topalian *et al.* 2012; Taube *et al.* 2014). PD-L1 IHC analysis is currently performed in the Cellular Pathology department of University Hospital of Wales, allowing Welsh NSCLC patients to benefit from anti-PD-1 immunotherapy in a stratified manner.

1.2.1.1 PD-L1 expression as a biomarker for immunotherapy treatment stratification

Currently, there are two anti-PD-1 mAbs to treat human cancers that are both NICE- and FDA-approved: pembrolizumab used to treat NSCLCs and melanomas (as mentioned above in relation to treatment stratification in NSCLC), and nivolumab which is primarily used to treat melanomas (NICE, 2016a; NICE, 2016b; Pai-Scherf *et al.* 2017; Fda.gov, 2020a). The mode of action of these anti PD-1 mAbs in improving a cancer patient's ability to detect the tumour and initiate an immune response is shown in figure 2. In contrast to pembrolizumab, the effectiveness of nivolumab as an immunotherapy agent has been deemed to be independent of PD-L1 expression (NICE, 2016b). There are a number of additional anti-PD-1 drugs that are FDA-approved for the treatment of a variety of cancer types, which are both PD-L1 expression dependent (e.g. atezolizumab in NSCLC) and independent (Fda.gov, 2020a).

PD-L1 expression is considered to be sub-optimal as a biomarker for immunotherapy response, as some patients who test positive for PD-L1 expression may not respond to immunotherapy, and conversely, there are PD-L1 expression negative patients whom still respond (Robert *et al.* 2015; Ribas *et al.* 2016). Specifically, 2016 trial data for NSCLC patients treated with pembrolizumab, showed that patients with high PD-L1 expression (>50%) still only had a 45% response rate to this immunotherapy (Sul *et al.* 2016). There are two IHC technical factors that could be affecting this lack of consistency in PD-L1 as a biomarker; firstly, the multiple PD-L1 IHC assays available have different sensitivities for detecting PD-L1 expression (Rimm *et al.* 2017; Hirsch *et al.* 2017), and secondly, the scoring of PD-L1 expression across all assays is variable dependent on the pathologist performing the assessment (Rimm *et al.* 2017). Other factors that impact on the clinical utility of PD-L1 expression as a biomarker are biological factors rather than technical-related nuances as discussed above (table 1).

Table 1: Biological factors contributing to PD-L1 expressing tumours showing a lack of response to anti-PD-1 immunotherapy (amended from Cottrell and Taube 2018).

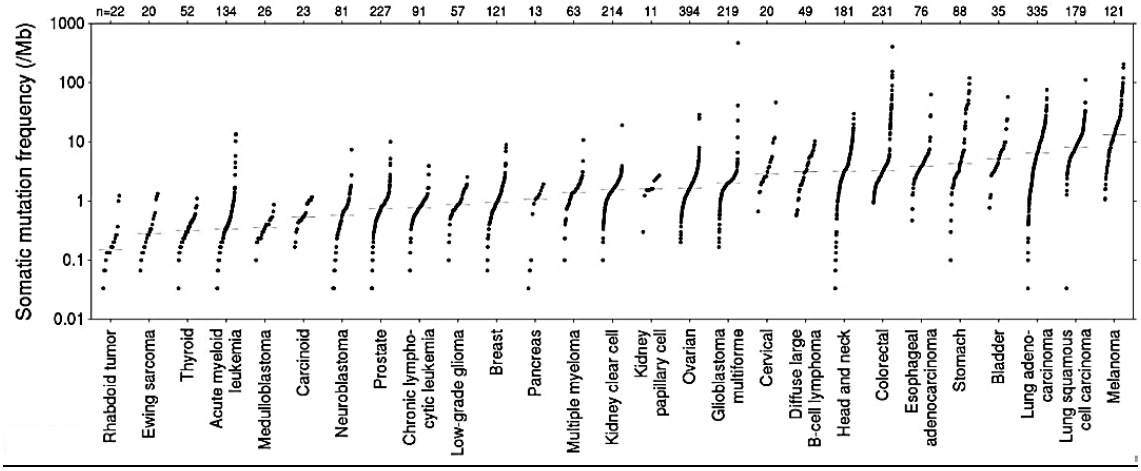
1	The sample analysed may not be representative of the whole tumour owing to heterogeneity within the tumour.
2	PD-L1 expression can change over time, therefore the sample analysed may not be representative of the tumour at the time of immunotherapy treatment.
3	Tumour could express other antigens, asides from PD-L1, which act as inhibitors of the cellular immune response.

1.2.1.2 The emergence of TMB as a potential biomarker for immunotherapy treatment stratification in lung cancer patients

Owing to the imperfect nature of PD-L1 expression as a biomarker, alternative biomarkers, such as TMB, are being investigated as alternatives for immunotherapy treatment stratification. Studies were published in 2015-16, focussing on the reasons

behind variable immunotherapy responses in melanoma and NSCLC patients treated with anti-PD-1 drugs (Rizvi *et al.* 2015; Hugo *et al.* 2016). Work by Lawrence *et al.* (2013) and Alexandrov *et al.* (2013) had already illustrated that the somatic variant frequency in lung cancers and melanomas was high, but was also highly variable between samples of the same cancer type (figure 3). Rizvi *et al.* (2015) hypothesised that the number of somatic variants within a NSCLC tumour may influence the patient's response to immunotherapy.

Figure 3: Variability in somatic variant frequencies across different tumour types. Data is provided for 3083 tumour-normal pairs from 27 different tumour types established using WES (Lawrence et al 2013). Vertical axis represents the total number of somatic variants in the exome; each dot represents a tumour-normal pair. Paediatric tumours, such as rhabdoid tumours and Ewing sarcomas, have low number of somatic mutations (furthest left on X axis); lung cancers and melanomas have a high somatic mutation frequency (furthest right on X axis). The degree of TMB variation within a tumour type is illustrated by the vertical range of each column of dots, with a high degree of TMB variation being reported in lung cancers (0.1-100 variants per Mb) and melanomas in particular.



Rizvi *et al.* (2015) and Hugo *et al.* (2016) utilised WES of NSCLC and melanoma patient samples respectively in order to identify genomic variations within these tumour samples. These studies found that tumours with a higher non-synonymous variant burden were more likely to respond to anti-PD-1 immunotherapy (pembrolizumab or nivolumab) and were linked to improved patient survival in both the melanoma and lung cancer populations studied (Rizvi *et al.* 2015; Hugo *et al.* 2016; figure 4). The findings of Hugo *et al.* (2016) did not meet statistical significance, which may have been linked to the fact that the study did not focus on a single immunotherapy drug, but included patients receiving either pembrolizumab or nivolumab. Despite both drugs being anti-PD-1 mAbs, it is already known that the effectiveness of nivolumab is independent of PD-L1 expression (NICE, 2016b); therefore, perhaps the Hugo *et al.* (2016) data could be skewed by a similar independency between nivolumab effectiveness and TMB status.

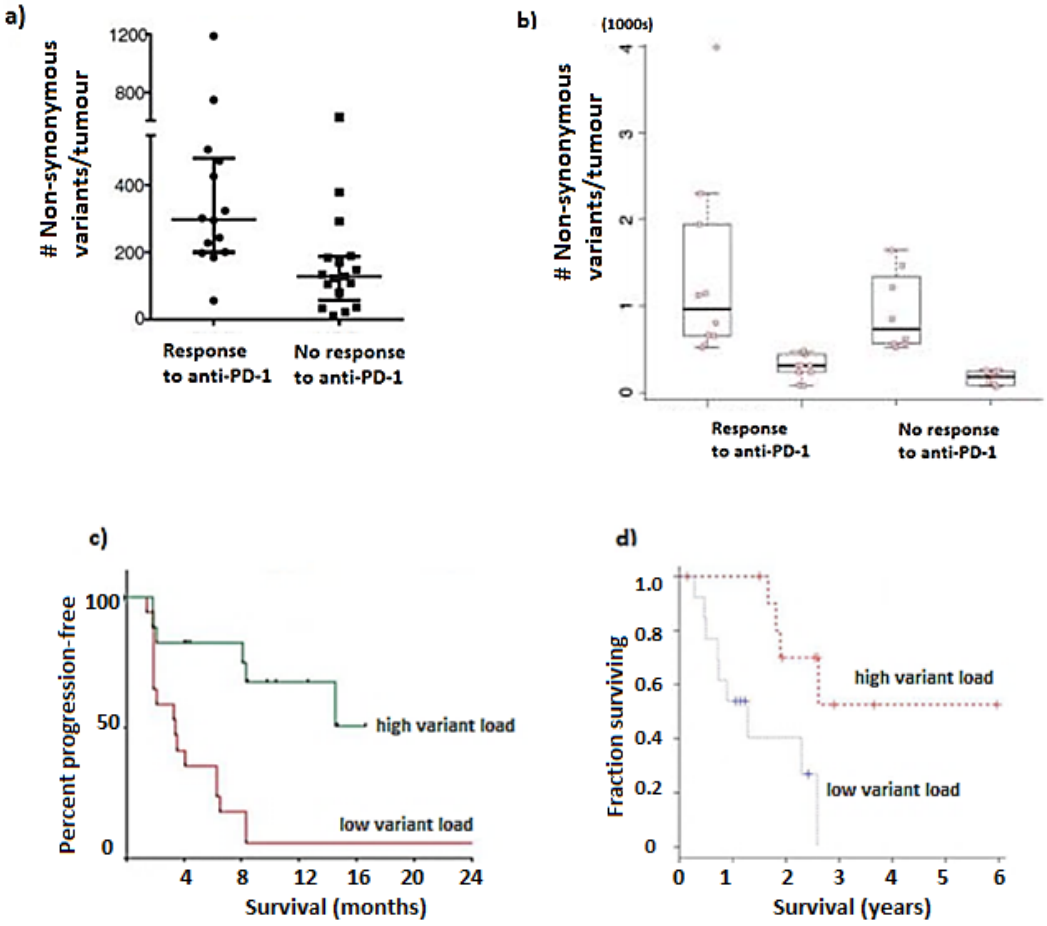
Figure 4: Correlation between variant number and immunotherapy response in published datasets of NSCLC and melanoma patients.

a) Correlation between non-synonymous variant rate and immunotherapy response (adapted from Rizvi *et al.* 2015). Plot showing the number of non-synonymous variants in 31 NSCLC patients with either response (defined as partial or stable response lasting >6 months; 14 patients) or lack of response (17 patients) to anti-PD-1 drug pembrolizumab. Median 299 variants/tumour in the response group versus 127 variants/tumour in the non-response group (Mann-Whitney *P* = 0.0008).

b) Correlation between non-synonymous variant rate and immunotherapy response, with data being divided into low and high variant load groups (determined as below/above the group median) (adapted from Hugo *et al.* 2016). Plot showing the number of non-synonymous variants in 38 melanoma patients with either response (response patterns were based on irRECIST; 21 patients) or lack of response (17 patients) to anti-PD-1 drugs (pembrolizumab or nivolumab). Median 495 variants/tumour in the response group versus 281 variants/tumour in the non-response group (Mann-Whitney *P* = 0.30).

c) Correlation between non-synonymous variant rate and survival in 31 anti-PD-1 (pembrolizumab) treated NSCLC patients (adapted from Rizvi *et al.* 2015).

d) Correlation between non-synonymous variant rate and survival in 38 anti-PD-1 (pembrolizumab or nivolumab) treated melanoma patients (adapted from Hugo *et al.* 2016).



Following the Rizvi *et al.* (2015) study highlighting TMB as a potential biomarker for immunotherapy response in NSCLC patients, WES to assess TMB status was retrospectively used within 2 lung-focused clinical trials, CheckMate-026 (Carbone *et al.* 2017), and CheckMate-032 (Hellmann *et al.* 2018b), involving the anti-PD-1 drug nivolumab. The correlation between immunotherapy response/improved lung cancer patient survival and high TMB as described by Rizvi *et al.* (2015) was replicated in these trials. This nivolumab-based trial data does not support the previous proposal that there is independency between nivolumab effectiveness and TMB status, suggesting that the relatively small dataset of Hugo *et al.* (2016) could have resulted in lack of statistical significance within this study.

Importantly Checkmate-026 highlighted a combined benefit of using both PD-L1 status and TMB levels to predict patient response to immunotherapy, with high TMB + high PD-L1 expression patients having the best outcomes in this trial (Carbone *et al.* 2017). This combined predictive benefit of TMB and PD-L1 data has since been replicated in other studies, both for NSCLC (Castellanos *et al.* 2019), and head and neck cancer (Seiwert *et al.* 2018). This research thesis will identify whether the same combined benefit of PD-L1 expression status and TMB estimation is replicated within a small Welsh NSCLC patient cohort.

Mounting evidence for the clinical utility of TMB as an immunotherapy response biomarker led the CheckMate-227 trial to be the first trial to prospectively evaluate progression free survival (PFS) in high TMB lung cancer patients (Hellmann *et al.* 2018a). Based on preliminary data showing improved PFS in high TMB patients receiving combined immunotherapy (the PD-1 antibody nivolumab, plus the CTLA-4 antibody ipilimunab), in June 2018 the FDA approved the application for the use of this drug combination for treatment of NSCLC patients with high TMB (Bristol-Myers Squibb, 2018). Unfortunately, the overall survival data based on TMB status that emerged later in 2018 was not statistically significant, which led Bristol Myers Squibb to withdraw the FDA licence application for combined immunotherapy usage in high TMB NSCLC patients in January 2019 (Targeted Oncology, 2019).

Despite the lack of statistical significance of the CheckMate-227 overall survival data noted above, research into TMB as a biomarker continued. Most notably, the KEYNOTE-158 trial investigated the utility of TMB assessment for predicting pembrolizumab response in a range of solid tumours, including small cell lung cancer, and demonstrated objective response rates in 29% (30/102) of TMB high patients compared to only 6% (43/688) of TMB non-high patients (Marabelle *et al.* 2020). Based on this trial data, in June 2020, treatment stratification of pembrolizumab based on the use of TMB became the first example of an FDA-approved service using TMB as a biomarker; specifically, this approval was granted for the use of pembrolizumab in patients with unresectable or metastatic TMB-high solid tumours, with a requirement for TMB assessment to be performed using an FDA-approved NGS panel with a TMB high threshold of 10 variants/Mb (Marcus *et al.* 2021).

Interest in TMB as a biomarker remains high in the research environment, as demonstrated by a ClinicalTrials.gov website search performed on 9th June 2022, which identified twenty-two currently active clinical trials evaluating the utility of using TMB status to refine immunotherapy use amongst NSCLC patients (ClinicalTrials.gov, 2022). Future NICE approval of drugs based on TMB level may be forthcoming if ongoing prospective trials yield positive associations between TMB and immunotherapy

response, particularly given the fact that the first FDA approval of TMB as a biomarker has already been granted (Marcus *et al.* 2021). Alongside the clinical trial research that is currently ongoing into the utility of TMB as a biomarker, a 2019 survey by the International Quality Network for Pathology (IQN Path) to assess the global state of TMB testing at this timepoint identified that, despite TMB not being FDA- or NICE-approved as a biomarker at the time of the survey, TMB testing was already being offered by 69 labs worldwide (Fenizia *et al.* 2021). The testing was being performed in these labs for a combination of research purposes (56 labs) and for clinical applications (57 labs) (Fenizia *et al.* 2021).

1.3 TMB as a biomarker in the clinical setting

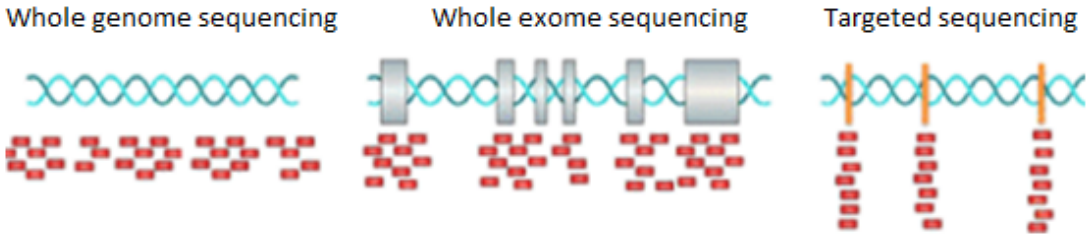
Many of the early research studies and trials noted so far in relation to the measurement of TMB (Ritzi *et al.* 2015, Carbone *et al.* 2017, Hellmann *et al.* 2018b) relied on the use of WES to interrogate tumour DNA and a matched normal DNA from a patient. Both the choice of technology and the dual sample testing have practical implications for genomic laboratories, should these labs be required to deliver TMB-based services for improved NHS patient care; these issues are discussed below.

1.3.1 TMB detection methods: WES vs targeted NGS panels

The gold standard for measuring TMB could be considered to be WES as this technique identifies variants across the whole exome of the tumour sample (approximately 22,000 coding genes over 30Mb) and uses this value to determine variant load. However, the high levels of both human resources and IT support needed to process samples by WES and interrogate and store WES data, serve as potential barriers to WES use within an NHS setting (Phgfoundation.org, 2011; Fancello *et al.* 2019).

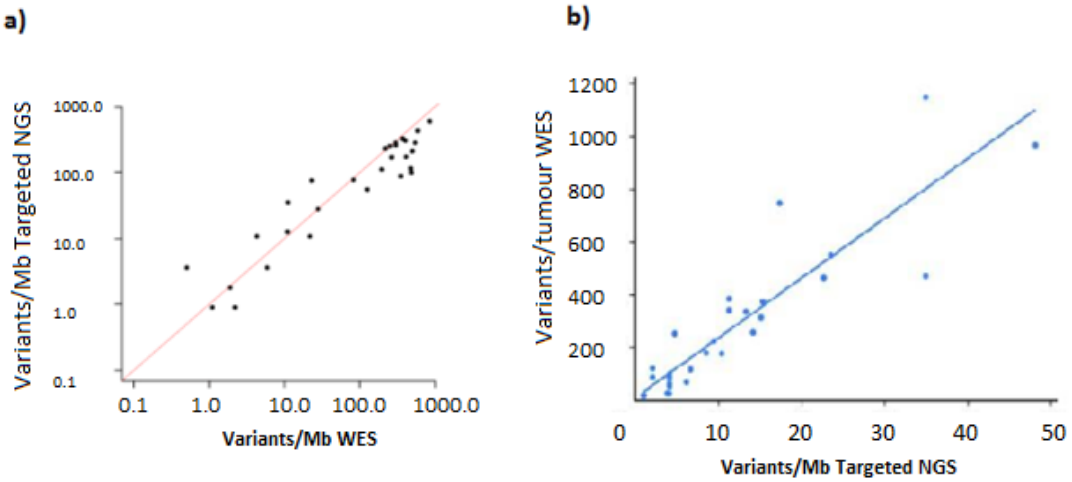
In terms of WES for solid tumour sample analysis, there is an additional practical issue pertaining to the limited amount of tumour material available, coupled with the DNA input requirements of WES. Hellmann *et al.* (2018b) noted that only 61% of lung patients in the CheckMate-032 trial had sufficient tumor biopsy material to generate the 150ng DNA required for WES. To be feasible within the clinical setting therefore, TMB quantification would ideally be performed using a targeted NGS panel where both DNA requirements and staffing and computational resource needs are reduced (figure 5). It is therefore perhaps unsurprising that the results of the IQN Path survey in 2019 showed that 72% (50/69) of the labs already analysing TMB at a research or clinical level were using targeted sequencing approaches (Fenizia *et al.* 2021).

Figure 5: Illustration of target enrichment in WES and targeted NGS panels. Target enrichment is the process by which a subset of genes or regions of the genome are isolated and sequenced. In WES, the target enrichment process selects for all coding genes of the exome, whilst in targeted NGS panels the enrichment focuses on a specific subset of genes according to the panel design. The DNA double helix is shown; the grey boxes represent genes that are targeted in WES; the orange regions represent genes or gene regions that are targeted for enrichment within an NGS panel. Red regions illustrate the multiple overlapping sequencing reads that would be generated from each of these sequencing approaches.



The accuracy of TMB measurement using targeted NGS panels compared with WES has been investigated in a number of studies across a range of tumour types (including lung), with good correlation between TMB levels ascertained by the different methodologies (Rizvi *et al.* 2018; Chalmers *et al.* 2017; figure 6). Accuracy of panel-derived TMB estimation in relation to WGS has also been demonstrated; Pestinger *et al.* (2020) found that the TMB data from the 1.94Mb Illumina TSO500 panel was comparable to data generated from WGS ($R^2 = 0.9$). Importantly, many studies and clinical trials have also successfully used targeted NGS panels to demonstrate the potential clinical utility of TMB as a biomarker for immunotherapy response (Campesato *et al.* 2015; Johnson *et al.* 2016; Kowanetz *et al.* 2017; Hellmann *et al.* 2018a). This evidence supporting the utility of targeted NGS panels in estimating TMB supports the use of such panels in this research thesis to evaluate utility in a Welsh patient cohort.

Figure 6: Correlation between TMB levels in targeted NGS panels and WES. a) 29 cancer patients determined by WES (30Mb) and the FoundationOne targeted NGS panel (1.1Mb). Estimates of TMB correlated well with an R^2 value of 0.74 (Chalmers *et al.* 2017); b) 49 anti-PD-1-treated NSCLC patients determined by WES (30Mb) and 3 targeted NGS panels (ranging from 0.98-1.22Mb in size). Estimates of TMB by these approaches correlated well (Spearman p = 0.86; *P* < 0.001) (Rizvi *et al.* 2018).

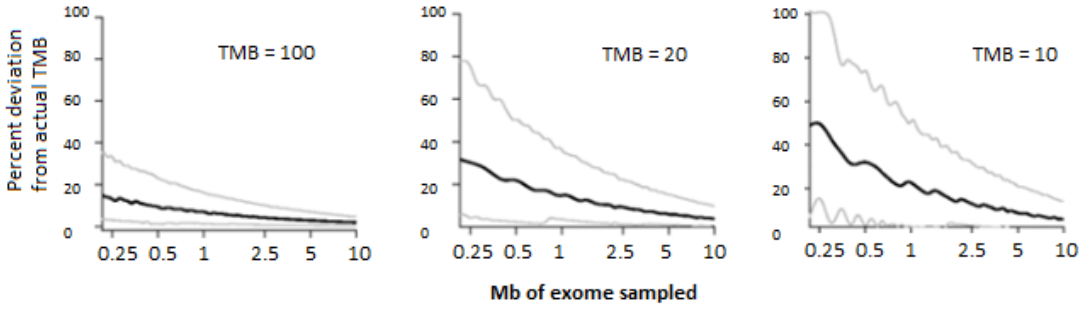


1.3.1.1 Critical size of targeted NGS panels for TMB detection

The sizes of the panels used in the targeted NGS-based TMB studies already mentioned (Rizvi *et al.* 2018, Chalmers *et al.* 2017, Pestinger *et al.* 2020, Campesato *et al.* 2015, Johnson *et al.* 2016, Kowanetz *et al.* 2017, and Hellmann *et al.* 2018a), are in the range 0.91-1.94Mb, representing the interrogation of 3-6% of the exome. Although a number of these panels have been deemed to be effective at TMB measurement by comparison to WES or WGS (Chalmers *et al.* 2017; Rizvi *et al.* 2018; Pestinger *et al.* 2020), a series of published works show that the size of a targeted panel is critical to TMB estimation, and that not all panels are equally effective in their ability to quantify TMB (Chalmers *et al.* 2017; Buchhalter *et al.* 2019; Campesato *et al.* 2015).

Chalmers *et al.* (2017) computationally interrogated the TMB data derived from WES of tumour samples in order to estimate the number of variants that would be identified using NGS panels of size 0.2Mb to 10Mb, at TMB levels of 100, 20, and 10 variants per Mb. They found that the deviation in the number of variants identified by the 'simulation' panels versus WES was lower for high TMB samples than for low TMB samples (figure 7). Meaning that, perhaps unsurprisingly, smaller targeted panels could be expected to have utility in detection of high TMB samples but would be unsuitable for the detection of low level TMB as the panel would likely significantly over/under-estimate TMB levels. This observation was supported in a study by Hatakeyama *et al.* (2018), where TMB values in 2000 cancer patients were shown to be over-estimated for samples with lower TMB levels when a targeted 1.6Mb panel (409 genes) was used compared to using WES.

Figure 7: Estimation of TMB using targeted NGS panels. Results from three computer simulations predicting the performance of NGS panels of various sizes (x axis: 0.2Mb to 10Mb) in the detection of TMB at three different levels (TMB = 100, 20, 10 variants per Mb). Median observed deviation is shown in black and 10% and 90% confidence intervals are shown in grey. These plots show that the standard deviation decreases, and thus accuracy of TMB calling increases, as the size of the NGS panel increases (Chalmers *et al.*, 2017).



From the simulation data, Chalmers *et al.* (2017) concluded that, based on the inaccuracy of TMB estimation compared to WES, panels of <0.5Mb would be unsuitable for TMB detection. Data from another NGS panel vs WES simulation study led Buchhalter *et al.* (2019https://onlinelibrary.wiley.com/doi/full/10.1002/ijc.31878) to conclude that panels had to be >1.5Mb in size to estimate TMB accurately. The work by Hatakeyama *et al.* (2018), however, suggests refinement of this panel lower size limit to >1.6Mb, given that the panel used in the Hatakeyama *et al.* (2018) study was

1.6Mb in size and yet showed some inaccuracy in TMB measurement (compared to WES) for low TMB samples. Confusingly, concluding that a panel size of 1.6Mb is required for accurate TMB detection challenges the data from Chalmers *et al.* (2017) and Rizvi *et al.* (2018), which showed that panels of 0.98-1.22Mb in size produced TMB data that accurately correlated with WES data (figure 6).

The variation in apparent critical NGS panel size in terms of accuracy of TMB estimation compared to WES, could be explained by the findings of Budczies *et al.* (2019) who noted that TMB measurement may depend on multiple factors including the size of the panel, and perhaps the gene content of the panel. Therefore, differences in panel size and panel content are predicted to influence the accuracy of TMB quantification in comparison to WES. The accuracy of an NGS panel for TMB assessment compared to WES would also be influenced by how the TMB calculations were performed which is discussed later in this literature review.

The impact of different gene contents of panels in relation to TMB detection levels has not been specifically investigated in a practical setting, however, Budczies *et al.* (2019) simulated panels of differing gene constitutions to evaluate how effective these panels were at determining variant load. This simulation found that a panel composed of oncogenes and tumour suppressor genes, perhaps unsurprisingly given the causative nature of such genetic variants in cancer, detected far more variants than a panel made up of randomly selected genes (Budczies *et al.* 2019). As the commercially available NGS panels being used to estimate TMB have generally been designed for the identification of clinically relevant variants in tumour samples, their gene content is oncogene/tumour suppressor heavy, which could result in an over-estimation of TMB. Bearing this in mind, the correlation in TMB values generated from targeted panels and WES/WGS as noted previously (Rizvi *et al.* 2018; Chalmers *et al.* 2017; Pestinger *et al.* 2020) suggests that the impact of the gene composition of targeted panels, at least in these studies, is not significant enough to distort TMB estimations compared to the WES/WGS TMB values. Although, interestingly, Chalmers *et al.* (2017) excluded known somatic variants in COSMIC from the TMB calculation, and Pestinger *et al.* (2020) removed driver variants prior to TMB estimation; both of these NGS data filtering steps likely contributed to an improved correlation with the WES/WGS-generated TMB scores by compensating for the gene composition of these targeted panels.

Given the variable permutations of oncogenes and tumour suppressor genes within a panel, it is unsurprising that commercially available NGS panels have highly variable gene contents as noted by Melendez *et al.* (2018), whom highlighted that the two FDA-approved NGS panels, MSK-IMPACT and FoundationOne, only share approximately 50% of genes. The panels investigated in this thesis target a different set of genes and interrogate a different proportion of the genome (table 5), which will allow the impact of panel size and panel gene content on TMB estimation and clinical utility to be evaluated within this research study.

1.3.2 Sample requirements for TMB detection

The early WES-based studies and trials investigating the use of TMB as a biomarker utilised two samples, tumour DNA (typically from an FFPE tissue sample) and a matched normal DNA (e.g. extracted DNA from a patient blood sample or from normal tissue), from each cancer patient to allow genomic data between the samples to be compared and germline (hereditary) genetic variations to be excluded (Rizvi *et al.*

2015; Hugo *et al.* 2016; Cristescu *et al.* 2017; Kowanetz *et al.* 2017). In this way, only somatic (tumour-specific) variations were evaluated within the remit of determining the TMB status of each tumour sample.

A TMB-based immunotherapy stratification service within an NHS setting would preferably involve the analysis of a single tumour sample per patient to keep costs as low as possible, and to align with existing cancer care pathways in which matched normal specimens are not routinely obtained. Fortunately, matched normal DNA is not essential in order to eliminate germline variants from sequencing data. Data from the tumour can be interrogated bioinformatically to remove germline variants via comparison to reference databases available online, which have been generated via the genetic analysis of thousands of individuals in datasets such as the 1000 Genomes Project (Lek *et al.* 2016; Fu *et al.* 2012; The 1000 Genomes Project Consortium 2015); the use of such datasets is commonplace in Genomics laboratories, including AWMGS, in the evaluation of variants identified by sequencing of tumour specimens, so would be an obvious choice within a TMB clinical service. An alternative to germline datasets is the use of computational algorithms to predict and subsequently remove the germline variants within a tumour sample (Sun *et al.* 2018; Chalmers *et al.* 2017; Chan *et al.* 2019; table 2). For example, Sun *et al.* (2018) developed an algorithm based on evaluation of allele frequencies at >3500 single nucleotide polymorphism (SNP) sites; in this algorithm, the allele frequencies at these SNP sites within a tumour sample are compared to the allele frequencies expected if the variants were of germline origin, with similar allele frequencies signifying variants of likely germline origin. This algorithm was validated by Sun *et al.* (2018) as correctly predicting the germline/somatic origin of >95% of variants across 30 tumour samples.

Table 2: Comparison of a number of technical parameters of WES (Ritzi *et al.* 2015; Hellmann *et al.* 2018), and targeted NGS panels (Hellmann *et al.* 2018; Pestinger *et al.* 2019) for TMB estimation (table adapted from Chan *et al.* 2019). Of particular note is the varying methods of data manipulation used in order to remove germline variants from the dataset, and the varying definitions of TMB based on the technical capabilities of each panel.

	WES – Rizvi *et al.* (2015)	**WES data from CheckMate 032 trial (Hellmann *et al.* 2018b)**	**FoundationOne CDx targeted panel data - CheckMate 227 trial (Hellmann *et al.* 2018a)**	**Illumina TSO500 targeted panel (Pestinger *et al.* 2020)**
No of genes interrogated	22,000 gene coding regions	22,000 gene coding regions	324 cancer-related genes	523 cancer-related genes
Types of variants captured	Missense variants	Missense variants and indels	Missense variants and indels	Missense variants and indels
Method of removal of germline variants	Tumour and blood samples sequenced	Tumour and blood samples sequenced	Bioinformatics algorithms	Use of germline variant database
Capture region (tumour DNA)	30Mb	30Mb	1.8Mb	1.94Mb
TMB definition	Number of non-synonymous somatic missense variants in the sequenced tumour genome	Number of somatic missense variants in the sequenced tumour genome	Number of somatic, coding variants (synonymous and non-synonymous), short indels per Mb of tumour genome	Number of somatic, coding variants (synonymous and non-synonymous), short indels per Mb of tumour genome

1.3.2.1 TMB detection using circulating tumour DNA (ctDNA) samples

The use of cancer patient blood samples as an alternative to tissue samples for TMB evaluation has been explored in a number of studies, including studies focussed on lung cancer patients (Gandara *et al.* 2018). These blood-based studies rely on the analysis of ctDNA, which is cell-free DNA that is shed from the patient's tumour cells as a result of apoptosis or necrosis and released into the circulation.

Blood samples have the benefit of being relatively easy to obtain and eliminate the need for invasive biopsies. ctDNA analysis is already used in the NHS diagnostic environment, including within AWMGS, to aid the treatment stratification of NSCLC patients via detection of clinically relevant variants within the ctDNA sample for patients in whom there is insufficient biopsy material available to use in genomic analysis. The utility of ctDNA analysis in NSCLC patients in particular stems from the fact that lung biopsy material can be scarce owing to the limited volume of tumour tissue and the multitude of diagnostic tests required within the lung cancer pathway. Indeed, Lim *et al.* (2015) identified that up to 30% of NSCLC patients have insufficient tumour material to perform the range of tests required to aid diagnosis and treatment decisions. Another benefit of blood samples over tumour tissue is that the ctDNA is representative of the whole tumour, whilst tumour biopsies represent only a specific

area of the tumour, meaning that any potential heterogeneity within the sample could be missed by analysing the tumour biopsy.

The benefits of blood analysis could make ctDNA an attractive option for use in TMB estimation, however, this area of research is still in its infancy and is outside the scope of this thesis. One area for consideration prior to establishment of a ctDNA-based TMB service relates to the selection of suitable methodologies for ctDNA analysis. Methodology selection would have to consider that the ctDNA yield from a blood sample of a lung cancer patient is typically low at 60ng (based on AWMGS experience using Promega Maxwell 16 extraction), and ctDNA fragment lengths are small at around 134bp (Underhill 2021). Therefore, an NGS panel with the ability to amplify small DNA fragments using limited DNA would have to be selected for a ctDNA-based service. The commercially available NGS panels used routinely in current NHS service delivery at AWMGS are designed to amplify larger DNA fragments and have higher DNA input requirements so would have limited utility in ctDNA analysis.

Another important factor to consider prior to using blood samples to measure TMB relates to the observation made by Davis *et al.* (2017) who concluded that there was a low correlation between the TMB values generated from paired tissue and blood samples (mainly from lung and breast cancer patients). Correlation is dependent on when the tissue sample was taken compared to the blood sample, as any differences in timing of the sample collections could impact on the TMB scores obtained. ctDNA levels in cancer patients can vary based on the rate of tumour shedding of ctDNA into the bloodstream, which is affected by the size and location of the tumour. Quantity of ctDNA is also known to be related to more advanced disease (Bettegowda *et al.* 2014). Therefore, TMB estimates from blood analysis could be lower than tumour analysis-based estimations owing to limited ctDNA shedding, whilst tumour-based TMB scores could be over- or under- estimations based on the tumour sample being a skewed representation of the whole tumour. Despite potential differences between ctDNA and tissue based TMB estimations as noted by Davis *et al.* (2017), the B-FIRST trial did show that TMB values derived from ctDNA samples from NSCLC patients could be used to accurately identify patients who benefitted from the anti-PD-L1 drug atexolizumab in terms of an improved overall response rate (Gandara *et al.* 2018); therefore, supporting a potential utility of the use of ctDNA for TMB-based immunotherapy stratification. It is worth noting that as well as the ctDNA-specific issue of appropriate technology selection highlighted above, the issues addressed in this literature review regarding how to calculate TMB are also relevant to a ctDNA-based TMB service and would need resolving prior to the delivery of any clinical TMB service.

1.3.3 Calculating TMB levels

TMB levels were initially reported in WES studies as the number of variants in the tumour genome (Rizvi *et al.*, 2015) although this was more commonly referred to as 'variants per megabase (Mb)' following the use of targeted panels (Hellmann *et al.* 2018a) (table 2). For both WES and targeted panels, this 'variants per Mb' TMB value can be obtained by a simple calculation based on the number of somatic variants identified in X Mb of sequencing (X being the amount of the genome interrogated). Although this seems quite a straight-forward strategy, and implies that TMB levels could be easily compared across platforms, this is not the case in practice owing to

differences in the types of variants included in TMB estimations (table 2), which is discussed below.

1.3.3.1 Variant types included in TMB estimates

One of the earliest TMB-focused WES studies in NSCLC by Rizvi *et al.* (2015) defined TMB as the number of non-synonymous coding variants within the tumour; the study did not include indels in the TMB estimate, and also did not include variants that were intronic, non-coding, or synonymous in the TMB calculation. Zeng and Bromberg (2019) estimated that there are 10,000 synonymous variants in every human genome (3200Mb in size), which equates to 0.0003% of the genome harbouring a synonymous variant. The premise behind excluding synonymous variants from TMB estimates is that such variants are deemed unlikely to be directly involved in creating neoantigens on the tumour cell surface, which is of primary interest in the prediction of immunotherapy response (Melendez *et al.* 2018). Interestingly, the first prospective trial into the link between TMB and progression free survival, CheckMate-227, did not remove synonymous variations from its TMB calculations (Hellmann *et al.* 2018a; table 2). The reasoning behind their inclusion was that the trial deemed the presence of synonymous variants to be an indicator of mutation rate in the tumour genome, therefore considered this to be a worthwhile addition (Hellmann *et al.* 2018a). This thesis has been designed to investigate the utility of including synonymous variants in TMB estimations as this is one variable that will be altered in the TMB calculations performed within the patient cohort investigated.

Targeted NGS panels allow an increased depth of sequencing compared to WES owing to the vastly reduced genomic area interrogated (around 2Mb compared to 30Mb in WES), which allows improved detection of indels. As such, many studies utilising targeted panels include indels in their TMB estimations (table 2). Budczies *et al.* (2019) suggested that TMB estimates could actually be improved by the inclusion of indels in variant load estimates.

There is no formal guidance relating to the variant types to include in a TMB estimate, which would complicate the use of TMB as a biomarker in the clinical setting, and is one of the drivers for this research study in terms of providing data that could generate such best practice guidance. As well as the technical practicalities in determining TMB levels, there are a number of biological factors influencing TMB levels, which could potentially complicate the use of TMB as a biomarker (table 3); these elements are outside the scope of this thesis.

Table 3: Biological and social factors influencing TMB estimations in NSCLC patients.

1	**Sex of patient:** TMB is higher in men (Alexandrov *et al.* 2016).
2	**Smoker status:** TMB is higher in smokers (Alexandrov *et al.* 2016; Sha *et al.* 2020).
3	**Stage of cancer progression:** TMB levels have been shown to be lower in early stage lung cancer patients compared to patients with late stage disease (Zhang *et al.* 2018).
4	**Presence of specific variants in the tumour can be associated with low TMB:** The presence of *EGFR*, *ALK*, *ROS1*, or *MET* exon 14 variants in NSCLCs generally correlates with low TMB levels (Spigel *et al.* 2016).
5	**Presence of specific variants in the tumour are linked to immunotherapy response:** Some gene variants are associated with a better immune response (Alexandrov *et al.* 2013); therefore, the presence of such a variant could mean that a patient could respond well to immunotherapy, irrespective of their TMB status (Chan *et al.* 2019). Conversely, variants in other genes have been linked to insensitivity to immunotherapies in some patient groups (Zaretsky *et al.* 2016).

1.3.3.2 Defining high TMB

TMB is a continuous variable. One of the major controversies within the area of TMB quantification is the lack of a defined threshold for 'TMB high' (table 4), which would be critical within a TMB clinical service as would separate the patients whom would receive immunotherapy treatment from those that would not. As previously illustrated (figure 3), somatic variant frequencies, and therefore by extrapolation TMB levels, vary enormously between tumour types, therefore there is likely a requirement for TMB thresholds to be different dependent on the sample type, as noted by Chan *et al.* (2019) and Strickler *et al.* (2021). Table 4 shows that for a selection of the lung-focused targeted NGS publications referenced within this literature review, the TMB high threshold was set at between 7 and 20 variants/Mb; this is in line with a similar critique of TMB threshold variation performed by Heeke and Hofman (2018).

Table 4: Comparison of the TMB high thresholds across a number of lung-based studies* using either WES or targeted NGS panels for TMB estimation. TMB high thresholds for WES have been converted to variants/Mb using a WES panel size of 30Mb to aid comparison to the targeted panel thresholds. *Note: Chalmers *et al.* (2017) and Pestinger *et al.* (2020) were studies interrogating multiple tumour types including lung.

Reference	TMB quantification methodology	TMB high threshold
Ritzi *et al.* (2015)	WES	>209 variants (equivalent to 7 variants/Mb)
Carbone *et al.* (2017): phase 3 trial - CheckMate-026 data	WES	>243 variants (equivalent to 8 variants/Mb)
Hellmann *et al.* (2018b): phase 1/2 trial - CheckMate-032 data)	WES	>248 variants (equivalent to 8 variants/Mb)
TARGETED NGS PANELS		
Campesato *et al.* (2015)	Foundation Medicine Comprehensive Genome Profiling (CGP) panel	≥7 variants/Mb
Campesato *et al.* (2015)	HSL-CGP (bespoke designed panel)	≥13 variants/Mb
Chalmers *et al.* (2017)	FoundationOne panel (1.1Mb)	>20 variants/Mb
Hellmann *et al.* (2018a) (phase 3 trial: CheckMate-227 data)	FoundationOne CDx (1.8Mb)	≥10 variants/Mb
Pestinger *et al.* (2020)	Illumina TSO500 panel (1.94Mb)	>10 variants/Mb

In the early WES study by Rizvi *et al.* (2015) and the targeted panel analysis by Campesato *et al.* (2015), the TMB high threshold was simply set as the median value within the patient cohort. For example, in the targeted panel analysis performed by Campesato *et al.* (2015) using the Foundation Medicine CGP panel, the median number of non-synonymous somatic variants/Mb was 9 and 5 for tumours from patients with immunotherapy response and no-response respectively, and the TMB high threshold was set mid-way between these values at ≥7 variants/Mb. In a similar manner, for the retrospective exploratory analyses performed on the CheckMate-026 and -032 WES data, TMB high thresholds were defined as the upper tertile of the patient cohort data (Carbone *et al.* 2017; Hellmann *et al.* 2018b). Heeke and Hofman (2018) noted that there is no biological justification for such percentile-based threshold setting as used in these studies.

The first prospective trial evaluating the use of TMB as a biomarker in lung cancer (CheckMate-227) used a TMB high threshold of ≥10 variants/Mb (Hellmann *et al.* 2018a). This threshold was set based on the findings from the CheckMate-568 trial (Ramalingam *et al.* 2018), which utilised the same FoundationOne CDx targeted NGS panel and evaluated patients receiving the same combined immunotherapy as the CheckMate-227 study. In the CheckMate-568 trial, the TMB high threshold of ≥10 variants/Mb was statistically derived using a ROC curve (Ramalingam *et al.* 2018). Using this threshold, Ramalingam *et al.* (2018) showed that ORR of NSCLC patients receiving nivolumab and ipiliumumab increased as TMB increased (TMB ≥10 variants/Mb, n=48, 43.8% ORR; TMB <10 variants/Mb, n=50, 12.0% ORR), and plateaued after 10 variants/Mb. A threshold of 10 variants/Mb also defined high TMB

in the KEYNOTE-158 study (Marabelle *et al.* 2020), which also used the FoundationOne CDx targeted NGS panel, although importantly evaluated patient response to pembrolizumab rather than the combination of nivolumab and ipiliumumab assessed in the CheckMate-568 (Ramalingam *et al.* 2018) and CheckMate-227 (Hellmann *et al.* 2018a) trials. This 10 variants/Mb threshold demonstrated utility in the KEYNOTE-158 study for prediction of patient response (Marabelle *et al.* 2020), suggesting that a validated TMB threshold for a particular tumour type and NGS panel combination could have utility independent of the immunotherapy/immunotherapy combination used.

Based on the results of the KEYNOTE-158 trial, FDA approval for the use of TMB assessment across a range of solid tumours to guide pembrolizumab stratification has been granted using a TMB high threshold of 10 variants/Mb (Marcus *et al.* 2021). Interestingly, the approval is based on the use of FDA-approved NGS panels, which include the FoundationOne CDx panel but also include the MSK-IMPACT panel, and yet the same TMB high threshold is recommended (Marcus *et al.* 2021). Importantly, based on the CheckMate-227 and -568 trial data, Pestinger *et al.* (2020) used a TMB high threshold of ≥10 variants/Mb when comparing the utility of the Illumina TSO500 panel with WGS for TMB assessment (table 4). All colorectal patients assessed in the Pestinger *et al.* (2020) study were classified as high TMB in a concordant manner using both the targeted panel and WGS. This therefore suggests that the TMB threshold of ≥10 variants/Mb does have utility in determining TMB levels using other NGS panels, asides from the FoundationOne CDx NGS assay (Pestinger *et al.* 2020), as well as having utility in both NSCLC (Ramalingam *et al.* 2018; Hellmann *et al.* 2018a) and colorectal (Pestinger *et al.* 2020) patient cohorts. This contradicts the suggestion of Chan *et al.* (2019) and Strickler *et al.* (2021) that different TMB thresholds would likely be required for different tumour types. The utility of the same TMB high threshold in both lung and colorectal tumour types aligns to the similarities in the somatic variant frequencies between these two cancer types (Lawrence *et al.* 2013; illustrated in figure 3), and suggests that this threshold would also have utility in the evaluation of TMB in melanomas with a similar somatic variant frequency to lung and colorectal cancers. Within the KEYNOTE-158 study, the 10 variants/Mb threshold showed utility in immunotherapy response prediction across a range of tumour types including small cell lung cancer, cervical cancer and thyroid (Marabelle *et al.* 2020). However, based on the Lawrence *et al.* (2013) WES-generated data illustrated in figure 3, two of these tumour types alone have differing somatic variant frequencies, with cervical cancer having around 1-10 variants/Mb and thyroid cancers generally having 0.1-1variants/Mb. It is therefore interesting that a single 10 variants/Mb threshold has shown utility across these two tumour types (cervical and thyroid), and indeed across the other seven tumour types evaluated by Marabelle *et al.* (2020).

1.4 Summary of literature review

The literature review provides a summary of the issues surrounding the use of PD-L1 expression status as an immunotherapy response prediction biomarker, and highlights published evidence that demonstrates the utility of TMB, both alone or in combination with PD-L1 high expression status, as a potential biomarker for immunotherapy treatment stratification in lung cancer patients. The literature review describes the areas of TMB estimation that lack consensus in the research environment, namely: the lack of consistency in the variants included within TMB calculations, the absence of a validated TMB high threshold, and the lack of clarity regarding which targeted NGS panels provide accurate TMB estimation (in terms of comparison to the gold standard WES/WGS approach), with the suggestion that panel utility is influenced by the amount of the genome interrogated as well as the gene content of the targeted genomic region.

1.5 Research hypothesis

1.5.1 Primary research question

Does TMB in combination with PD-L1 expression analysis have clinical utility as a biomarker for anti-PD-L1 immunotherapy treatment response in a Welsh lung cancer patient cohort?

1.5.2 Hypothesis

Data from this study will emulate the findings of international research groups (Carbone *et al.* 2017; Peters *et al.* 2017; Seiwert *et al.* 2018) demonstrating that TMB status in combination with PD-L1 expression data can act as an anti-PD-L1 immunotherapy treatment response biomarker, by accurately stratifying patients in this Welsh lung cancer patient cohort into responder and non-responder groups.

1.6 Research thesis aims and objectives

As the literature review reveals, the potential utility of TMB as a biomarker of immunotherapy response has been noted within at least six clinical trials (Kowantz *et al.* 2017; Carbone *et al.* 2017; Hellmann *et al.* 2018b; Hellmann *et al.* 2018a; Marabelle *et al.* 2020). The primary objective of this research project is to assess whether such TMB clinical utility extends across a small Welsh lung cancer patient cohort. The primary focus will be on demonstrating if TMB assessment in combination with PD-L1 expression analysis can accurately stratify patients into immunotherapy responder and non-responder groups. However, other aspects of clinical utility will be considered within this thesis relating to the feasibility of TMB service implementation within the NHS environment. The assessment of feasibility, including cost of testing and availability of External Quality Assurance (EQA) schemes, is essential knowledge to understand the true potential of or barriers to the establishment of a TMB clinical service within the NHS.

The small cohort size (n=17) will limit the weighting of the findings in terms of there being limitations in the translation of any demonstrated clinical utility of TMB assessment within this cohort to the Welsh population as a whole. This thesis can therefore be viewed as a pilot study, the findings from which can be used as a platform to perform future larger-scale research regarding the utility of TMB as a biomarker within the Welsh population.

As pembrolizumab is only used currently in a stratified manner (based on PD-L1 status) within the treatment of NSCLC patients within the NHS, all patients in this Welsh cohort will have been previously identified as having PD-L1 high expressing tumours. It is therefore the combined utility of both TMB status and PD-L1 expression levels that is being assessed within this thesis rather than the utility of TMB alone as a biomarker of immunotherapy response. The literature review highlighted that the combined benefit of TMB level and PD-L1 status in predicting immunotherapy response has been noted previously in three studies (Carbone *et al.* 2017; Peters *et al.* 2017; Seiwert *et al.* 2018). In assessing the utility of TMB as a biomarker, this thesis will use different targeted NGS panels, different TMB calculations, and different TMB thresholds to address some of the controversies within the area of TMB assessment, which adds to the secondary aims of this research thesis.

1.6.1 Aims of this thesis

1. Produce novel data regarding the impact of different targeted NGS panels of different size and gene content on TMB estimation. This will be done by directly comparing the TMB quantifications of lung tumour samples from the same cohort of patients using three different NGS panels.
2. For each of the NGS panels, perform TMB quantifications of the tumour samples using a range of TMB calculations based on altering the variants counted within the estimations, including investigating the impact of the inclusion/exclusion of synonymous variants on TMB score.
3. Evaluate the utility of TMB quantification, calculated from each permutation of NGS panel and TMB calculation method, for immunotherapy response prediction, using different TMB high thresholds to define the TMB high patient group, thus identifying an optimal set of panel/analysis/threshold parameters that maximises the utility of TMB assessment in this cohort.
4. Evaluate the clinical utility of TMB assessment to guide immunotherapy stratification in terms of the feasibility of TMB service implementation within the NHS.

1.7 Justification of the methodology

1.7.1 Sample numbers

The patient cohort size within this thesis is dictated by the cost of NGS and the funding available for this research; the project costings are noted in appendix 2. Ideally, all samples will be processed on each of the three NGS panels but, as well as costings, another factor that will impact on this is the volume of DNA obtained from each sample, which was noted as a project risk within the innovation proposal (appendix 2).

The NGS panels used require a minimum amount of DNA ranging from 25ng-100ng; limited DNA may mean that samples can only be analysed on one or two of the panels.

The study design relies upon a minimum number of eight samples being analysed on each of the three NGS panels in order to allow meaningful statistical analysis. Eight represents the minimum number of observations required to perform the Spearman rank correlation test, which is the analysis that will be performed to describe the association between TMB value and immunotherapy response. The analysis of a minimum of 12 samples on more than one NGS panel is required to ensure the statistical significance of the paired t-test evaluation of any differences in TMB values obtained from the different panels. Importantly, the expected difference between TMB levels measured by different panels is unknown so power calculation is not possible within the scope of this project.

1.7.2 Selection of targeted NGS panels

This research will perform TMB quantification of FFPE samples from lung cancer patients using three commercially available targeted NGS panels, designed by Agilent, Illumina and Nonacus, which is a novel evaluation not identified in the literature review. The panels selected are all over the 1.5Mb size noted by Buchhalter *et al.* (2019https://onlinelibrary.wiley.com/doi/full/10.1002/ijc.31878) as being essential for accuracy of TMB measurement, but the Nonacus panel is below the 1.6Mb size accuracy threshold suggested by the findings of Hatakeyama *et al.* (2018) (table 5). The results of this research will therefore add to the debate regarding impact of panel size on the accuracy of TMB measurement (Bucchalter *et al.* 2018).

Adding to the simulated work of Budczies *et al.* 2019, the use of three panels will allow the impact of gene panel content on TMB estimation to be evaluated. Although the panels selected have a large oncogene/tumour suppression gene content (table 6), the gene content differs (table 5; full gene list of each of the three panels can be found in appendix 5). The differing gene panel contents supports the point raised in the literature review where it was noted that panels used for TMB detection can have hugely variable gene contents (Melendez *et al.* 2018).

Table 5: Comparison of size and gene content of the three NGS panels evaluated in this thesis. Panels are: Illumina TruSight™ Oncology 500 panel, Agilent SureSelect Community Design Glasgow Cancer Core panel, and Nonacus Cell3™ Target: Pan Cancer panel.

	Agilent	Illumina	Nonacus
Panel size	1.7 Mb	1.94 Mb	1.58 Mb
Number of genes covered	174	523	524
% of genes shared with Agilent panel	na	32%	29%
% of genes shared with Illumina panel	95%	na	68%
% of genes shared with Nonacus panel	87%	68%	Na

Table 6: Gene contents of the Illumina, Nonacus and Agilent panels in relation to the most commonly mutated genes in lung cancer (mycancergenome.org, 2022a). Green/red indicates that the gene is/is not targeted by the panel. Oncogenic driver mutations in lung cancer most commonly occur in the *KRAS* and *EGFR* genes (Chevallier *et al.* 2021); the variant hotspots (representing locations where variants are most likely to occur) of both genes (*EGFR* exons 18-21 and *KRAS* exons 2-4) are targeted by all three panels.

	TP53	*KRAS*	*EGFR*	*CDKN2A*	*STK11*	*KEAP1*	*KMT2D*	*RB1*	*ATM*	*PIK3CA*	*NF1*	*SMARCA4*	*CDKN2B*	*ARID1A*	*RBM10*	*ALK*
Illumina																
Nonacus																
Agilent																

As the samples from the Welsh Cancer Bank (WCB) are of unknown TMB status, the accuracy of TMB assessment will be evaluated in this study by using the data generated from the Illumina panel as a benchmark with which to assess the accuracy of the other panels in TMB quantification. This bench-marking is possible as the Illumina TSO500 panel was shown by Pestinger *et al.* (2020) to provide TMB values comparable to WGS, which, like WES, could be considered as a 'gold standard' approach to TMB scoring given that the presence of variants across the whole genome is evaluated.

As well as panel size, panel gene content, and demonstrated WGS correlation being factors in panel selection in this study, NGS panel choice was also influenced by the experience of AWMGS in working with these panels. Both Illumina and Agilent panels had been used in the laboratory before, albeit with some minor protocol differences, therefore staff were familiar with the methodologies and were already appropriately trained. This will therefore provide a level of quality assurance to the results obtained.

All three NGS panels are hybridisation capture-based target enrichments. Hybridisation capture relies on the use of probes to capture target sequences in a DNA library, and uses low input amounts of DNA, which is often a limitation of working with FFPE material and is relevant in the context of this project in terms of attempting to maximise the number of samples that could be sequenced on all three NGS panels to maximise the data obtained and to ensure statistical analysis of the data is possible.

Another consideration in panel selection was ensuring the panels selected had utility within the existing lung cancer care pathway, for which tests are already funded by the Welsh Health Specialised Services Committee (WHSCC) and by NHS England in English Genomics laboratories. A panel that contains all of the lung clinically relevant genes associated with NICE-approved NSCLC treatment stratification (NICE, 2022) as well as the genes on the Cancer Test Directory (england.nhs.uk, 2022), which describes the genomic tests commissioned by NHS England for cancer patients, would have greater utility in an NHS laboratory, beyond the scope of TMB detection (table 7).

Table 7: Comparison of Illumina, Nonacus and Agilent panel gene contents to existing NSCLC service requirements. The Cancer Test Directory (england.nhs.uk, 2022) gene list (as of 31 October 2022) is the same as the NICE-approved treatment-related genes (NICE, 2022) with the addition of the *MET* gene. Green/red indicates that the gene is/is not targeted by the panel.

	EGFR	*KRAS*	*BRAF*	*ALK*	*ROS1*	*RET*	*NTRK1*	*NTRK2*	*NTRK3*	*MET*
Illumina										
Nonacus										
Agilent										

1.7.3 Evaluation of immunotherapy response

Immunotherapy responses for each of the patients in this Welsh cohort (n=17) will be obtained through medical record review, where response is recorded using common terminology defined by RECIST 1.1 guidance (Eisenhauer *et al.* 2009): complete response (CR), partial response (PR), progressive disease (PD), stable disease (SD), incomplete response (IR). To perform statistical analysis of the data, a quantifiable measurement of response will be required, for which number of days survival post-PDL1 test date will be used. PD-L1 IHC analysis is performed at the time of lung cancer diagnosis therefore effectively represents the date of diagnosis within the patients in this study. Patients in the cohort will be selected to ensure that PD-L1 IHC assessment occurred in 2017. Therefore, in line with the <10% 5-year survival rate of lung cancer patients (NICE, 2021), by the completion of this thesis in 2022 there will be clear survival data for each patient as sadly the majority of the cohort will be deceased. Patient survival will therefore be calculated from PD-L1 test date to the date of death or, for the small number of expected surviving patients, the date of death will be replaced by the date of final medical record review in 2022.

1.7.4 TMB measurement

Some of the major controversies of TMB as a biomarker come from the fact that there is no standardised method for TMB measurement and there is no agreed threshold to identify immunotherapy responders (Fancello *et al.* 2019). Within this research thesis, dual analysis of the NGS data will be performed to include/exclude the presence of synonymous variants as this is a key area of difference within TMB-based publications (Rizvi *et al.* 2015; Hellmann *et al.* 2018a). The inclusion of synonymous variants within TMB estimations was a recommendation of Fancello *et al.* (2019) who proposed that this provided a better approximation of TMB when extrapolated across the whole genome. There is no published evidence against TMB estimations including indels, therefore these will be included in the TMB calculations within this thesis.

As well as assessing the impact of synonymous variants on TMB quantification, investigation into the impact of sequencing artefacts on TMB estimations will be evaluated. Sequencing artefacts are sequence changes that although present in the sequencing data are not present in the original sample; the presence of artefacts in solid tumour sequencing data would result in an over-estimation of TMB. The exclusion of potential sequencing artefacts prior to TMB assessment is not addressed specifically within the publications identified in this literature review. This is likely owing to the fact that there are a number of strategies that can be employed to

minimise sequence artefacts (Do and Dobrovic 2015) and these are commonplace practices within AWMGS and generally across the diagnostic and research communities. Also, within AWMGS, one of the key steps within NGS data analysis of tumour samples is to evaluate the variants identified using a specific set of criteria to determine if they are genuine variants present in the tumour; the aim of this evaluation is to exclude any variants that are deemed to be likely sequencing artefacts. The importance of this artefact removal step in the diagnostic laboratory is to minimise so-called false positive variant calls to ensure that diagnostic, prognostic and therapeutic decisions are correctly made based on the genuine genetic variants within a sample.

Artefacts can arise in a number of different areas within the sample workflow. Firstly, FFPE DNA samples are prone to deamination artefacts (C>T/G>A errors) owing to the tissue fixation process (Do and Dobrovic 2015https://academic.oup.com/clinchem/article/61/1/64/5611545). Secondly, PCR steps exist within many NGS protocols, including the library preparation protocols of the Illumina, Nonacus, and Agilent NGS panels used in this thesis, and this can introduce DNA polymerase errors. Polymerase base substitution error rates differ based on specific assay conditions, but have been quoted by Potapov and Ong (2017) as being in the range of 1/3,200 to 1/300,000 errors/base. Based on this estimated error range, in the 1.94Mb Illumina panel used in this project, 6-593 of the variants identified in a sample could be polymerase error-derived artefacts; these false positive variant calls would inflate TMB estimations. Notably, PCR duplicates have been shown to be much less common in hybridisation capture approaches, such as the three NGS panels used in this thesis, compared to amplicon enrichment protocols (Samorodnitsky *et al.* 2015). Finally, errors in short-read alignment within the sequencing protocol can cause read-end artefacts and strand bias artefacts (Koboldt 2020); the frequency of sequencing artefacts using NovaSeq technology (as used in this thesis) has been reported to be 0.1% (Ma *et al.* 2019).

Based on the difference between the AWMGS solid tumour analysis approach of excluding sequencing artefacts from NGS data, and published TMB assessment methods, which do not, as far as the literature review identified, perform additional artefact removal, this thesis will evaluate the impact of potential sequencing artefacts on TMB estimations. TMB quantifications of the patient cohort will be performed using calculations that either include or exclude sequencing artefacts.

The Institut Curie TMB tool (Github, 2022ahttps://github.com/bioinfo-pf-curie/TMB) will be used for the calculation of TMB values. This decision was based on the fact that this tool is freely available and would therefore have utility in both NHS and research applications going forwards. The tool is also transparent and can be easily manipulated, which provides the user with complete knowledge and control over the variants included within the TMB assessment. Notably, it is important within this thesis that the metrics of the TMB calculations performed across the three panels are the same so that results can be compared.

The TMB values of each patient sample assessed using each permutation of NGS panel and TMB calculation method will be translated into TMB high/low statuses based on the use of a TMB high threshold. As the use of a TMB high threshold of 10 variants/Mb has been validated in four independent studies (Ramalingam *et al.* 2018; Hellmann *et al.* 2018a; Pestinger *et al.* 2020; Marabelle *et al.* 2020), this threshold will be used in

this study. Receiver Operating Characteristic (ROC) curve generated thresholds will also be evaluated, as this method has proven utility as was the method used by Ramalingam *et al.* (2018) to determine the 10 variants/Mb cut-off, and Fancello *et al.* (2019) also showed the utility of this method. The use of a TMB threshold is key to assessing the utility of TMB as a biomarker and answering the primary research question, as it allows the patient survival (post PD-L1 test) of the TMB high and TMB low groups to be compared and analysed using appropriate statistical methods. If the hypothesis is correct, namely that this study's findings will emulate the findings of other publications (Carbone *et al.* 2017; Peters *et al.* 2017; Seiwert *et al.* 2018), then a positive correlation between increasing TMB score and longer survival would be expected. It also allows differences in survival (post PD-L1 test) between the PD-L1 expressor cohort and the TMB-high/PD-L1 expressor cohort to be compared. Differences in survival between these paired comparisons would indicate potential utility of TMB as a predictor of immunotherapy response. The sensitivity of the immunotherapy response predictions using TMB assessment in the PD-L1 expressor cohort in this study will be evaluated based on comparing the predicted responder (TMB high) and predicted non-responder (TMB low) outcomes to the actual immunotherapy responses (based on RECIST 1.1 criteria) of these patients.

Comparison and statistical evaluation of each permutation of NGS panel/TMB calculation/TMB threshold addresses one of the aims of the study by enabling the possible identification of a set of optimal criteria for TMB quantification that provides differentiation between immunotherapy responders and non-responders.

1.7.5 Assessment of the feasibility of TMB service implementation within the NHS

The evaluation of the feasibility of TMB assessment in the NHS is important as, if TMB is NICE-approved as an immunotherapy response predictor in the future, the practicalities of implementing a TMB service would have to be addressed within NHS Genomic laboratories, including AWMGS. The evaluation of the feasibility of a TMB service within the NHS will therefore encompass an assessment of the cost effectiveness of TMB analysis using targeted NGS, as well as an assessment of the availability of External Quality Assurance (EQA) schemes for TMB calculation. Investigation into these areas ensures that a broad understanding of the clinical utility of TMB quantification within the NHS is obtained.

The assessment of cost effectiveness will be done by considering both the cost of performing NGS analysis, based on the panels used within this study, as well as the utility of these targeted panels within the existing lung cancer patient pathway. There are a number of NSCLC clinically actionable genes for which testing is already provided in AWMGS (and other UK Genomics laboratories) namely: *EGFR, KRAS, BRAF, ALK, ROS1, RET, NTRK1, NTRK2, NTRK3*, based on the NICE-approval of relevant lung-cancer targeted drugs (NICE, 2022). This genetic analysis is already funded in the UK, therefore, if the panels used in this study target these NSCLC clinically actionable genes this could effectively make the introduction of TMB assessment a cost neutral service in the NHS.

The reason for investigating the availability of TMB-focussed EQA schemes is that EQA is an essential requirement of NHS Genomics laboratories in relation to maintaining

ISO151589 UKAS medical laboratory accreditation. EQA schemes distribute samples of known genotype/status to participating laboratories, measure the quality/accuracy of lab results, and allow labs to monitor performance against other labs. EQA schemes aim to ensure that patients receive results of good quality encompassing both the correct result and correct clinical interpretation of that result by providing quality assessment for participating laboratories. Any new service introduced into AWMGS requires a means of demonstrating external quality assurance, with an established EQA scheme being the most straight-forward means of providing this quality assurance. Given the variation in elements of TMB assessment, the approach to the establishment of an EQA scheme could be complex, so it will be interesting to see how such a scheme may have been developed.

1.8 The value of further research into the utility of TMB and PD-L1 assessment as a combined biomarker for immunotherapy stratification

The existing PD-L1 expression biomarker used both in the NHS and internationally for pembrolizumab stratification lacks accuracy in the identification of patients most likely to respond to this immunotherapy (Sul *et al.* 2016). TMB is a promising combinatory biomarker which, based on existing literature (Carbone *et al.* 2017), could be used to improve treatment stratification in PD-L1 high expressing lung cancer patients ensuring that patients access the most appropriate treatments in a timely manner. TMB could therefore be an area of high interest within the NHS for its potential to improve patient outcome.

There is currently no requirement for TMB assessment in NHS Genomic laboratories owing to there being no existing NICE-approved TMB-based stratifications of immunotherapy use; the AWMGS has no experience of TMB quantification. However, in June 2020, the first FDA-approval for the use of TMB as a biomarker for pembrolizumab stratification in solid tumours was granted (Marcus *et al.* 2021), increasing the likelihood that similar NICE-approval may be imminent. UK Genomic laboratories are required to deliver any genomic testing aligned to the use of NICE-approved drugs within 60 days of the drug being approved. Therefore, any NICE-approvals based on the use of TMB as a biomarker would require timely validation, implementation and delivery of an NHS service for TMB assessment.

To ensure precious NHS resources are used appropriately, it would be preferable for NHS Genomic labs to have guidance regarding the most clinically appropriate methods for TMB assessment. Currently there is no such guidance and the choice of NGS panel, TMB calculations and TMB high thresholds used within TMB research publications is extremely varied. Each of these variables will be evaluated in this thesis with the aim of identifying if there is a set of parameters (panel, calculation, threshold) for TMB estimation that provide the most accurate differentiation between immunotherapy responders and non-responders, at least within this patient cohort. This could guide the TMB quantification methods used in future research studies and NHS service validation efforts, and could form the basis of future best practice guidance in TMB assessment to ensure delivery of genetic testing of the highest possible standards (acgs.uk.com, 2022). The availability of best practice guidelines would facilitate the

translation of TMB assessment into the NHS. At a local level, the expertise gained from this research ensures that the AWMGS is well positioned to deliver NGS-based TMB analysis should a service be required in the future.

Direct comparison of TMB scores generated from different NGS panels using the same sample cohort was a novel area of TMB research at the time of thesis conception. The different NGS panels evaluated in this thesis will provide valuable additional insight into the impact of targeted panel size and gene content on TMB estimation. The evaluation of different methods of calculating TMB within the same patient cohort will provide novel data observing the impact on the TMB score of the inclusion/exclusion of synonymous variants, and inclusion/exclusion of sequencing artefacts.

The utility of combined TMB and PD-L1 assessment for the stratification of immunotherapy use in a Welsh population has not been explored to date. This pilot study will provide novel data from a small Welsh patient cohort, which could highlight a need for further research studies evaluating more expansive cohorts of Welsh patients to strengthen any findings from the pilot.

Chapter 2: Methods

Method sections 2.1 to 2.9 focus on the evaluation of clinical utility of TMB in terms of establishing the impact of TMB status on immunotherapy response within this Welsh patient cohort. Method section 2.10 describes the feasibility of the implementation of a TMB-based service in the NHS.

2.1 Patient selection

Patient samples were retrospectively identified and selected from the Welsh Cancer Bank. Patients had NSCLC and had all received anti-PD-1 pembrolizumab in the first line setting following PD-L1 positive IHC results. Treatment response and overall survival data were obtained through medical record review. The measurement of treatment response in this patient cohort broadly followed the categories defined in the RECIST 1.1 criteria (Eisenhauer *et al.* 2009; table 8), which evaluates response based on the size of the tumour as well as considering the tumour marker levels and size of non-target lesions. The RECIST 1.1 criteria are used for assessing cancer responses in trial settings. Whilst the treatment response terminology from RECIST 1.1 is used in routine NHS practice, as indicated from this Welsh patient cohort, local clinical consultation revealed that the guidelines are not followed rigidly in terms of the detailed assessment of size changes of the tumour as this is very time consuming and not practical within the routine clinical care setting.

FFPE tumour specimens were sourced from 17 of the selected patients from the WCB, which has ethics approval from Wales Research Ethics Committee to collect and issue biomaterials for projects using anonymised samples. 12 patients had stage 4 cancer, and 5 patients had stage 3 cancer. 5 x 10uM unstained tissue sections were received for each patient, along with 1 x 4uM H&E stained slide with the area of highest neoplastic cell content indicated by a trained Histopathologist. Number of days survival post-PDL1 test date was used as a quantifiable measure of immunotherapy response. Patient survival was calculated from the date of the PD-L1 IHC test, which for this cohort was over a 2-year time-frame of between August 2017 and May 2019, to the date of death or (where no date of death) until 11/04/2022, which is when the final data collection from WCB was performed and the patient was confirmed as alive. Patients were classified as responders if they experienced partial or complete responses or stable disease by RECIST 1.1 (Eisenhauer *et al.* 2009).

Table 8: RECIST 1.1 criteria for assessment of treatment response (Eisenhauer *et al.* 2009). CR = complete response; PR = partial response; PD = progressive disease; SD = stable disease; IR = incomplete response.

Response assessment	RECIST guideline, version 1.1
Target lesions	
CR	Disappearance of all target lesions and reduction in the short axis measurement of all pathologic lymph nodes to ≤10 mm
PR	≥30% decrease in the sum of the longest diameter of the target lesions compared with baseline
PD	≥20% increase of at least 5 mm in the sum of the longest diameter of the target lesions compared with the smallest sum of the longest diameter recorded **OR** The appearance of new lesions, including those detected by FDG-PET
SD	Neither PR nor PD
Non-target lesions	
CR	Disappearance of all non-target lesions and normalization of tumor marker levels
IR, SD	Persistence of 1 or more non-target lesions and/or the maintenance of tumor marker levels above normal limits
PD	The appearance of 1 or more new lesions or unequivocal progression If patient has measurable disease, an increase in the overall level or substantial worsening in non-target lesions, such that tumor burden has increased, even if there is SD or PR in target lesions If no measurable disease, an increase in the overall tumor burden comparable in magnitude with the increase that would be required to declare PD in measurable disease (eg, an increase in pleural effusions from trace to large, or an increase in lymphangitic disease from localized to widespread)

2.2 DNA extraction

Macrodissection was performed prior to DNA extraction of each sample. The area of highest neoplastic cell content was scraped off the unstained slides using a sterile scalpel blade; this process minimises the amount of normal cellular material within the DNA extraction and increases the sensitivity of the testing.

DNA extraction was performed on the 17 patient samples using Promega Maxwell 16 (Promega UK Ltd, Southampton UK), a benchtop nucleic acid extraction robot allowing the simultaneous automated extraction of up to 16 samples, with the Maxwell® RSC FFPE DNA Kit (Promega UK Ltd, Southampton UK) as per supplier protocol (appendix 6). DNAs were eluted in 72ul of nuclease free water, and DNA was quantified according to the supplier protocol (appendix 7) using the Qubit fluorometer high sensitivity assay (ThermoFisher Scientific, Loughborough UK) on the Qubit™ Flex fluorometer (ThermoFisher Scientific, Loughborough UK); the high sensitivity assay is able to quantify DNA between 0.2-100ng. 260:280nm absorbance data was generated using the Nanodrop™ 2000 Spectrophotometer (ThermoFisher Scientific, Loughborough UK) to highlight the presence of any protein or RNA contaminants in the DNA as required

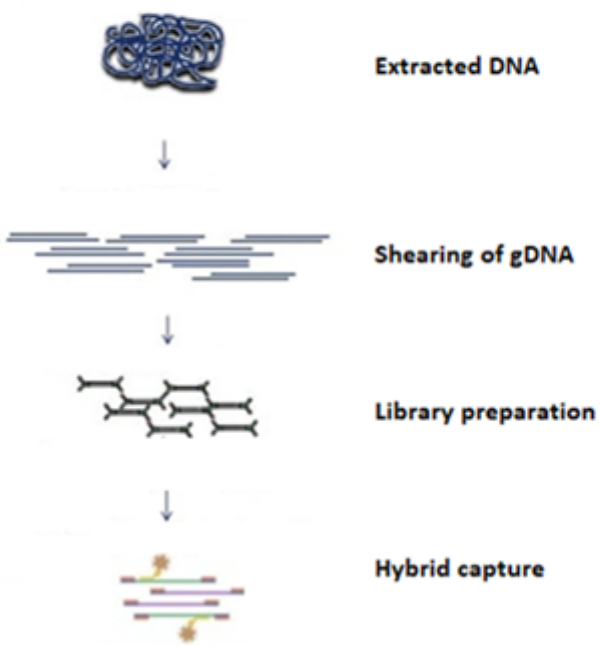
for results interrogation. Some samples were used for other research projects so <72ul was available following DNA extraction.

2.3 Next Generation Sequencing: Target enrichment

Hybridisation capture-based target enrichment was performed using three NGS panels: Illumina TruSight™ Oncology 500 panel (Illumina, Cambridge UK), Agilent SureSelect Community Design Glasgow Cancer Core panel (Agilent Technologies LDA UK Limited, Stockport UK), and Nonacus Cell3™ Target: Pan Cancer panel (Nonacus Limited, Birmingham UK) (figure 8). These three panels have differences in gene content (tables 5 and 6 in introduction chapter).

Where sufficient DNA (175ng) was available, target enrichment was performed on all three NGS panels (n=12); this sample number ensured the minimum requirements for paired t test statistical evaluation of data were met. Based on DNA availability, some DNA samples were processed on Nonacus alone (n=1), Illumina alone (n=1), or dual enrichment was performed: Agilent and Nonacus (n=1), Agilent and Illumina (n=1), Nonacus and Illumina (n=1). The minimum study requirements of analysing eight samples on each of the NGS panels was met enabling the Spearman correlation test to be used for statistical analysis of data.

Figure 8: Overview of steps within a hybridisation capture-based target enrichment (adapted from Qiagen.com, 2022). Extracted DNA is sheared enzymatically or mechanically within the first step of a hybrid capture target enrichment protocol to generate fragments of approximately 200bp in length, as short-read sequences are required for sequencing on the Illumina NovaSeq™. Library preparation involves end repair and A-tailing of the fragmented DNA, prior to adapter ligation. Adapter ligation links an adapter to the DNA fragments; these adapters are multi-purpose as they contain indexes (or barcodes) that are applied to different patient samples allowing multiple samples to be simultaneously sequenced, and the adapters enable the sequences to attach to a flow cell for sequencing. In the Illumina, Agilent and Nonacus protocols, a pre-capture PCR step is then performed [cycle number = 15, 12, and 7-12 (dependent on input DNA) respectively]. Hybrid capture with target-specific probes then takes place, which enriches for the target regions. Within the Illumina, Agilent and Nonacus protocols, a PCR step is then performed (cycle number = 18, 11, and 7-16 respectively) to amplify the enriched libraries prior to sequencing.



2.3.1 Sequencing library preparation: Illumina

Targeted enrichment of 15 samples (DNA input: 100ng) was performed using the Illumina TruSight™ Oncology 500 panel (Illumina, Cambridge UK) as per supplier protocol (appendix 8). This panel targets 523 genes implicated in the pathogenesis of solid tumours, representing 1.94Mb of the genome.

2.3.2 Sequencing library preparation: Agilent

Targeted enrichment of 14 samples (DNA input: 50ng) was performed using the Agilent SureSelect Community Design Glasgow Cancer Core panel (Agilent Technologies LDA UK Limited, Stockport UK) as per supplier protocol (appendix 9). This panel targets 174 genes implicated in the pathogenesis of solid tumours, representing 1.7Mb of the genome.

2.3.2 Sequencing library preparation: Nonacus

Targeted enrichment of 15 samples (DNA input: 25ng) was performed using the Nonacus Cell3™ Target: Pan Cancer panel (Nonacus Limited, Birmingham UK) as per supplier protocol (appendix 10). This panel targets 524 genes, including 116 cancer driver genes, and 345 genes in vital cancer signalling pathways, representing 1.58Mb of the genome.

2.4 Next Generation Sequencing and bioinformatic data analysis

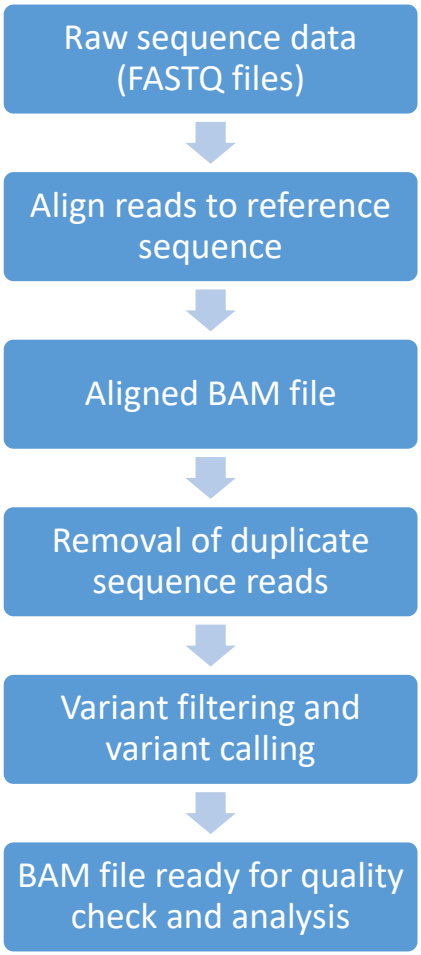
All libraries were sequenced using paired end sequencing on the Illumina NovaSeq™ system as per supplier protocol. Illumina technology utilises reversible dye terminator SBS chemistry involving reiterative cycles of single base incorporation, imaging and cleavage of the terminator chemistry (Meldrum *et al.* 2011). The sequencing data from the three panels was run through the AWMGS SomaticEnrichment pipeline v2.0.0 (Github, 2022bhttps://github.com/AWGL/SomaticEnrichment), which has been validated in-house for clinical diagnostic use in the analysis of data from hybridisation capture-based panels for the detection of somatic variation including single nucleotide variants and indels (summarised in figure 9).

Once the data had been run through the SomaticEnrichment pipeline v2.0.0 (Github, 2022b), a manual Quality Control check was performed using the fastQC tool (bioinformatics.babraham.ac.uk, 2022) to check the quality of the sequencing in terms of the Q30 Phred score, and a depth of coverage tool to check that targeted regions contained coverage above a 250x threshold. According to evidence described by Petrackova *et al.* (2019), a coverage depth of 250x should be sufficient to detect alleles down to a 5% variant allele frequency with a threshold of variant supporting reads ≥5. The use of this 250x sequencing threshold minimises false negatives, and the use of a 13-variant read threshold in order to deem a variant as genuine ensures that false positive rates are minimised.

The Q30 Phred score relates to the quality scores that are assigned to each base during sequencing. A Q score of 30 is equivalent to the probability of an incorrect base call 1 in 1000 times, which relates to a base call accuracy of 99.9%. Q30 is considered the benchmark for quality in NGS (illumina.com, 2011). Following this manual Quality Control check, additional annotation of the sequencing VCF files using the SnpEff tool (pcingola.github, 2021), which predicts the effects of genetic variants on genes and proteins, was performed as per the requirements of the Institut Curie TMB tool (Github, 2022a).

Figure 9: Illustration of the AWMGS bioinformatics pipeline for solid tumour NGS analysis (figure adapted from Koboldt 2020). The SomaticEnrichment pipeline v2.0.0 (Github, 2022b) is used for NGS data analysis. Raw sequencing data is aligned to the GRCh37 version of the reference genome using BWA-mem (bio-bwa.sourceforge.net, 2022) for sequence mapping. Duplicate reads that have originated from the same DNA sequence molecule are identified using GATK (Picard) MarkDuplicates tool (gatk.broadinstitute.org, 2022a) and are removed from the BAM file. The Mutect2 (gatk.broadinstitute.org, 2022b) tool is used https://gatk.broadinstitute.org/hc/en-us/articles/4409917447707-Mutect2**for variant filtering (e.g. filtering out unmapped reads) and variant calling (e.g. naming of single nucleotide variants and indels). Removal of germline variants within the SomaticEnrichment pipeline is**

achieved using the flag `--genotype-germline-sites false` with reference to the Genome Aggregation Database (gnomad.broadinstitute.org, 2022). The data is now ready for quality checks, which are described in the main text.



2.5 Tumour Mutational Burden estimation

TMB (variants/Mb) was calculated for each sample on each of the successful NGS runs using the Institut Curie TMB tool (Github, 2022ahttps://github.com/bioinfo-pf-curie/TMB). The Institut Curie TMB tool is a versatile tool that filters variants from inputted vcf files according to criteria set by the user, before using the remaining post-filtering variants to calculate the TMB of each sample. This tool has several required inputs and several optional arguments, which were set to accommodate the needs of this project. The required inputs included: a vcf file annotated using both Mutect2 and SnpEff, sample ID, and a panel-specific bed file (hg19 reference genome; sourced from Illumina, Nonacus and Agilent technical support) in order to calculate the size of the genomic region interrogated. An additional parameter specifying 'coding regions only' was added to the calculation of the genomic region interrogated, in line with the fact that only coding regions were analysed for variant detection.

The optional arguments used within the Institut Curie TMB tool (Github, 2022a) are noted in table 9. These were used to ensure the quality of the data produced, minimising false negatives and false positives. Following this data filtering step, data

was exported to generate a spreadsheet per patient that included all of the remaining variants and the TMB score.

Table 9: Parameters used to filter variants within the sequencing data for use of the Institut Curie TMB tool (Github, 2022a). Any parameters which align with the default settings of this tool have been noted.

Parameter	Setting	Justification
Variant allelic frequency (sensitivity of assay)	0.05 (default)	A 5% variant is considered the threshold of clinical utility within existing AWMGS lung cancer NGS services. This threshold correlates with the expected level of sensitivity of the NGS assays when a minimum coverage of 250x is achieved (Petrackova *et al.* 2019).
Minor allele frequency (the frequency of the second most frequent allele for a given SNP in a population)	0.001 (default)	This low threshold gives confidence in the variant allele being somatic in origin (Koboldt 2020).
Minimum depth (number of reads required to call variant genuine)	13	Reflects a 5% variant being detected at minimum coverage of 250x (Petrackova *et al.* 2019); therefore, minimises false positives.
Minimum alternative allele depth	2 (default)	Minimises false positives.
Filter low quality variants (i.e. not PASS)	'true'	Removes poor quality variants, minimising false positive variant calls.
Filter non-coding variants	'true'	All coding alterations, including short variant alterations, base substitutions, indels, and silent alterations were all counted in the TMB estimation, whereas other non-coding alterations were excluded in line with this project design.

Data was run through the TMB tool twice to provide two datasets per panel, once in which the optional argument of 'filter synonymous variants' was set to true and once where this criterion was set to false.

The TMB value (variants/Mb) generated from the TMB tool is in essence the total number of variants counted (as per variant criteria noted in table 9 above) divided by the size of the coding region of the panel in Mb (802,968 bases for Illumina panel and 835,198 bases for Nonacus panel).

2.6 Evaluation of sequence artefacts

Sequence artefacts are sequence changes identified in sequencing data that are not present in the original sample. These artefacts can arise in a number of different areas within the sample workflow as discussed in the introduction. Do and Dobrovic (2015) identified a number of strategies that could be used to limit the number of sequence artefacts from FFPE DNA. A number of these proposals have been incorporated into

this methodology with the aim of minimising the number of sequencing artefacts in the datasets and maximising the accuracy of the TMB estimations (table 10), as sequencing artefacts will artificially raise the TMB value.

Table 10: Strategies employed to minimise sequence artefacts within the dataset (adapted from Do and Dobrovic 2015). In terms of the DNA extraction, a 2-minute 80^{0}C incubation is the first step of the Maxwell protocol (appendix 6), and this is followed by a 30-minute proteinase K incubation step, both of which serve to potentially reduce artefacts as highlighted in the table.

Step	Strategy
DNA extraction	Macrodissection of tumour-enriched areas as determined by pathologist.
	Use of sufficient tissue, whenever possible, to maximise DNA yield.
	Heat treatment to remove formaldehyde-induced crosslinks and to facilitate subsequent tissue digestion with proteinase.
	Extended proteinase K treatment to digest tissue and to remove proteins cross-linked to DNA.
DNA assessment	Assessment of double stranded DNA quantity using Qubit fluorometer high sensitivity assay (ThermoFisher Scientific, Loughborough UK).
Library preparation	Minimise number of PCR cycles.
	Adhere to supplier protocols and use recommended DNA inputs to maximise DNA templates in the sequencing reaction.
	Hybridisation-capture based target enrichment allows the recognition of the initial templates in sequence reads using their unique start and end sites.
Identification of genuine sequence variants	Implement minimum variant allele frequency parameter in data analysis to minimise inclusion of low level (<5%) sequence variants (including artefacts) in TMB estimation.

In addition to these measures for reducing artefacts, this thesis aims to investigate the utility of further interrogation of sequencing data to identify and remove likely sequencing artefacts from the datasets prior to TMB estimation in line with existing AWMGS NGS workflows for solid tumour sample analysis. To this end, the two datasets per NGS panel (one with synonymous variants removed, and one with synonymous variants included) were each interrogated using Excel functionality to identify the variants shared frequently by patients within a dataset. Variants common to many patients within each panel cohort are likely to represent SNPs, hotspot variants, or artefacts (Bewicke-Copley *et al.* 2019). To identify the most likely false-positive variant calls within the datasets, the most commonly shared variants (n = 54) from each dataset were interrogated in dbSNP (Sherry *et al.* 1999) and visualised in IGV v2.10.3 (igv.org, 2022). This represented between 3% and 7% of all shared variants within each of the four datasets. Variants were classified as a polymorphism if the alternate allele frequency provided by the ALFA project European dataset (available via dbSNP; Sherry *et al.* 1999) was >0.1, representing an allele with estimated population frequency of 1% (Karki *et al.* 2015). The ALFA-generated allele frequencies are based on data from over two million subject entries within the NCBI database of Genotypes and Phenotypes (dbGAP) (Phan *et al.* 2020). Within IGV, a selection of criteria was used

to determine whether the variants, with alternate allele frequencies <0.1, were likely to be artefacts (table 11).

Once the selection of shared variants from each of the two datasets per NGS panel (one with synonymous variants removed, and one with synonymous variants included) had been analysed using dbSNP and IGV, and likely artefacts had been identified, this data was used to estimate the number of artefacts within each of the four datasets. The TMB scores for each patient in each of the two datasets per NGS panel (with and without synonymous variants) were adjusted to account for the removal of predicted artefacts.

Table 11: Features suggestive of an artefact rather than a genuine variant within the tumour sample (AWMGS unpublished data). The first criteria is described by Koboldt 2020.

1	Variant may not be present in both forward and reverse strands
2	Variant may not be present in overlapping reads
3	The sequencing around the variant may be of poor quality
4	Variant may be in a homopolymer repeat sequence or repetitive region.
5	The location of the variant may have multiple low-level sequencing changes such as other nucleotides substitutions.
6	Variant may be a deamination artefact (C>T or G>A)

2.7 TMB high threshold setting

The four Illumina and four Nonacus datasets (TMB estimation with/without the inclusion of synonymous variants, and with/without the removal of sequencing artefacts) were interrogated using a minimum of 11 different TMB high thresholds (table 12). The lowest TMB high threshold used across all panel/analysis datasets was 10 variants/Mb, which is the threshold that has proven utility in stratifying immunotherapy responders and non-responders in TMB studies by Ramalingam *et al.* (2018), Hellmann *et al.* (2018a), Pestinger *et al.* (2020) and Marabelle *et al.* (2020). The other TMB high thresholds investigated were chosen based on the TMB scores of that particular NGS panel/analysis dataset (explained further below). The TMB high thresholds evaluated were a minimum of 10 variants/Mb from one another, e.g. 90 variants/Mb and 100 variants/Mb thresholds were evaluated in all but one of the 8 panel/analysis combinations.

The sensitivity and specificity of each TMB high threshold in relation to the correct classification of patients into responder/non-responder groups was determined. The TMB high threshold most effective at determining immunotherapy response in the patient cohort, based on maximising true positive rate (to maximise sensitivity) and minimising false positive rate (to maximise specificity), was identified using ROC-curve analysis. Owing to the evaluation of sensitivity and specificity of each threshold, it was not necessary to evaluate the same thresholds for each panel/analysis combination as, sensitivity and/or specificity will only change if a TMB score of a patient within the cohort lies between the last thresholds investigated. For example, if 2 patients are in a cohort and one has a TMB score of 100 variants/Mb and the other has a TMB score of 150 variants/Mb, a TMB threshold of between 110 and 140 variants/Mb will have the same sensitivity and specificity. Effectively, it could therefore be considered that TMB

high thresholds were evaluated within this project at 10 variants/Mb intervals between the lowest and highest TMB high thresholds (shown in table 12), although in practice this was not necessary as illustrated in the example above.

Table 12: The range of TMB high thresholds investigated in each of the NGS panel/analysis combinations to determine sensitivity and specificity for prediction of immunotherapy response in each patient cohort. The highest threshold was between 140-440 variants/Mb, which varied dependent on the panel/analysis combination.

TMB calculation parameters	Range of TMB high thresholds investigated (variants/Mb)			
	Illumina		Nonacus	
TMB estimation including synonymous variants with removal of artefacts	10	300	10	320
TMB estimation including synonymous variants without removal of artefacts	10	400	10	440
TMB estimation excluding synonymous variants with removal of artefacts	10	290	10	140
TMB estimation excluding synonymous variants without removal of artefacts	10	290	10	290

2.8 Statistical analysis

Mean TMB estimations from the cohort of patients (n=13) whose tumour samples were analysed on both the Illumina and Nonacus panels, were evaluated for any statistically significant differences using a paired t-test. A total of four paired t-tests were performed to account for the four different analysis strategies used, namely TMB estimation with/without the inclusion of synonymous variants, and with/without the removal of sequencing artefacts. This evaluation provides information regarding the importance of panel selection and variant selection on TMB estimations, which links to two of the key aims of this project.

Scatter plots were used to visualise the relationship between immunotherapy response (response plotted as post-PD-L1 assessment survival in days) and TMB estimation (variants/Mb) from all Illumina (n=15) and Nonacus (n=15) NGS library preparations. A total of eight scatter plots were required to visualise both panel datasets across the four different analysis strategies (namely TMB estimation with/without the inclusion of synonymous variants, and with/without the removal of sequencing artefacts). Spearman rank correlation coefficient (r_s) calculations were performed on each of the eight datasets to investigate any association between TMB score (variants/Mb) and immunotherapy response, with the strongest association being identified by the highest r_s value.

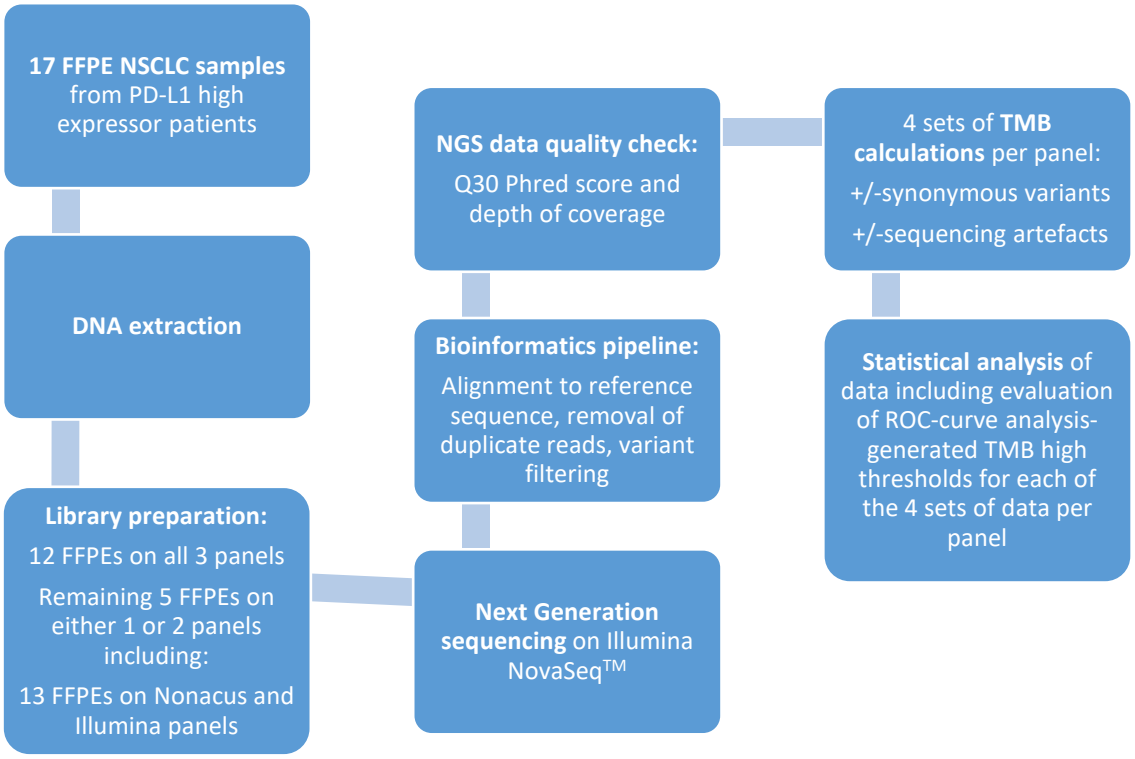
The panel/analysis package that gave the strongest association between TMB score and immunotherapy response using Spearman rank analysis was then further evaluated. Using a ROC-curve generated TMB high threshold, the patient cohort was divided into TMB high and TMB low groups and the median survival time of each group was calculated. A log rank test was used to identify any statistical significance between the Kaplan Meier survival curves of the TMB high and TMB low patient groups. This

evaluation provides the key piece of evidence to answer the primary research question of this thesis regarding whether there is clinical utility in the use of TMB status in combination with PD-L1 expression as a biomarker for immunotherapy response, as clinical utility would be suggested by a significant difference in survival times between the TMB high + PD-L1 expressor (>50%), and TMB low + PD-L1 expressor (>50%) patient groups.

The potential benefit of using PD-L1 expression status in combination with TMB score for immunotherapy response prediction compared to the use of PD-L1 expression status alone, was investigated using the Kruskal-Wallis test. This was performed specifically for the cohort of patients evaluated using the panel/analysis package that gave the strongest association between TMB score and immunotherapy response using Spearman rank analysis. The median survival of the PD-L1 >50% group (n= 15), and the TMB high + PD-L1 >50% group (n=10) was compared and evaluated for statistical significance. The expectation was that, in accordance with published data in non-Welsh cohorts, the combined TMB and PD-L1 assessment would be an improved biomarker for immunotherapy response prediction demonstrated by a statistically significant increase in survival of the TMB high + PD-L1 >50% group compared to the PD-L1 >50% group.

2.9 Summary of methods

Figure 10: Flow chart summarising the methods within this thesis used to answer the primary research question. The three NGS panels utilised are the Illumina TruSight™ Oncology 500 panel, the Agilent SureSelect Community Design Glasgow Cancer Core panel, and the Nonacus Cell3™ Target: Pan Cancer panel.



2.10 Evaluating the feasibility of TMB assessment in a clinical setting

2.10.1 Cost of a TMB service within AWMGS

The cost of TMB assessment per patient was calculated for the panels in this study that successfully generated sequencing data. These calculations considered the cost of the targeted panel kit, the sequencing costs, and staff time associated with the testing process, based on costings generated at AWMGS.

2.10.2 Investigation into the availability of EQA schemes focussed on TMB assessment

The investigation into the availability of TMB-based EQA schemes was performed by a website search of the two key providers of EQA schemes used by UK Genomics laboratories (including AWMGS): UK NEQAS (National External Quality Assessment Service; ukneqas.org.uk, 2022) and EMQN (European Molecular Genetics Quality Network; emqn.org, 2022). EQA schemes provide quality assurance for laboratories by assessing the performance of laboratories in the analysis and interpretation of a set of pre-validated samples. Participation in such EQA schemes is a requirement of an ISO15189 accredited laboratory.

Chapter 3: Results

3.1 DNA extraction

DNA extraction was successful from all 17 of the WCB-sourced FFPE tumour specimens from PD-L1 high expressor (>50%) NSCLC patients. Qubit concentrations were in the range of 2.5ng/ul – 50.0ng/ul, with a median DNA concentration of 10.8ng/ul; Nanodrop evaluation of a selection of samples showed that the 260:280 absorbance ratios ranged from 1.87-1.92 (excluding a poor-quality outlier at 260:280 of 0.54). Although each Promega Maxwell extraction provided 72ul of DNA, many of the 17 DNA samples were used in other research projects; the total quantity of DNA remaining for each patient sample was between 25ng and 1900ng, with a median quantity of 515ng. For all 17 samples, sufficient DNA was obtained to perform NGS using at least one of the target enrichment panels. There was enough DNA in 12 of the samples to perform target enrichment and sequencing using all 3 NGS panels (175ng required).

3.2 Sequencing library preparation

Dependent on the volume of DNA obtained from the FFPE extractions, sequencing library preparation was performed for between one and three of the NGS panels under investigation within this thesis (table 13).

Table 13: Patient cohort investigated in this thesis dependent on DNA availability. Sample IDs are noted across the top of the table. Green indicates that the sample was sequenced on the panel; red indicates that there was insufficient DNA to sequence the sample on the panel. DNA input requirements of each panel are: Illumina 100ng; Agilent 50ng; Nonacus 25ng.

	20M70071	20M70072	20M70073	20M70075	20M70076	20M70078	20M70079	20M70081	20M70084	20M70088	20M70090	20M70091	20M70074	20M70086	20M70087	20M70077	20M70089
Illumina																	
Agilent																	
Nonacus																	

3.2.1 Agilent

Sequencing failed on the Agilent panel for 14/14 patients, with all of the 1.7Mb panel achieving <250x coverage at each base position. This poor coverage across all regions of the panel was identified at the sequencing data Quality Control step, meaning no additional processing of the Agilent data using the Institut Curie TMB tool was performed. The root cause of this sequencing failure is not known. The quality control steps within the library preparation did not indicate any problems with the set-up. Importantly, the AWMGS laboratory has experienced previous issues with poor coverage of another Agilent NGS panel that were never resolved; the overlapping nature of the protocols for these 2 Agilent panels suggests that these panel failures could be linked, perhaps in terms of a common set-up error within the laboratory caused by lack of experience in this protocol.

3.2.2 Illumina

Q30 Phred sequencing quality score was 91.2% for the Illumina panel sequencing run. This indicates that 91.2% of bases sequenced on this sequencing run had a predicted quality score of 30 or more, which translates to a 1 in 1000 (0.1%) chance of error at each base position. This QC score predicts a low level of sequencing errors across the 15 samples that were processed on the Illumina panel, which gives confidence in the accuracy of the NGS data generated.

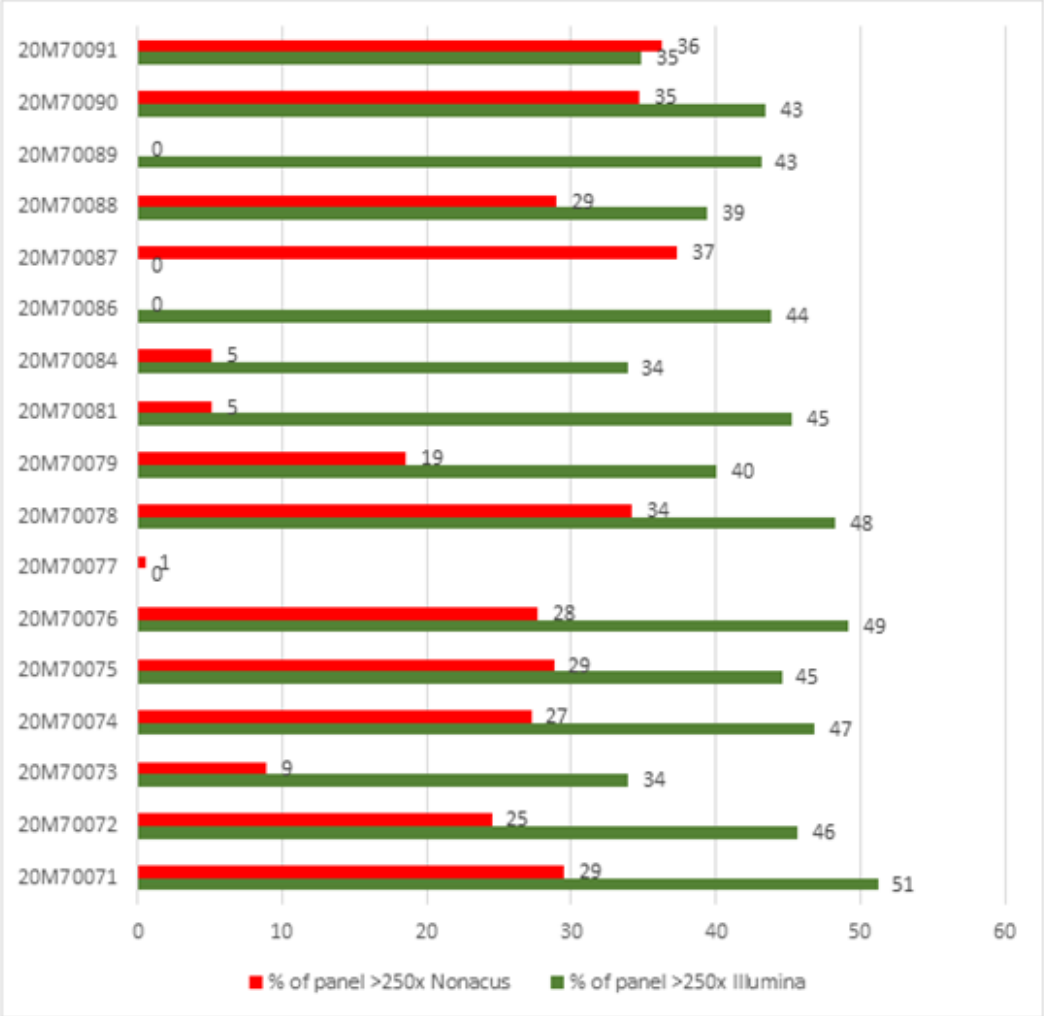
The second element of the manual Quality Control check in which the panel coverage was evaluated to check that targeted regions contained coverage above the 250x showed that across the 15 patient cohort a minimum of 34% of the Illumina panel had achieved the 250x threshold (figure 11). Based on the Q30 score and the coverage data, all 15 samples passed the Quality Control step and could continue on to the next stage of analysis to calculate TMB estimation.

3.2.3 Nonacus

Q30 Phred sequencing quality score was 83.5% for the Nonacus panel sequencing run, indicating that 83.5% of bases had a 0.1% chance of error in the base call.

Figure 11 shows the coverage achieved per sample. Despite only 27% (4/15) of the patients achieving 250x across >30% of the Nonacus panel, rising to 73% (11/15) of patients who achieved 250x coverage across 10% of this panel, all 15 samples continued on to the next stage of analysis to calculate TMB estimation and maximise data collection.

Figure 11: Comparison of Illumina and Nonacus panel coverage for all samples sequenced. 15 patient samples were prepped using the Illumina TruSight™ Oncology 500 panel and the Nonacus Cell3™ Target: Pan Cancer panel and sequenced on the Illumina NovaSeq™. The percentage of bases in the panel achieving over 250x sequencing coverage is shown for each patient. Note: owing to insufficient DNA, samples 20M70087 and 20M70077 were not sequenced using the Illumina panel, and samples 20M70089 and 20M70086 were not sequenced using the Nonacus panel.



3.2.4 Comparison of panel coverage

In general, sequencing coverage was more uniform between the samples sequenced using the Illumina panel with the percentage of bases being sequenced to 250x lying within a range of 34-51% across all 15 patients, with a mean coverage of 43% (figure 11). There was generally poorer coverage of samples prepped using the Nonacus panel, where the percentage of bases being sequenced to 250x was 1-37% across the 15 patient samples analysed, with a mean coverage of 23%.

Closer interrogation of sequencing coverage at the gene level indicated that all 15 samples sequenced on the Illumina panel covered 100% of the *EGFR* variant hotspot regions (exons 18-21) to a minimum of 250x, and an average of 97.8% of the *KRAS* variant hotspot regions (exons 2-4) were covered to 250x in the patient cohort. Using the Nonacus panel an average of 73.8% of the *EGFR* hotspots and 72.2% of the *KRAS* hotspots achieved 250x coverage in the 15-patient cohort. Review of sequencing

coverage data collated from a selection of lung cancer-related genes (*CDKN2A, PTEN, NRAS, RET, KRAS, ERBB2, PIK3CA, EGFR, MET, BRAF*) showed that the average coverage to 250x of this gene selection in the Illumina cohort was 55% compared to an average of 39% in the Nonacus panel cohort (data not shown).

3.3 Tumour Mutational Burden estimation

TMB estimation was performed for both the Illumina and Nonacus datasets using the Institut Curie TMB tool (Github, 2022a). Ultimately, four datasets were produced for each panel based on the inclusion/exclusion of synonymous variants (+synonymous/-synonymous) and the inclusion/exclusion of sequencing artefacts (+artefacts/-artefacts) within the TMB calculations.

3.3.1 Impact of NGS panel on TMB score

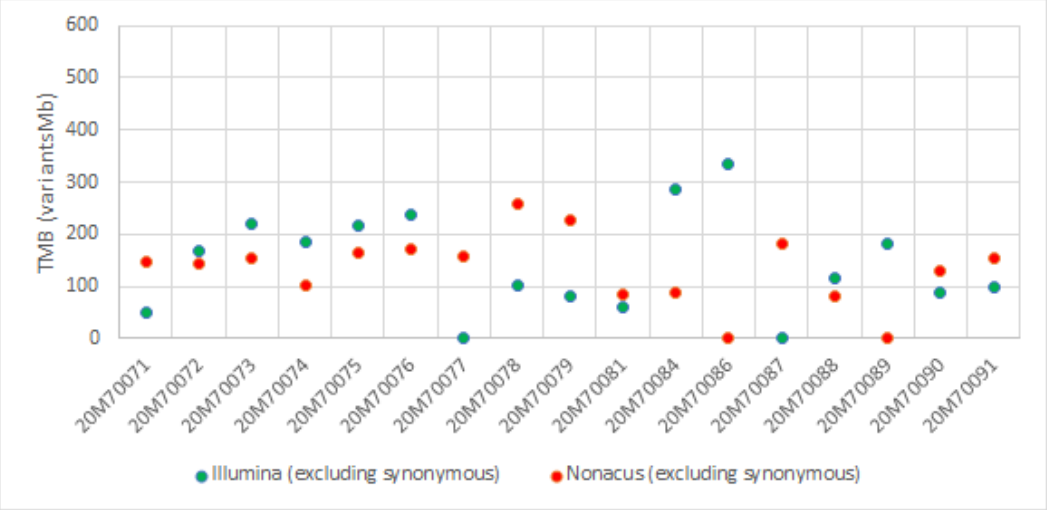
The same 13 tumour samples were sequenced on both the Illumina and Nonacus panels. The number of variants identified in each of these tumour samples varied dependent on the NGS panel used. When no additional artefact removal was performed (artefact removal is discussed later in results chapter), the Nonacus panel detected a higher number of variants per patient in 46% (6/13) of patients in the -synonymous dataset and in 85% (11/13) of patients in the +synonymous dataset (appendix 11). The average number of variants detected per sample across the 13 samples analysed on both NGS panels was greater when Nonacus targeting was used, although this difference between panels was relatively insignificant in the -synonymous dataset (-synonymous: Nonacus mean 122 and Illumina mean 118; +synonymous: Nonacus mean 215 and Illumina mean 148). The larger discrepancy seen in the number of variants detected between the Illumina and Nonacus panels using +synonymous conditions, can be related to the higher number of synonymous variants detected in the Nonacus-targeted samples (table 14).

The difference in the number of variants detected in the 13 samples sequenced on both the Illumina and Nonacus panels results in differing TMB scores (variants/Mb) for each patient dependent on the panel used (figure 12), as importantly the effective genome size interrogated within the targeted sequencing is similar for both panels (802,968 bases for Illumina and 835,198 bases for Nonacus). Focusing again on the +artefacts datasets, the Nonacus panel generated a higher TMB score per patient in 46% (6/13) of patients in the -synonymous +artefacts dataset and in 85% (11/13) of patients in the +synonymous +artefacts dataset.

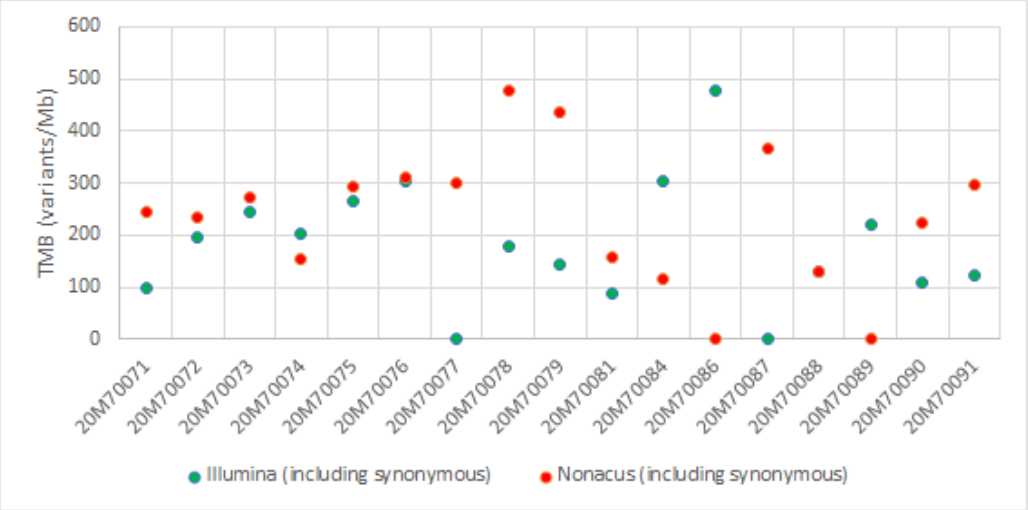
In accordance with the general higher variant number in Nonacus-targeted samples when using +synonymous (+artefacts) conditions, the average TMB score across the 13 samples analysed on both panels was higher across the Nonacus-targeted +synonymous samples (+artefacts: Nonacus mean 258 variants/Mb and Illumina mean 184 variants/Mb). This increase was also seen under -artefacts conditions: Nonacus mean 193 variants/Mb and Illumina mean 178 variants/Mb. The impact of this variation in TMB scores on overall TMB status (high/low) and subsequent treatment stratification of patients is highlighted later in the results chapter.

Figure 12: Scatter graphs to illustrate the impact of NGS panel on TMB estimation. The two datasets shown represent the: a) -synonymous and b) +synonymous TMB estimations. No artefact removal has been performed on this data. Note: Samples 20M70087 and 20M70077 were not sequenced using the Illumina panel, and samples 20M70089 and 20M70086 were not sequenced using the Nonacus panel; TMB scores for these patients are noted as zero.

a)



b)



3.3.2 Impact of synonymous variants on TMB score

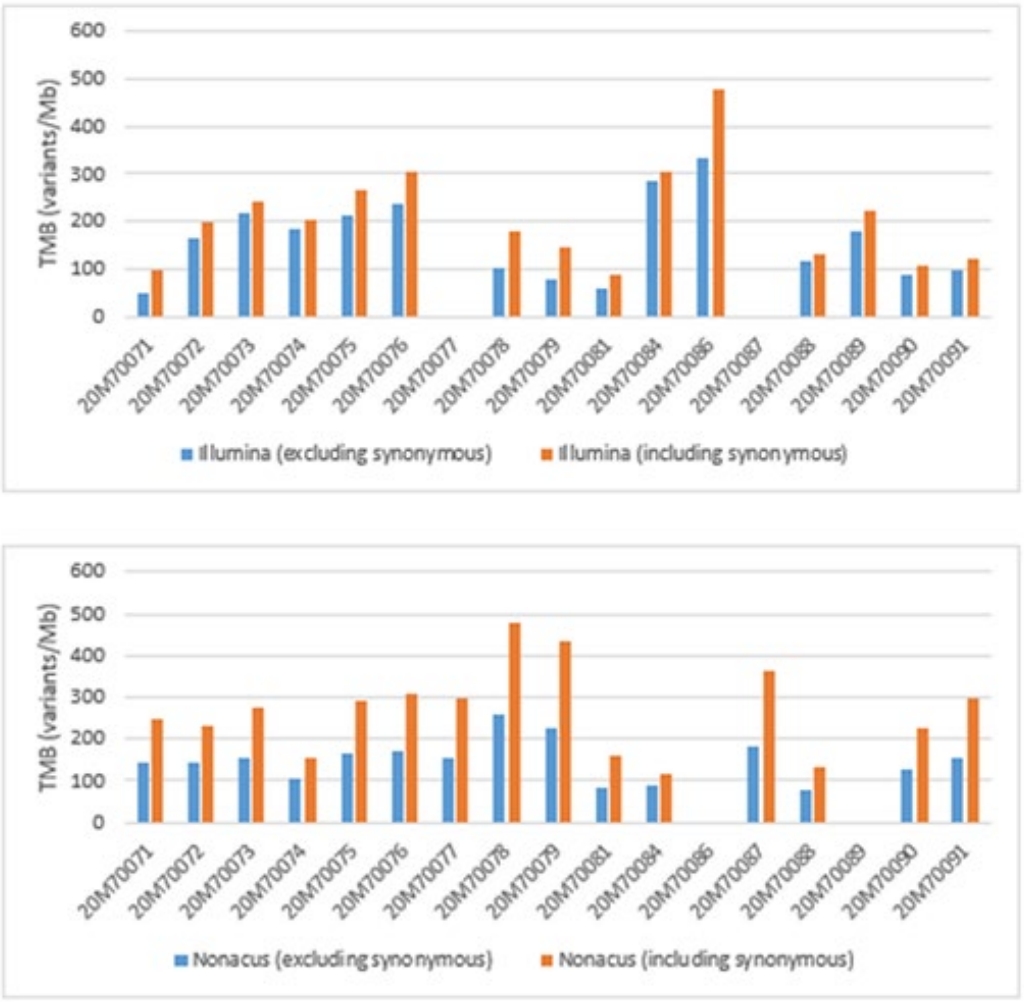
The number of identified variants increased in both the Illumina and Nonacus datasets (n=15 for both) as a whole with the inclusion of synonymous variants. The increases were remarkably different dependent on the panel used (table 14). Table 14 shows that the number of variants detected in the 13 patients sequenced across both panels followed the same trend as when analysing the 15 patient datasets, in terms of the differences between the number of variants detected per panel, and in terms of the impact of the inclusion or exclusion of synonymous variants on variants counted.

Table 14: Summary of the number of variants identified across the 15 patients sequenced on either the Illumina or Nonacus panels. This data shows the percentage increases in the number of variants detected between the -/+ synonymous datasets. It should be noted that 13 of each panel-specific cohort of 15 patients were the same, i.e. were sequenced on both panels; the variant data for these 13 patients is shown. Note: Additional artefact removal has not been performed in this data.

	Illumina sequencing data			**Nonacus sequencing data**		
	- synonymous	**+ synonymous**	**% increase**	**- synonymous**	**+ synonymous**	**% increase**
Number of variants identified across all 15 patients	1948	2482	**27%**	1866	3355	**80%**
Number of variants identified across same set of 13 patients	1534	1920	**25%**	1584	2800	**77%**

In the Illumina dataset there was a 27% increase in the number of variants called by the analysis pipeline when synonymous variants were counted, whilst the inclusion of synonymous variants in the Nonacus dataset resulted in an 80% increase in variant number (table 14). When visualising the variant data on a patient level rather than across the cohort as a whole, corroborative information is gathered in that the number of variants increased in all patients (n = 15) in both the Illumina dataset (35% average increase, correlating to 12-116 synonymous variants per patient) and Nonacus (77% average increase, correlating to 25-184 synonymous variants per patient) datasets with the inclusion of synonymous variants (appendix 11). The resulting TMB scores (variants/Mb) increased by 35% per patient on average in the Illumina dataset compared to 77% per patient on average in the Nonacus dataset as a result of the inclusion of synonymous variants within the calculations (figure 13).

Figure 13: Bar charts to illustrate the impact of inclusion of synonymous variants on TMB estimation. The graphs show the differing TMB scores (variants/Mb) generated from the Illumina and Nonacus datasets (total number of patients = 17; 15 patients sequenced on each panel; 13 patients sequenced on both panels) when synonymous variants are included or excluded from the TMB calculations. Note: Samples 20M70087 and 20M70077 were not sequenced using the Illumina panel, and samples 20M70089 and 20M70086 were not sequenced using the Nonacus panel.



3.3.3 Impact of artefacts on TMB score

Evaluation of data from each of the four NGS panel datasets (Illumina and Nonacus +/- synonymous) using Excel functionality was performed to identify shared variants within the patient cohorts of each dataset. The most commonly shared variants (54 in total across the 4 datasets) were interrogated in dbSNP and IGV to identify likely artefacts based on the criteria described in the methods chapter. See appendices 12 and 13 for full details of the variants interrogated and classified.

In the Nonacus and Illumina datasets, a varying percentage of shared variants were identified as artefacts following dbSNP and IGV interrogation (table 15-a). The Nonacus datasets had a much higher proportion of artefacts compared to the Illumina datasets based on the limited analysis of shared variants performed, and artefact rate was shown to be higher in the -synonymous datasets than in the +synonymous datasets. The artefact prevalence, based on the shared variants interrogated, was used to estimate the number of artefacts within each of

the four complete datasets (table 15-b). Artefact levels of 2-3% for the Illumina panel and 25-42% for the Nonacus panel were predicted.

Table 15: Interrogation of a proportion of shared variants within the two Illumina and two Nonacus datasets in dbSNP and IGV: a) to provide an estimation of artefact frequency across the shared variants of each dataset; b) to predict the % of artefacts within each 15-patient dataset using the %s determined in a). In a) the percentages represent the number of interrogated variants classified as artefact as a percentage of the total number of shared variants investigated in IGV.

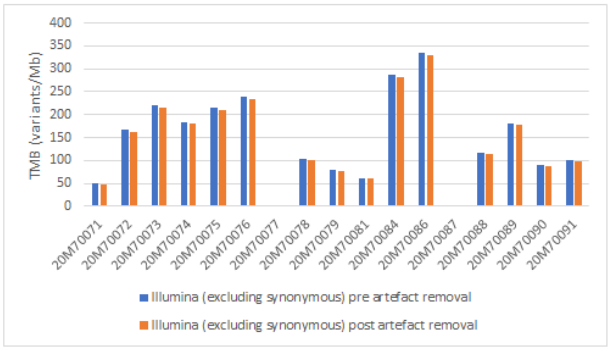
a)	Illumina sequencing data		Nonacus sequencing data	
	- synonymous	+ synonymous	- synonymous	+ synonymous
% of interrogated shared variants classified as artefact	17% (1/6)	15% (2/13)	62% (8/13)	36% (8/22)

b)	Illumina sequencing data		Nonacus sequencing data	
	- synonymous	+ synonymous	- synonymous	+ synonymous
I: Number of non-unique variants (i.e. each shared by >1 patient) across all 15 patients sequenced	210	441	1253	2373
II: Predicted number of non-unique variants within the 15 patient dataset representing artefacts (based on % in 'a') [a*I]	36	66	777	854
III: Total no of variants identified across all 15 patients sequenced	1948	2482	1866	3355
IV: Predicted % of variants within the 15-patient dataset representing artefacts [(II/III)%]	2%	3%	42%	25%

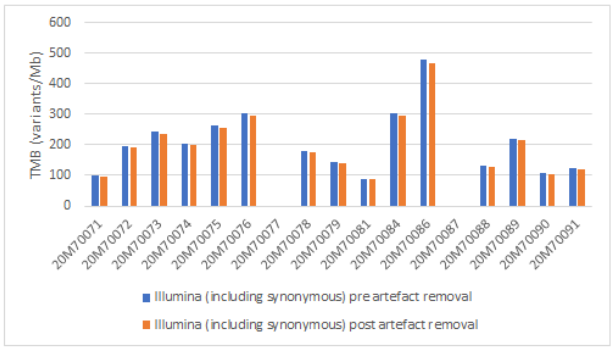
Re-calculation of TMB scores following removal of artefacts from the two Illumina and two Nonacus datasets (+/-synonymous) was performed using the predicted artefact levels in table 15-b (figure 14; appendix 11). The impact of differences in TMB scores on the TMB high/TMB low classification of patients will be described later in the results chapter.

Figure 14: Bar charts to illustrate the impact of artefact removal on TMB estimation. Variation in TMB score when predicted artefacts are removed from the total number of variants identified per patient in both the Illumina and Nonacus datasets. The number of variants detected in each patient was adjusted according to the predicted % of artefacts within each dataset (refer to table 15-b). Note: Samples 20M70087 and 20M70077 were not sequenced using the Illumina panel, and samples 20M70089 and 20M70086 were not sequenced using the Nonacus panel.

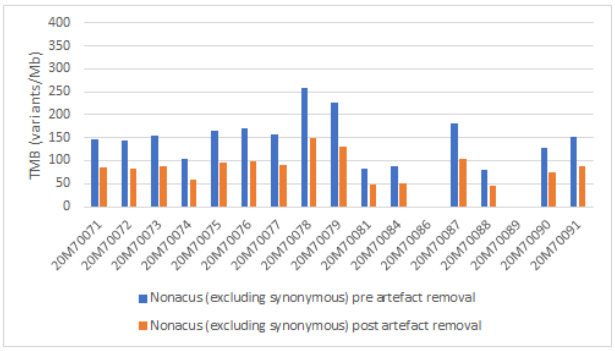
a)



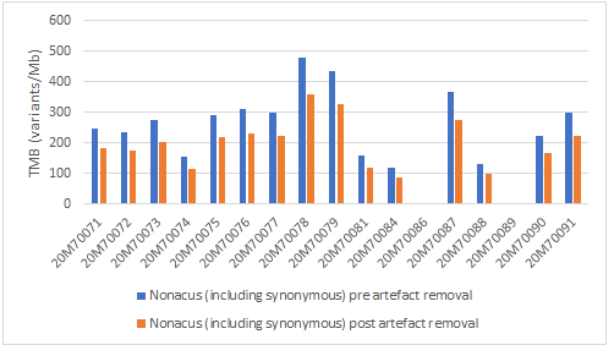
b)



c)



d)

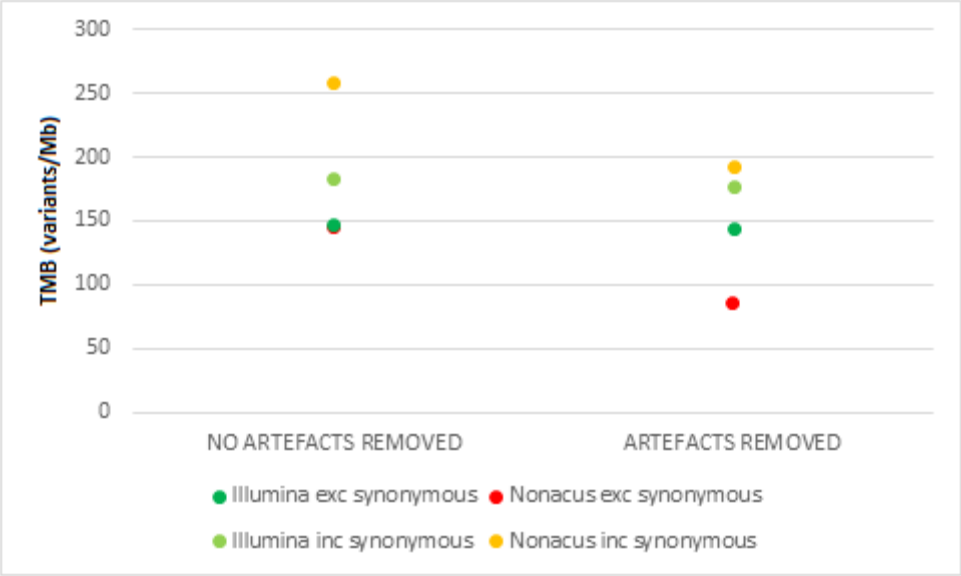


3.4 Statistical analysis

3.4.1 Comparison of TMB scores from Illumina and Nonacus NGS panels using paired t-tests

Within the cohort of patients (n=13) analysed on both Illumina and Nonacus panels, there is variation in TMB estimations when either the Illumina or Nonacus panel is used for sample interrogation (figure 12). This is true when artefacts are included or removed (figure 14), and is also true when synonymous variants are included or excluded from the calculations (figure 13). These differences in TMB estimations are summarised in figure 15 below. The statistical significance of the differences in TMB scores within this 13-patient cohort when different panels and different variant calculations are used, was determined using a selection of paired t-tests.

Figure 15: Variation in cohort mean TMB score using different panel/analysis conditions. The mean TMB scores are shown, calculated for the same patient cohort (n=13) analysed using different NGS panels and with different variant calculations applied. Note: The Illumina and Nonacus -synonymous datapoints are over-laid in the +artefacts dataset representing mean TMB scores of 147 and 146 respectively.



The t-test compares the dataset mean TMB scores (as shown in figure 15) to the null hypothesis stating that the mean difference between the mean TMB scores from Illumina and Nonacus enrichment for the patient cohort will be zero. Two-tailed t-tests were used as the alternative hypothesis is that the Illumina and Nonacus mean TMB scores for the cohort are not equal (no emphasis on which panel may have a lower mean). The degrees of freedom is 12 (n-1) and the t critical value is 2.17881283 for all datasets (+/-synonymous variants, and +/-artefacts). A t-value of zero would equate to there being no difference between the mean TMB scores in each dataset. The p-value gives the probability that the t-value observed will be larger than the t critical value.

Table 16: Paired t-test evaluation of variation in TMB estimations under different panel/analysis conditions. Evaluation of the statistically significant difference (level of significance α = 0.05) between the mean TMB scores generated from the Illumina and Nonacus panel datasets (n=13) when different methods of variant counting were applied (+/- synonymous variants and +/- artefacts).

	+ artefacts		- artefacts	
	- synonymous	**+ synonymous**	**- synonymous**	**+ synonymous**
t- value (2 s.f.)	0.038	-2.00	2.50	-0.49
p-value (2 d.p.)	0.97	0.069	0.028	0.63

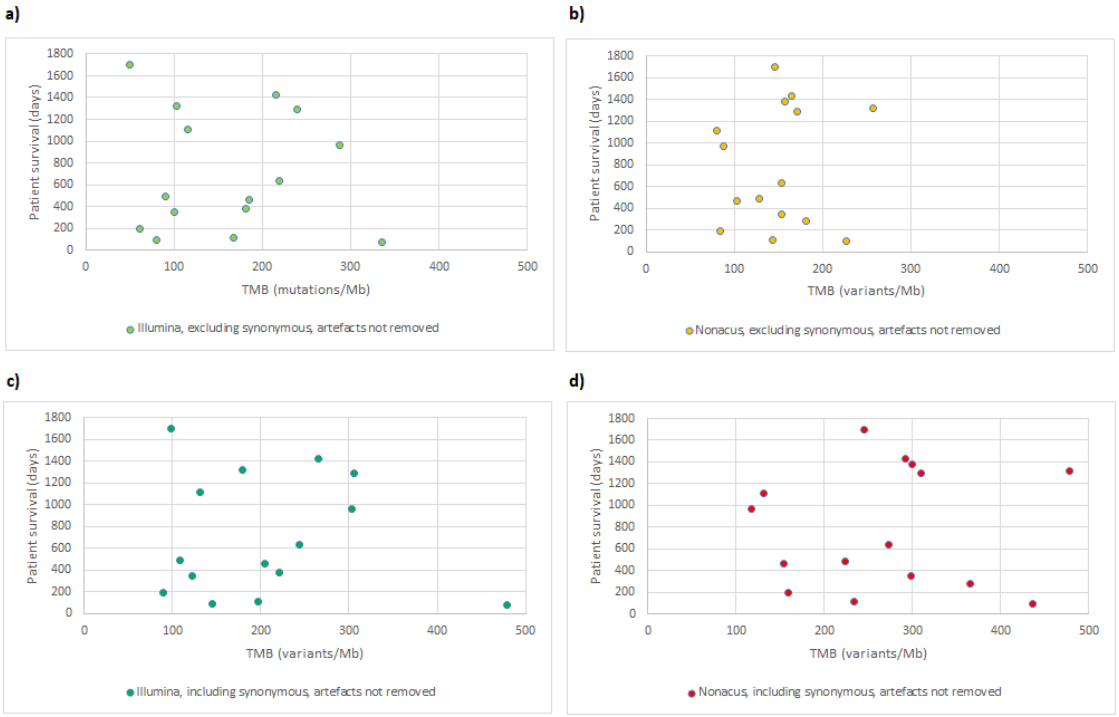
The difference between the mean cohort (n=13) TMB scores generated from Illumina and Nonacus (values of 144 and 85 variants/Mb respectively) is statistically significant within the -synonymous -artefacts dataset (table 16). In this dataset, the two-tail p-value is less than the 0.05 significance level, therefore the null hypothesis is rejected. Comparison of the t value in the -synonymous -artefacts dataset to the critical t value shows that the t value is greater than the critical t value.

3.4.2 Association between TMB scores and immunotherapy response

The association between TMB score and immunotherapy response is shown in the scatter graphs (figure 16) for the Ilumina and Nonacus +artefacts datasets. When artefacts were excluded, the datapoints in each of the four datasets shifted to the left to different degrees dependent on the number of artefacts within the panel/analysis dataset which reduced the TMB scores (plotted on the x axes) (data not shown).

Figure 16 demonstrates scattering of datapoints. E.g. in the Illumina cohort (-synonymous variants + artefacts; figure 16-a) patient survival >800 days was associated with TMB scores of between 50 and 288 variants/Mb; patients in the same cohort with similar TMB scores had survival of <200 days.

Figure 16: Scatter plots to visualise the relationship between TMB and immunotherapy response. Graphs a-d represent +artefacts datasets using: a) Illumina panel -synonymous; b) Nonacus panel -synonymous; c) Illumina panel +synonymous; d) Nonacus panel +synonymous. Survival is calculated as the number of days between the PD-L1 test to the date of death, or (where no date of death) until 11/04/2022 when patient was confirmed as alive by WCB. 15 patients were analysed on each panel.

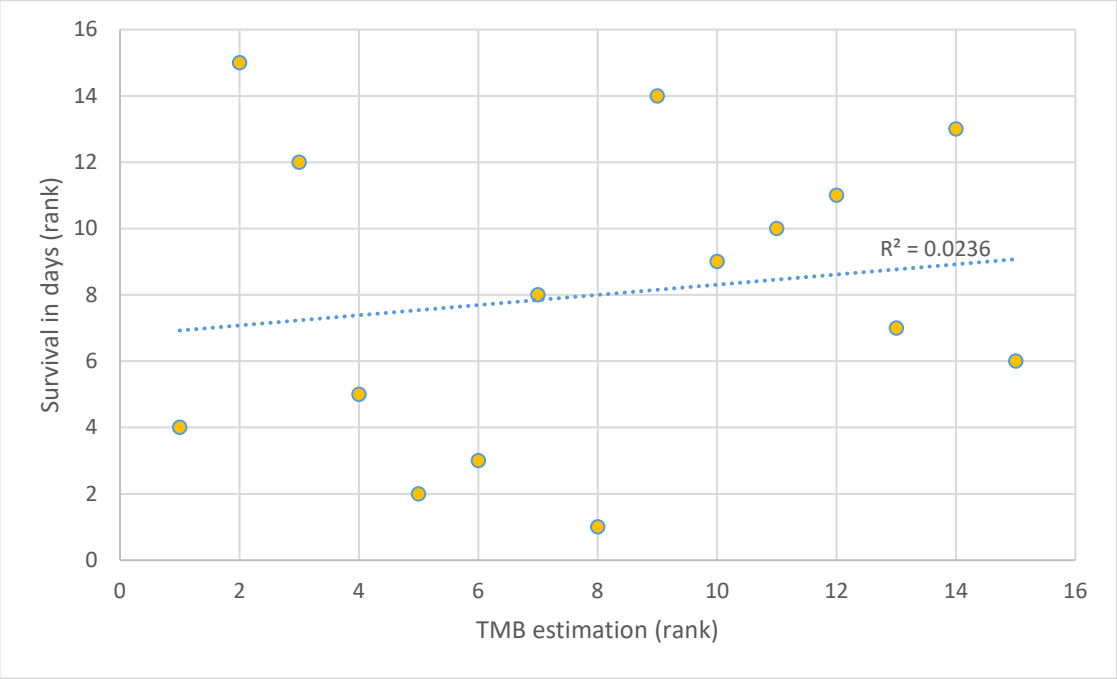


To describe the association between TMB score and immunotherapy response, the Spearman rank correlation coefficient (r_s) was calculated for each Nonacus and Illumina dataset, +/-synonymous and +/-artefacts. The r_s for all of the Nonacus and Illumina datasets show a weak positive correlation between TMB score and immunotherapy response (table 17). There was a trend for a stronger positive correlation, demonstrated by higher r_s values, between TMB score and immunotherapy response using the Nonacus panel compared to the Illumina panel. The strongest association between TMB score and patient survival ($r_s = 0.154$) was demonstrated using the Nonacus panel -synonymous +artefacts conditions (figure 17).

Table 17: Spearman correlation coefficient evaluation of the association between TMB score and immunotherapy response under different panel/analysis conditions. R_s values are noted in the table to describe the association between TMB score and immunotherapy response (measured in terms of survival time post PD-L1 test) within the patient cohort (n=15) using the Nonacus and Illumina NGS panels, +/-synonymous and +/-artefacts.

			Spearman correlation coefficient (r_s) to 3d.p.
Illumina	**+ artefacts**	**- synonymous**	0.007
		+ synonymous	0.011
	- artefacts	**- synonymous**	0.007
		+ synonymous	0.011
Nonacus	**+ artefacts**	**- synonymous**	0.154
		+ synonymous	0.132
	- artefacts	**- synonymous**	0.046
		+ synonymous	0.111

Figure 17: Spearman rank correlation between TMB score and immunotherapy response using Nonacus -synonymous variants +artefacts. Immunotherapy response is measured in terms of survival in days, which is calculated as the number of days between PD-L1 test to date of death or 11/04/2022. TMB values were measured in variants/Mb across the patient cohort (n=15).



3.4.3 Evaluation of TMB high thresholds for accurate immunotherapy response patient stratification

3.4.3.1 10 variants/Mb TMB high threshold

Using the 10 variants/Mb TMB high threshold suggested by Ramalingam *et al.* (2018), sensitivity, the number of patients who responded to immunotherapy (based on clinical record RECIST-1 data) who were correctly identified as TMB high (for both

Illumina and Nonacus datasets under all analysis conditions) was 100%. However, specificity was 0%, with 9/9 (100%) non-responders (according to RECIST-1 classification) being incorrectly classified as responders (i.e. TMB high) using the TMB estimations generated from both Illumina and Nonacus datasets under all analysis conditions. These sensitivity and specificity figures relate to the fact that all TMB scores of patients in this mixed cohort of responders and non-responders were greater than the 10 variants/Mb TMB high threshold (see figure 14).

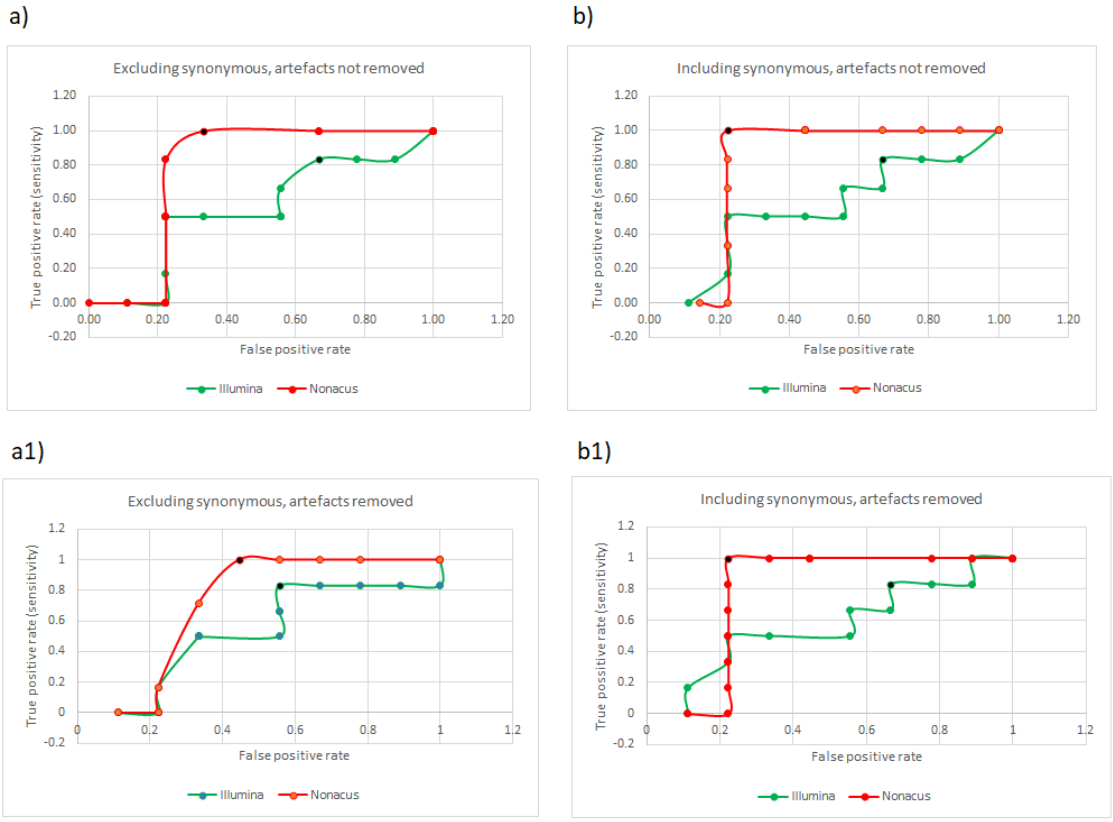
3.4.3.2 ROC-curve generated TMB high thresholds

Maximising sensitivity and specificity:

ROC curve analysis was used to determine the TMB high threshold most effective at determining immunotherapy response in this Welsh patient cohort, based on maximising true positive rate (to maximise sensitivity) and minimising false positive rate (to maximise specificity) (figure 18).

The optimal ROC-curve generated TMB high thresholds, showed that the maximum sensitivity and specificity for accurate responder/non-responder classification (responders defined using RECIST 1.1 criteria encompassing CR, PR and SD; Eisenhauer *et al.* 2009) was achieved using Nonacus +synonymous variants +/-artefacts (figure 18; table 18).

Figure 18: ROC curve analysis showing false positive rates and true positive rates for immunotherapy response classification using various TMB high thresholds. TMB high thresholds ranged from a lowest value of 10 variants/Mb (data point at true positive rate of 1 and false positive rate of 1) to a highest value within the range 290-440 variants/Mb dependent on the panel/analysis combination (TMB thresholds used are shown in appendix 14). The false positive rate is the ratio of negative samples (non-responders) that are incorrectly classified as responders [FP/(FP+TN)]. The true positive rate is the ratio of positive samples (responders) that are correctly classified [TP/(TP+FN)]. Responders were defined using RECIST 1.1 criteria (Eisenhauer et al 2009) encompassing CR (complete responders), PR (partial responders) and SD (stable disease). Graphs a and b represent +artefacts datasets (n = 15); graphs a1 and b1 represent -artefacts datasets (n = 15) when: a) -synonymous and b) +synonymous. The TMB high thresholds with optimal sensitivity and specificity is shown by use of a black datapoint.



Variability in ROC-curve generated thresholds in this study:

Different optimal TMB thresholds were identified by ROC curve analysis depending on the panel/analysis parameters used (table 18). The optimal ROC-curve generated thresholds for Nonacus varied between 80 variants/Mb and 240 variants/Mb, whilst the Illumina thresholds varied in the range 100-130 variants/Mb. This threshold variation relates to the variation in TMB scores observed using each set of analysis parameters (+/-synonymous variants and +/-artefacts). The optimal TMB high thresholds fluctuate in line with the mean TMB score for the patient cohort; so, the lowest mean TMB score (across the four different analyses) for the Illumina cohort was 144 variants/Mb using the -synonymous -artefact analysis, and these analysis settings

also had the lowest TMB threshold for Illumina. Equally, the highest mean TMB score for the Nonacus cohort was 244 variants/Mb when the +synonymous + artefact analysis was used, and this analysis package also had the highest Nonacus TMB threshold aligned to it.

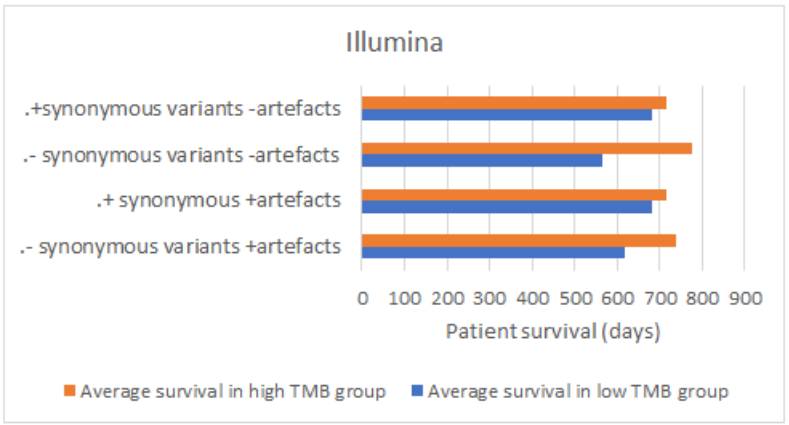
Table 18: Optimal ROC-curve generated TMB high thresholds for accurate prediction of immunotherapy response based on maximising sensitivity and specificity in this patient cohort (n = 15). Sensitivity and specificity have been determined in relation to the accuracy of responder/non-responder classification (responders defined using RECIST 1.1 criteria encompassing CR, PR and SD; Eisenhauer *et al.* 2009).

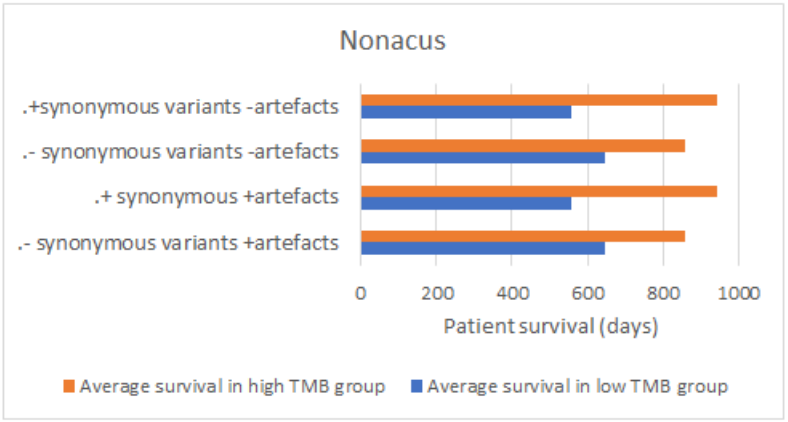
			TMB threshold (variants/Mb)	Sensitivity (TP)	Specificity (1-FP)
Illumina	**+ artefacts**	**- synonymous**	>100	83%	33%
		+ synonymous	>130	83%	33%
	- artefacts	**- synonymous**	>100	83%	44%
		+ synonymous	>120	83%	33%
Nonacus	**+ artefacts**	**- synonymous**	>130	100%	67%
		+ synonymous	>240	100%	78%
	- artefacts	**- synonymous**	>80	100%	56%
		+ synonymous	>180	100%	78%

Utility of TMB as a predictor of immunotherapy response:

Using the optimal ROC-curve generated TMB high thresholds in table 18, the utility of TMB as a biomarker for immunotherapy response prediction can be observed by comparing the assigned TMB status of the patients with the immunotherapy responses (survival post PD-L1 test) (figure 19). In all of the eight panel/analysis combinations (Illumina +/- synonymous variants +/-artefacts; Nonacus +/- synonymous variants +/- artefacts), patient survival in the high TMB group is longer than in the low TMB group. The statistical significance of this observed difference in survival between TMB high and TMB low groups is assessed in section 3.4.4.

Figure 19: Variation in survival between high TMB patients and low TMB patients in this Welsh patient cohort across each of the 8 panel/analysis combinations. Mean survival within each panel/analysis cohort (n=15) is plotted, with survival representing survival post PD-L1 test to date of death or 11/04/2022. TMB high thresholds for each of the eight panel/analysis combinations were generated using ROC curve analysis (table 18).





3.4.3.3 Impact of the use of different NGS panels and different TMB calculations on TMB status

Data already presented in the results chapter shows that analysis of the same patient sample on different NGS panels generates different values for the number of variants within the tumour sample, and in some analysis settings (-synonymous variants -artefacts) this difference in TMB score between the Nonacus and Illumina panels is statistically significant. Using the optimal ROC-curve generated thresholds described above, the impact on TMB high/low status of the NGS panel used for TMB calculation can be seen. The data shows that in 4/13 (31%) patients, there was total discrepancy between the TMB status assigned to these patients dependent on whether the Illumina or Nonacus panel was used for TMB assessment (table 19-a), but the TMB status in these cases did not change based on the TMB analysis performed. In one of these samples (20M70071), the Nonacus-generated TMB values (under all analysis conditions) were only just greater than the TMB high threshold (ROC-curve generated). In a similar way, the Illumina-generated TMB values of sample 20M70088 were close to (within 20 variants of) the TMB high threshold in all analysis settings (table 19-a).

6/13 (46%) patients had the same TMB status regardless of the panel and analysis method used. In 3/13 (23%) patients, the TMB status differed dependent on both the

panel and analysis parameters used (table 19-b). Table 19-b shows that, in the three patient samples in which there were discrepancies in TMB status dependent on the panel and analysis parameters used, many of the TMB values generated for these samples were close to (within 20 variants of) the TMB high threshold for that specific panel/analysis combination.

Table 19: Variability in TMB status when different NGS panels and analysis parameters are used for TMB calculation. Green indicates TMB high, red indicates TMB low; a * indicates that the TMB value is close to (within 20 variants of) the TMB high threshold. The RECIST-1 classification of each sample is shown in the table; a tick indicates that RECIST-1 classification (CR = complete response; PR = partial response; NR = no response; Eisenhauer *et al.* 2009) correlates with the TMB status. a) Data from the four samples in the Welsh patient cohort that had different TMB status dependent on the NGS panel used. Note: The TMB status of these four samples did not change in either panel setting when different TMB analysis parameters (+/- synonymous variants and +/-artefacts) were used. b) Data from the three samples in this study that showed variability in TMB status dependent on both the panel and analysis parameters used for TMB quantification.

a)

	20M70071 (CR)	20M70074 (NR)	20M70084 (NR)	20M70088 (PR)
Illumina				√*
Nonacus	√*	√	√	

b)

	20M70072 (NR)	20M70079 (NR)	20M70091 (NR)
Illumina - synonymous + artefacts		√	*
Illumina + synonymous + artefacts		*	√*
Illumina - synonymous – artefacts		√	√*
Illumina + synonymous – artefacts		*	√*
Nonacus - synonymous + artefacts	*		
Nonacus + synonymous + artefacts	√*		
Nonacus - synonymous – artefacts	*		*
Nonacus + synonymous – artefacts	√*		

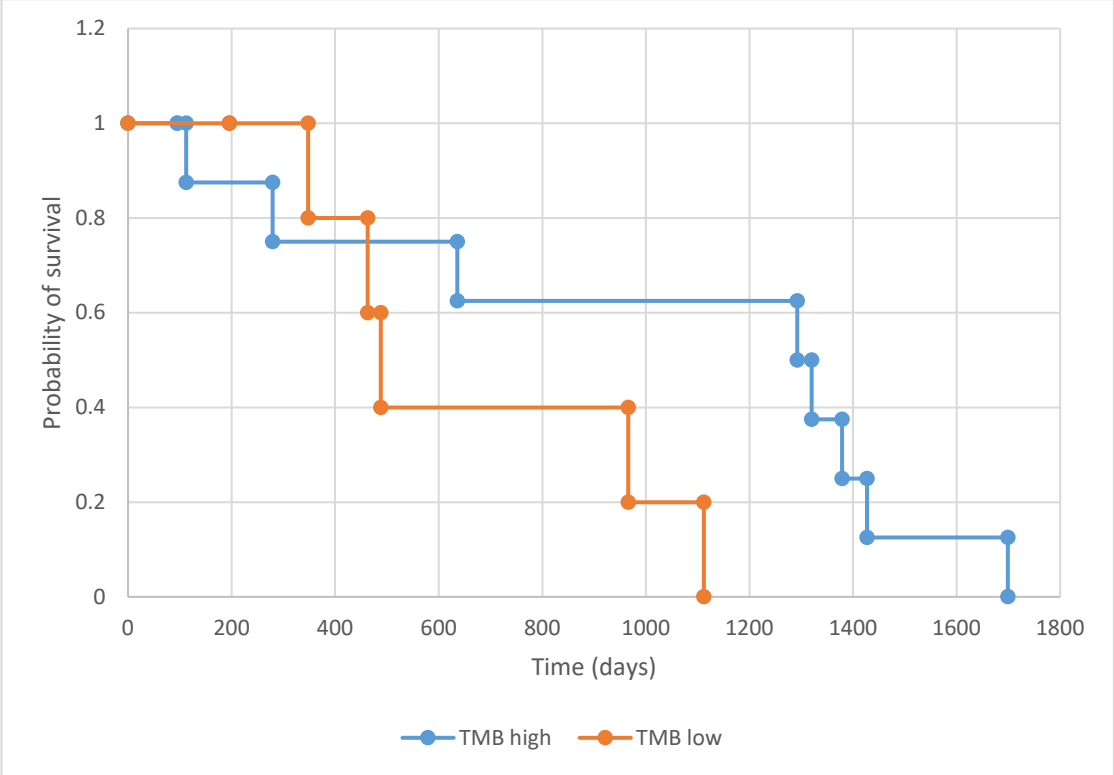
3.4.4 Association between TMB status and immunotherapy response

The weak positive correlation (r_s) between TMB score and immunotherapy response (survival post PD-L1 test) across all Nonacus and Illumina datasets has already been described earlier in the results chapter. The subsequent classification of patients into TMB high and TMB low groups based on the optimal ROC-curve generated TMB high threshold, allows the survival of patients within the TMB high and TMB low groups to be compared.

Kaplan Meier curves were used to visualise the difference in survival (calculated as survival post PD-L1 test to date of death or 11/04/2022) between the TMB high and TMB low groups of the Nonacus -synonymous variants +artefacts dataset (figure 20). This patient/analysis cohort was interrogated as, out of the eight different panel/analysis cohorts, this gave the strongest association (r_s) between TMB score and immunotherapy response.

The median survival time (probability of survival of 0.5 in figure 20), is 1292 days for the TMB high group, whilst median survival in the TMB low group is 488 days. From the Kaplan Meier curves, a log rank test was used to assesses whether the Kaplan Meier curves from these TMB high and TMB low cohorts are significantly different. The log rank (χ^2 critical value of 3.841 at 5% level of significance) showed that there was no statistically significant difference between the TMB high and TMB low survival curves (χ^2 =0.258, p=0.05).

Figure 20: Kaplan Meier curves illustrating the likelihood of survival within TMB high and TMB low cohorts. Survival represents survival post PD-L1 test to date of death or 11/04/2022. TMB estimations in this patient cohort (n=15) were performed on Nonacus - synonymous variants +artefacts. A TMB high threshold of 130 variants/Mb was used as generated from ROC curve analysis (table 18), with TMB high (n=9) and TMB low (n=6).



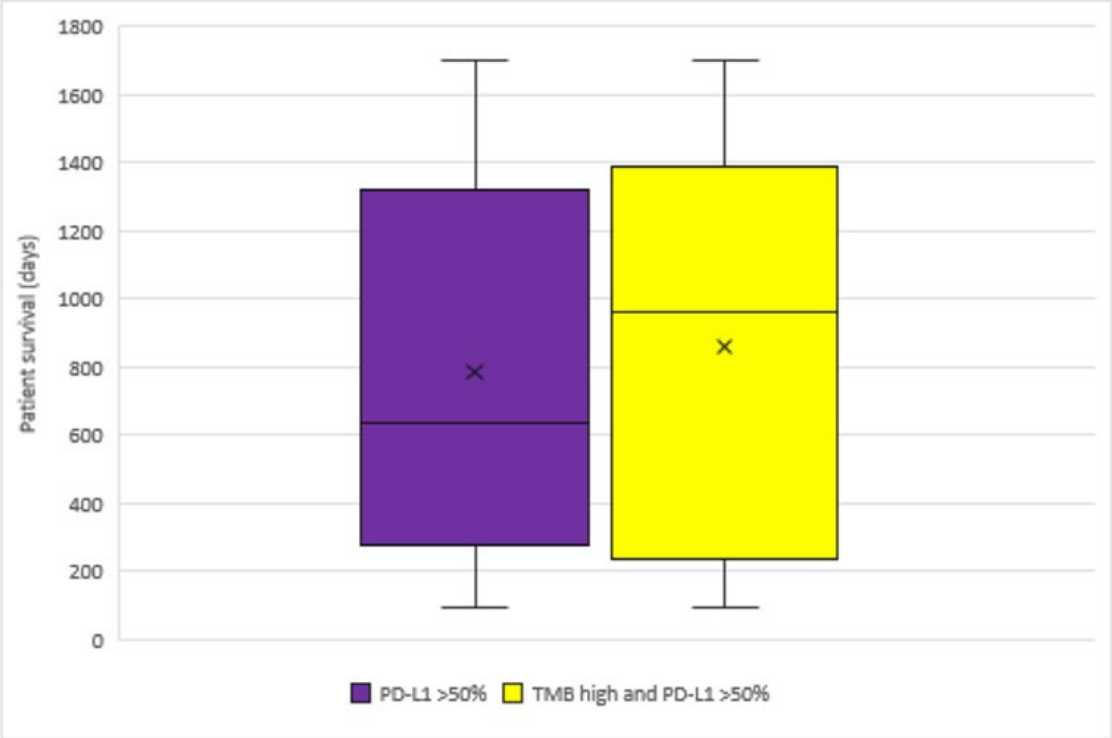
3.4.5 Evaluation of the utility of combining TMB status with PD-L1 status in prediction of immunotherapy response

Using PD-L1 expression alone to predict immunotherapy response within the total patient cohort (n=17) resulted in only 7/17 (41%) responders (encompassing RECIST 1.1 categories CR, PR and SD; Eisenhauer *et al.* 2009) within this PD-L1 high (>50%) expression group; this is a biomarker sensitivity of 41%. This percentage varies very little (range: 40-47%) when the cohort is restricted to the 13 samples analysed across both NGS panels, or reduced to either of the panel-specific 15-patient cohorts. This study data has already demonstrated that the sensitivity of TMB estimation for accurate responder/non-responder classification (responders defined using RECIST 1.1) within this PD-L1 expressor patient cohort ranges from 83% to 100%, dependent on the panel/analysis method used to determine the patient TMB scores (table 18).

To quantify and statistically analyse the combined utility of TMB and PD-L1 expression status, the survival times of the PD-L1 expressor group (n=15) were compared to the survival times of the TMB high + PD-L1 expressor group (n=10). The TMB high group evaluated is from the Nonacus -synonymous variants +artefacts dataset, which is the dataset that gave the most positive correlation between TMB value and immunotherapy response in Spearman rank analysis. Notably, to enable comparison here, the PD-L1 expressor group represented the 15 patients analysed on the Nonacus panel rather than the total 17 patient cohort in this study.

The box and whisker plots (figure 21) show that by assessing TMB value and PD-L1 expression in combination, within the patient group analysed on the Nonacus panel - synonymous variants +artefacts (n = 15), the median patient survival time is extended from 636 days (when only PD-L1 expression is evaluated) to 964 days. The mean survival time is also extended but by a smaller margin (787 days compared to 859 days) when both TMB and PD-L1 expression are evaluated. Variation in survival time was similar in both the PD-L1 high group (n=15) and the combined TMB high + PD-L1 high group (n=10). The Kruskal-Wallis test was used to determine if there was a statistically significant difference in patient survival between these two groups. The test revealed that the median patient survival did not show a statistically significant difference ($H = 0.077$, $p = 0.782$) between the PD-L1 >50% group and the TMB high + PD-L1 >50% group.

Figure 21: Box and whisker plots to visualise the effectiveness of PD-L1 high expression alone and the combined utility of TMB high status and PD-L1 high expression in predicting immunotherapy response. Immunotherapy response is demonstrated as survival, calculated as the number of days between the PD-L1 test to the date of death, or (where no date of death) until 11/04/2022 when patient was confirmed as alive by WCB. The TMB estimations in the TMB high + PD-L1 high group were generated using Nonacus -synonymous +artefacts, which is the dataset that gave the most positive correlation (r_s) between TMB value and immunotherapy response. A TMB high threshold of 130 variants/Mb) has been used as defined using ROC curve analysis (table 18). The 'x' on the plot indicates the mean patient survival for each patient group, and the median survival of each group is shown with a solid line. PD-L1 high group (n=15) and the combined TMB high + PD-L1 high group (n=10).



3.5 Evaluation of the feasibility of TMB within a clinical setting

3.5.1 Cost of a TMB clinical service

The calculated costings of a TMB service within AWMGS is shown in table 20, using the two NGS panels that provided successful sequencing data. These costings do not consider the additional costs that would be associated with validation of a TMB service, which would involve the determination of sensitivity, specificity, and reproducibility of the TMB assay using a range of clinical-grade samples over multiple NGS runs. Information provided from the Pathology laboratory at University Hospital of Wales (personal correspondence), where PD-L1 assessment by IHC is currently performed for NSCLC samples, noted that the cost of PD-L1 analysis is £100 per patient.

It is important to consider these costs alongside the panel's gene contents, which extends the utility of these panels to encompass both TMB evaluation and the delivery of existing NSCLC WHSCC-funded services (section 1.7.2).

Table 20: Estimated cost of a TMB service within the AWMGS laboratory.

Item	Additional information	Estimated cost/tumour sample
DNA extraction	Maxwell extraction	£12
Library preparation: Illumina TruSight™ Oncology 500 panel	Illumina TruSight™ Oncology 500 panel DNA kit for 48 samples = £8758.20	£182.46
Library preparation: Nonacus Cell3™ Target: Pan Cancer panel	Nonacus Cell3™ Target: Pan Cancer panel for 16 samples = £1389	£86.81
Sequencing: NovaSeq	NovaSeq Xp 2-Lane Manifold Pack (£1071.60), NovaSeq Xp 2-Lane Kit (£461.70), NovaSeq 6000 SP Reagent Kit (300 cycle) (£5371.56) for a total of 48 samples	£143.85
Other consumables	Target Pure NGS clean-up beads, 10ml (£157.50), UK Delivery, Dry Ice (£38), Beckman - AMPure XP DNA cleanup kit (60ml) (£975), Oxford Gene Technology- Dynabeads™ M270 Streptavidin, 2ml (£276) for a total of 48 samples	£30.13
Staffing	Based on an average cost dependent on staff grade of 42p per minute	£21.38
TOTAL for Illumina targeting TOTAL for Nonacus targeting		£389.82 £294.17

3.5.2 Availability of External Quality Assurance schemes for a TMB clinical service

A search of the NEQAS and EMQN websites showed that no TMB-based EQA scheme run by either of these bodies exists (website search performed on 31st October 2022). However, in 2021, EMQN collaborated with the International Quality Network for Pathology (IQNPath) on a project to assess the standardisation of TMB testing, which included running a pilot EQA scheme for TMB analysis in 2020 (Abate 2020). The results from this pilot scheme, reporting data from 23 participating laboratories, have recently been published (Abate *et al.* 2022).

The pilot results showed that there were seven labs using one specific NGS panel (Oncomine Tumour Mutation Load panel), but within this group different TMB calculation methods were used (+/- synonymous), which resulted in different TMB estimations between sites (Abate *et al.* 2022). The average TMB scores for each patient were 2-5 variants/Mb greater under +synonymous conditions than the average -synonymous values (Abate *et al.* 2022). The average increase in TMB score when synonymous variants were included varied between 15% and 47% across the five samples analysed (Abate *et al.* 2022). The maximum increase in TMB score of a single

sample between sites was 99% (relates to a sample with TMB -synonymous = 8.04 variants/Mb in one lab, and TMB +synonymous = 15.99 variants/Mb in another lab). The calculations +/- synonymous were not performed in the same laboratory in this EQA scheme, therefore there could be additional variation in these TMB scores owing to additional differences in the calculations performed.

Chapter 4: Discussion

The use of the immunotherapy, pembrolizumab, is stratified in current NHS practice in the UK using IHC-measured PD-L1 expression levels in NSCLC tumours (NICE, 2016a) to target the patients most likely to respond to this treatment. Tumour cells expressing high levels of PD-L1 (>50%) will generally have a better response to pembrolizumab (Topalian *et al.* 2012; Taube *et al.* 2014); however, technical complexities regarding PD-L1 expression analysis means that this biomarker is imperfect. TMB has been identified as both an alternative and complementary biomarker to PD-L1 expression analysis in NSCLC patients by better stratifying patients into responder and non-responder groups (Kowantz *et al.* 2017; Carbone *et al.* 2017; Hellmann *et al.* 2018b; Hellmann *et al.* 2018a; Marabelle *et al.* 2020). However, there is no consensus or best practice guidance associated with TMB evaluation in terms of which method to use for quantification of TMB, which variants to include in the TMB estimations, and what TMB high threshold to use to accurately define immunotherapy responders and non-responders. This means that there is a huge range of analyses used within TMB publications, which vary in these three elements (panel, analysis, and threshold). With the first FDA-approval of a TMB-stratified solid tumour service for pembrolizumab use having been issued in June 2020 (Marcus *et al.* 2021), the potential for NICE-approval of a TMB-based service is heightened. Such NICE approval would require TMB-based stratification to be delivered within NHS Genomics laboratories such as AWMGS in a timeframe of 60 days from approval. Prior to such service delivery being possible, guidance is required in terms of detailing recommendations for TMB analysis to enable clinically appropriate validations of TMB services to be performed within NHS Genomics laboratories.

4.1 Answering the primary research question

This study was performed at the AWMGS, which is the laboratory that would be responsible for implementing a clinical TMB service within NHS Wales. This study was initiated with the aim of answering the primary research question: Does TMB in combination with PD-L1 expression analysis have clinical utility as a biomarker for anti-PD-L1 immunotherapy treatment response in a Welsh lung cancer patient cohort? This is the first question to address prior to establishing a TMB service in AWMGS. There are a number of key findings in this study, namely: improved survival in TMB high + PD-L1 high patients, and improved accuracy of immunotherapy response prediction using TMB assessment alongside PD-L1 expression data, as well as the potential cost neutral status of a TMB service within the NHS, that potentially support the clinical utility of TMB and PD-L1 expression analysis as a combined biomarker within this setting. The limited cohort size, along with the absence of the control of variables (other than PD-L1 status) that could impact on TMB values within this cohort, reduces the ability to generalise these findings across the Welsh lung cancer population as a whole. Having noted the potential clinical utility of a combined TMB and PD-L1 biomarker, this research reports variation in TMB quantification based on the choice of panel and analysis performed, and demonstrates the clinical significance of this variation using different TMB high thresholds; these findings diminish the clinical utility of TMB as a biomarker, but this study showed that utility could be improved by the use of an intermediate TMB category (see later). The development of best practice guidance

regarding TMB assessment and an effective EQA scheme for TMB quantification would also both improve the clinical utility of TMB as a biomarker, both of which are discussed herein.

The key study findings that support the clinical utility of a combined TMB and PD-L1 expression biomarker are:

1. A weak positive correlation (r_s) between TMB score and immunotherapy response (measured in terms of survival time post PD-L1 test) in the PD-L1 high expressor patient cohort evaluated in this study. This correlation was observed in all NGS panel/TMB analysis combinations investigated (r_s range=0.007-0.154 dependent on panel/analysis).
2. In all NGS panel/TMB analysis combinations, mean patient survival (measured as survival post PD-L1 test) in the TMB high + PD-L1 high (>50%) group was extended compared to the mean survival of the TMB low + PD-L1 high (>50%) group when using ROC-curve generated TMB high thresholds; this difference in survival did not meet statistical significance within the single panel/analysis cohort evaluated (χ^2=0.258, p=0.05).
3. The combined use of TMB status (using ROC-curve generated TMB high thresholds) and PD-L1 expression analysis improved the sensitivity of responder/non-responder classification (responders defined using RECIST 1.1 criteria encompassing CR, PR and SD; Eisenhauer *et al.* 2009) from 41% when PD-L1 assessment alone was used, to 78-100% sensitivity in a combined biomarker approach; notably, sensitivity varied dependent on the panel/analysis combination.
4. The combined use of TMB estimation and PD-L1 expression analysis (TMB high + PD-L1 >50%) increased the median patient survival time of this predicted responder group compared to the use of PD-L1 high alone, although this survival difference lacked statistical significance (H=0.077, p=0.782)
5. Based on the three targeted NGS panels evaluated, the implementation of a clinical TMB service for NSCLC patients would be cost neutral using the Nonacus and Ilumina panels. This cost neutral status was based on these NGS panels targeting all NSCLC clinical actionable genes (NICE, 2022), namely *EGFR, KRAS, BRAF, ALK, ROS1, RET, NTRK1, NTRK2, NTRK3*, for which testing is already funded across the NHS, as well as the inclusion of the *MET* gene within these panels, which is an additional NSCLC-relevant gene noted in the NHSE Cancer Test Directory (England.nhs.uk, 2022).

Finding number 3 above relating to the improved sensitivity of responder/non-responder classification seen in this study when a combination of TMB and PD-L1 assessment is performed, confirms the thesis hypothesis by demonstrating that TMB status in combination with PD-L1 expression data can act as an anti-PD-L1 immunotherapy treatment response biomarker, by accurately stratifying patients in this Welsh lung cancer patient cohort into responder and non-responder groups. This emulates the findings of other research publications (Carbone *et al.* 2017; Peters *et al.* 2017; Seiwert *et al.* 2018), although the thesis findings and the published data regarding combined TMB and PD-L1 biomarker utility lack statistical significance, showing that further research studies and trials are required to demonstrate such utility prior to a clinical service based on a combined biomarker being considered.

The findings in this thesis that suggest the requirement for guidance regarding TMB assessment prior to TMB realising its potential as a biomarker in combination with PD-L1 expression analysis within a clinical environment are:

1. The observed lack of consistency in TMB score and TMB status (using ROC-curve generated TMB high thresholds) when the same set of 13 samples was analysed across two targeted NGS panels, Illumina and Nonacus, which would be of clinical significance if TMB was being used as a biomarker of immunotherapy response. This variation in TMB score dependent on panel used reflects the findings of the 2021 EMQN/IQNPath pilot EQA scheme for TMB assessment (Abate *et al.* 2022).
2. The observed variation in TMB score and TMB status (using ROC-curve generated TMB high thresholds) when TMB scores were calculated for the same set of 13 samples using different variant inclusion criteria, i.e. +/-synonymous variants and +/-artefacts, which (as for point number 1 above) would be of clinical significance in a clinical setting by altering the patients predicted to respond to immunotherapy. Again, such variation in TMB score dependent on the inclusion/exclusion of synonymous variants was also suggested from the 2021 EMQN/IQNPath TMB pilot EQA scheme findings (Abate *et al.* 2022).
3. The observed variation in the TMB high threshold that generated the greatest sensitivity and specificity of responder/non-responder separation both when the same panel was used for TMB estimation using different analysis parameters, and when different panels were used for TMB quantification. Variation in TMB high thresholds dependent on the panel used has been noted in the literature (table 4).

All of these key findings, in support of the clinical utility of TMB as a biomarker of immunotherapy response and those that represent potential barriers to its clinical utility, are discussed in more detail in section 4.2 below as the aims of this thesis are addressed. Each aim is addressed separately, with the final aim of the evaluation of feasibility of a TMB service in the NHS considering not just the cost of delivering TMB testing and the availability of EQA schemes, but also bringing together key findings of the other aims that impact on the feasibility of an NHS TMB clinical service.

4.2 Fulfilling the aims of this thesis

4.2.1 Study aim 1: Implications of targeted panel size and gene content on TMB scores

The Illumina and Nonacus panels both target >1.5Mb of the genome, which is the size noted by Buchhalter *et al.* (2019) as being essential for accuracy of TMB measurement based on the similarity of TMB values generated by WES and simulated panels >1.5Mb in size. However, the Nonacus panel is below the 1.6Mb size threshold suggested by Hatakeyama *et al.* (2018) as being the minimum panel size for accurate TMB quantification, again in terms of accuracy to WES-generated data. As the samples in this patient cohort had not been analysed for TMB prior to this study, the TMB values

of these samples was unknown, and therefore accuracy of TMB assessment, in terms of accuracy of panel-based TMB scoring compared to WES, cannot be directly assessed. However, TMB estimations from the Illumina TruSight™ Oncology 500 panel evaluated in this thesis have been shown to be concordant with those generated from WGS data (Pestinger *et al.* 2020). As this thesis data shows that TMB estimations from the Illumina and Nonacus panels differ (which was a statistically significant difference in one panel/analysis setting), by extrapolation, so Nonacus and WGS TMB estimates would also be predicted to differ. Therefore, this study predicts that the Nonacus panel, targeting <1.6Mb of the genome, does not accurately measure TMB in terms of lack of consistent TMB values generated from Nonacus and the proven WGS-concordant Illumina approach (Pestinger *et al.* 2020), of course this is assuming that the Illumina panel in this study is performing in a representative manner to WGS in this setting. This finding supports the suggestion made by Hatakeyama *et al.* (2018) of a 1.6Mb minimum panel size for accuracy of TMB assessment compared to gold-standard WES/WGS. However, based on the differences in the genes and gene regions (in terms of hotspots vs wider gene screening) targeted by different NGS panels (tables 2 and 5), extrapolating the lack of predicted accuracy in TMB estimation from this single Nonacus panel to all panels <1.6Mb is not appropriate. The Hatakeyama *et al.* (2018) study was also based on a single panel <1.6Mb in size.

Given that WGS/WES is considered to be the gold standard for TMB quantification, one could expect that there would be an improved clinical utility of TMB analysis using the Illumina panel compared to estimation using the Nonacus panel, based on the demonstrated concordance of Illumina-generated TMB scores with WGS-generated values (Pestinger *et al.* 2020). However, in this study it is the Nonacus panel that out-performs the Illumina panel in many areas that demonstrate potential improved clinical utility of the Nonacus panel, in terms of the Nonacus panel showing: the strongest correlation between TMB score and patient survival (Illumina r_s range: 0.007-0.011; Nonacus r_s range: 0.046-0.154); improved sensitivity and specificity for responder/non-responder classification (responders defined using RECIST 1.1 criteria; Eisenhauer *et al.* 2009); the greatest difference in survival between TMB high and TMB low groups (ROC-curve generated thresholds). These aspects of the Nonacus panel (1.58Mb) performance within this study support the clinical utility of this panel for TMB estimation, although with the caveat that these study findings are based on the evaluation of a small patient cohort. This conclusion does not mean that all panels >1.5Mb would necessarily demonstrate clinical utility for immunotherapy response prediction, as this utility will be impacted by factors other than panel size (including panel gene content and TMB threshold, as discussed later). No comment can be made regarding the utility of panels <1.5Mb for TMB estimation as no panels of this size were evaluated.

This data has raised an interesting discussion point regarding whether it is appropriate to consider the utility of a targeted panel for TMB assessment in terms of the concordance of TMB values with WES/WGS. The data suggests that a panel (i.e. Nonacus) could generate inaccurate TMB estimations (in relation to lack of concordance with WES/WGS) and yet these values could still demonstrate clinical utility in predicting immunotherapy response. Based on the discrepancy between proposed lack of WGS-concordance and observed clinical utility having been identified within this small Welsh patient cohort, there is no recommendation to change the minimum panel size requirement for TMB estimation from 1.6Mb (ensuring

consistency with WES-generated TMB values). This finding of an apparent discrepancy between WES correlation and clinical utility could be a factor of the imperfect nature of the comparison of Nonacus TMB estimations to WGS (i.e. via the Illumina intermediator), small study size, or could be linked to the multi-factorial nature of immunotherapy response which could impact on the responses demonstrated by the patients in this thesis. Alternatively, WGS/WES concordance may not always align with clinical utility; this would have to be evaluated in a larger study.

The conclusion regarding the proposed clinical utility of Nonacus is based on the analysis of only 15 patient samples, so the Illumina panel should not be excluded as a potential useful panel for TMB evaluation on the basis of these results, particularly as the utility of the Illumina panel in TMB assessment has been demonstrated in another study (Pestinger *et al.* 2020). One important aspect of panel utility to consider is the sequencing coverage achieved across the genes of the panel, and it was the Illumina panel that delivered better sequencing coverage in this Welsh study as discussed below.

The differences in TMB estimations from the Nonacus and Illumina targeted NGS panel evaluation of the same Welsh patient cohort (n=13) when the same TMB calculation parameters and method are used, supports the suggestion made by Budczies *et al.* (2019), that variation in TMB estimation would be expected between different targeted panels of differing size and gene content. Given that the panels differ in both size (Illumina: 1.94Mb; Nonacus: 1.58Mb) and gene content (68% of the genes targeted are the same in both panels; table 6), it is not possible to determine the contribution of each of these elements to the differences in TMB scores observed. The difference in TMB scores between these two panels could be a function of one or both of these variables. In this context, the fact that the Agilent NGS panel (1.7Mb, 174 genes) failed to generate any sequencing data over the minimum coverage threshold (250x) is disappointing, as this data could have provided more information regarding the impact of gene content on TMB quantification. The Agilent panel shares a large proportion of its gene content with the Illumina and Nonacus panels (95% and 87% respectively); however, the Illumina and Nonacus panels evaluate additional mutation hotspots of around 350 genes on top of this shared gene content (table 6). It would therefore have been interesting to observe the impact of the additional gene content of the Illumina and Nonacus panels on TMB quantification by comparing Illumina and Nonacus-generated TMB scores to those from Agilent. Although again, the differences in size of these three panels could also contribute to any variation in TMB score. However, ultimately, this study confirms that gene panel size and/or the gene panel content does impact on TMB quantification. The impact that this TMB score variation has clinically would be dependent on the impact on TMB status, which in turn is dependent on the use of appropriately validated TMB thresholds; these elements are discussed later.

Although the gene content of the Illumina and Nonacus panels is different, the panels target the majority of the most highly mutated genes in lung cancer (Mycancergenome.org, 2022a); however notably, the Illumina panel targets the *KMT2D* gene whilst the Nonacus panel does not (table 6). Variants in the *KMT2D* gene occur in 10% of NSCLCs (Mycancergenome.org, 2022b); therefore, the absence of this gene alone from the Nonacus panel highlights how gene panel content could alter the numbers of variants identified in a sample. Both panels target the hotspot regions of

the *KRAS* and *EGFR* genes (table 6), which are where the majority of oncogenic lung cancer driver mutations occur (Chevallier *et al.* 2021). This maximises the likelihood of variant detection in these commonly mutated genes when either of the Illumina or Nonacus NGS panels is used. However, sequencing coverage across the panel as a whole, and specifically coverage of the *KRAS* and *EGFR* hotspot regions, within the Welsh patient cohort was worse when using the Nonacus panel compared to the Illumina panel. When a minimum coverage of 250x is achieved, variants at 5% frequency can be detected (Petrackova *et al.* 2019). The poorer sequencing coverage of the Nonacus targeted samples is likely to result in a reduced sensitivity for detection of variants, particularly low-level variants (e.g. 5%) that are common in heterogeneous FFPE samples. It is important to note that the calculation of TMB scores within this thesis requires the number of variants identified to be divided by the size of the genome (in Mb) interrogated. Notably, when gene regions have poor sequence coverage they will still be counted as interrogated regions in the context of the generation of a TMB score, but the sensitivity of variant detection within these regions is reduced; this could result in an under-estimation of TMB scores in poorly covered samples, such as those samples that have been Nonacus-targeted. This finding suggests the need for a minimum panel coverage requirement to ensure TMB would not be under-estimated in a sample; this quality threshold would have to be evaluated within a panel validation setting. No such quality threshold was used in this study based on the lack of understanding regarding what this threshold would be, along with the added requirement to maximise the dataset in this already small patient cohort. Despite the lower coverage of samples analysed using the Nonacus panel and the lack of *KMT2D* gene targeting in the Nonacus panel, the number of variants detected by this panel was higher in 46% (6/13) of patients in the -synonymous +artefacts dataset and in 85% (11/13) of patients in the +synonymous +artefacts (compared to the same patients targeted using the Illumina panel); the number of variants detected was also higher in around one third of the other Nonacus datasets (+synonymous -artefacts and -synonymous -artefacts). Other factors could be influencing the number of variants detected by Nonacus including: the increased presence of sequencing artefacts in the Nonacus samples compared to the Illumina samples, and the other differences in genes targeted by the Illumina and Nonacus panels beyond the lung cancer-relevant genes, which are both discussed in more detail in the next section.

Importantly, since conceiving and initiating this thesis, a number of research papers have been published investigating the effect of NGS panel selection on TMB quantification, including two publications focusing on NSCLC (Heeke *et al.* 2020; Ramos-Paradas *et al.* 2021; Vega *et al.* 2021). These papers represent the first published datasets involving the analysis of the same samples on more than one NGS panel. The results from this thesis, in relation to there being variation in TMB score dependent on the panel used to analyse the sample, are concordant with these publications. Panel-dependent TMB variation was also noted in the TMB-focussed pilot EQA scheme in 2020 (Abate *et al.* 2022), which was the first EQA scheme to evaluate TMB estimation. Dependent on the TMB high threshold used, these differences in TMB quantification could have an impact on the treatment options available to the patient if TMB was used as a biomarker of immunotherapy response in a clinical service; this will be discussed in more detail later in relation to the findings of this thesis. Variation in TMB score as a consequence of the panel used for determination of this value means that TMB data is not directly comparable between labs using different NGS

panels; this factor will be discussed later in relation to the establishment of TMB-based EQA schemes.

4.2.2 Study aim 2: Impact of variant number calculations on TMB scores

In the context of a TMB-based clinical service, the variation in TMB score based on the TMB calculation performed (considering both +/-synonymous and +/-artefacts), as demonstrated in this thesis and by Abate *et al.* (2022), could impact on the clinical decisions made, as well as meaning that TMB estimations from different labs are not directly comparable if different TMB calculations have been used. The same issues have been highlighted already in relation to the use of different panels for TMB estimation.

The variation in TMB scores generated from a single panel using +/-synonymous and +/-artefact calculations strongly suggests that a single TMB high threshold with utility independent of analysis parameters would be difficult to find. Indeed, this statement is corroborated by the ROC-curve generated TMB high thresholds within this study, which varied according to analysis method as well as panel used (discussed later).

4.2.2.1 Impact of synonymous variants on TMB scores

The number of variants detected per patient within each panel cohort increased by an average of 77% for Nonacus-targeted samples and 35% for Illumina-targeted samples when synonymous variants were counted (artefacts included), which resulted in the same percentage increases in TMB scores (variants/Mb). These differences are comparable to the findings of the 2021 EMQN/IQNPath pilot EQA scheme (Abate *et al.* 2022) when the same five samples were analysed using the same targeted NGS panel (Oncomine Tumour Mutation Load) in seven different laboratories using +/-synonymous calculations. The thesis and EQA results demonstrate that it would be critical for a laboratory providing a TMB clinical service to maintain consistency in the TMB calculation method used within a live service to ensure consistent results in each TMB assessment performed, as TMB quantification of a single patient sample could yield different results dependent on the variants included within the TMB estimation. Such consistency in methodology is standard practice in an NHS Genomics laboratory setting where all processes are heavily standardised to ensure consistent high-quality results are produced. The TMB calculation used within a clinical lab setting would have to be validated in accordance with validation guidelines, as is standard practice in UK Genomics labs (acgs.uk.com, 2022), and would have to have demonstrated clinical utility for TMB assessment for provision of a clinical service in terms of the values generated by the panel/analysis combination providing identification of immunotherapy responders.

Perhaps the most obvious reason for the general increase in the number of synonymous variants within the Nonacus-targeted patient samples, and the inflated number of variants in general seen in the Nonacus-sequenced patients (inflated in terms of poor coverage observed), could be the difference in gene content of the panels as mentioned in the comparison of panel performance in section 4.2.1. As noted previously, the Nonacus panel does not target any key lung cancer-mutated genes in addition to those targeted by the Illumina panel, so this does not explain the increased variant prevalence of the Nonacus-targeted samples. However, additional genes within the panels have not been scrutinised for their potential mutable-potential

either in cancer cells or in 'normal' cells, and any differences here could account for the increase in variant prevalence observed in Nonacus-targeted samples.

Of course, the sequencing coverage across the panel could be influencing the variants detected per patient, with good coverage of hotspot mutation regions resulting in a better variant pick-up rate. The lower sequencing coverage of a selection of lung cancer related genes (*CDKN2A, PTEN, NRAS, RET, KRAS, ERBB2, PIK3CA, EGFR, MET, BRAF*) on the Nonacus panel would suggest a decreased sensitivity for variant detection within these genes within the Nonacus dataset, and does not support the possibility that increased synonymous variant detection in the Nonacus panel is due to increased NSCLC-related variant pick-up rate.

One explanation for the particularly high increase (77%) in variant calls per patient within the Nonacus +synonymous +artefacts dataset, could be that there is a larger number of synonymous artefacts within the Nonacus dataset which is inflating the synonymous variant counts in the Nonacus-targeted samples. The predicted percentage of artefacts in the Nonacus +synonymous dataset is 25% (data generated in this thesis; table 14), whilst within the lllumina +synonymous dataset the predicted artefact prevalence is 2-3% (table 14), therefore these artefact prevalence estimations support this explanation. The reduced Q30 Phred score of the Nonacus targeted dataset supports an increased presence of artefacts within the Nonacus data, which is discussed further in the next section.

Based on the discussion above, the 42% increase (77% Nonacus compared to 35% Illumina) in synonymous variants detected in this Welsh patient cohort via Nonacus sequencing compared to Illumina sequencing can be attributed in part to the predicted increased artefact frequency (25-42%; table 14) when the Nonacus panel is used, which can be linked to the lower quality of the Nonacus sequencing (Q30 score). Other factors, other than artefact prevalence, are likely to be influencing synonymous variant frequency as the artefact estimations in this study are likely to be over-estimations (discussed in section 4.3.2), so the true artefact prevalence alone is unlikely to account for the 42% difference in synonymous variant frequency seen between the Nonacus and Illumina targeted samples. The gene content of the panel is also likely to be influencing variant prevalence as already noted.

Based on the estimated prevalence of synonymous variants in the genome of 10,000 variants per 3200Mb genome (0.0003%) (Zeng and Bromberg 2019), the number of synonymous variants in each of the patients analysed in this study would be expected to be extremely low (<1 variant per patient) based on the 1.5-1.9Mb panels used. In reality, 12-116 synonymous variants were detected per patient in the Illumina cohort (n=15) and 25-184 in the Nonacus-targeted patients (n=15). The key factor in the grossly different variant frequencies in this study and those predicted by Zeng and Bromberg (2019) is that the publication data is based on a 'normal' genome rather than a cancer genome. Cancer is a result of the accumulation of genetic variants within a cell and cancer genomes are highly mutable, which reflects the high number of synonymous variants identified in the samples of this study. Added to this, is the fact that the targeted sequencing performed in this study is skewed towards genes that are highly mutable in cancers with both the Illumina and Nonacus panels having been designed for use in the interrogation of tumour samples. It is reassuring to note that the synonymous variant prevalence within the EMQN/IQNPath 2021 pilot EQA scheme can be estimated to be around 2-5 variants/Mb (Abate *et al.* 2022), which would

equate to the detection of 4-10 variants in a 1.9Mb panel, which is at the lower end of the synonymous variant range in the 1.9Mb Illumina panel data of this study. The prevalence of synonymous variants in the Nonacus panel is greater than would be expected based on the pilot scheme data (3-8 variants expected in a 1.5Mb panel), and the reasons for this are likely to echo those described above in relation to the estimated high prevalence of artefacts in the Nonacus dataset.

4.2.2.2 Impact of sequencing artefacts on TMB scores

The number of predicted variants representing artefacts within the 15-patient dataset was calculated within this study to be much higher for the Nonacus dataset (25-42%) than the predicted number of artefacts within the Illumina panel sequencing dataset (2-3%); percentages vary based on +/- synonymous variants (table 14). There are no clear published estimations of artefact prevalence in FFPE samples. Notably, the artefact frequencies predicted in this study could be impacted by a potential bias in the prediction calculations caused by the IGV interrogation of a larger number of variants from the Nonacus panel dataset than from the Illumina dataset. This uneven interrogation of shared variants across the two panel cohorts is discussed in section 4.3.2.

The artefacts identified could be either sample-specific artefacts (e.g. C>T/G>A deamination artefacts) or panel-specific artefacts (e.g. PCR-induced artefacts or read-end artefacts). Sample-specific artefacts would be expected to occur at similar frequencies regardless of the panel that was used for analysis. As 13/15 samples analysed on each panel were the same, the impact of sample-specific artefacts would be limited to the four samples analysed on only one of the panels. Evaluation of the ratio of absorbance at 260:280nm of the four samples run only on a single panel demonstrated that the samples were of similar quality using this metric, with the exception of one predicted poor-quality sample (20M70089) analysed only on the Illumina panel. Unfortunately, at the time of this study, AWMGS did not have access to a Bioanalyzer, which would have provided more information regarding the quality of the samples used in this study in terms of level of DNA fragmentation, as this method measures the size of DNA molecules. Fragmentation in FFPE samples is associated with a higher level of artefacts (Wong *et al.*, 2014), so Bioanalyzer data could have been aligned to sample-specific artefact prevalence within the samples in this study, and specifically could have identified if any differences in artefact prevalence would have been expected between the four samples analysed on only a single panel. Based on the 260:280 metric alone, the likelihood of sample quality differences being a causative mechanism for the increased artefacts in the Nonacus datasets (+/- synonymous variants) is reduced.

The use of the same Illumina NovaSeq™ sequencing protocol and sequencing analysis pipeline rules these out as potential sources of variation in the generation of artefacts. The question is therefore whether the Nonacus targeted approach could be more prone to introducing sequencing artefacts into the dataset than the Illumina TruSight™ Oncology 500 method. One piece of data that supports the answer to this question being yes (at least within the confines of this study) is the difference between the Q30 Phred scores from each panel dataset. This quality score for the Nonacus dataset was lower than the Q30 value for the Illumina dataset, indicating that there is a greater likelihood of there being sequencing errors within the Nonacus-generated data. These sequencing errors would account for an increase in artefacts in the Nonacus dataset,

although it is not clear to what degree this would impact on artefact prevalence. It is also important to note here that the AWMGS staff who supported in the delivery of the NGS workflows in this thesis, are more familiar with the Illumina NGS panel and workflow than the Nonacus workflow, therefore this lack of experience in the use of the Nonacus panel could have influenced the poorer quality of results obtained from this panel. Also, owing to the COVID pandemic, training and support from the Nonacus technical support teams was limited to a virtual forum; this is also true for the Agilent panel which failed to generate any sequencing data. This could have impacted on the panel performance in terms of there being a potential increased risk of set-up errors in the unfamiliar library preparation processes. It is possible that improvement in coverage could be obtained through the optimisation of the targeting workflows within the AWMGS laboratory, or through additional training from specialist technical support.

Another potential source of artefacts in the datasets comes from the PCR step that occurs within the library preparation protocols, which can introduce DNA polymerase errors into the sequencing. The number of PCR cycles is limited in both the Illumina and Nonacus NGS protocols to avoid excess PCR duplicates that can lead to false-positive sequencing errors, but notably the Illumina protocol had a higher number of cycles than the Nonacus library preparation protocol. This fact effectively rules out this PCR step as a potential cause of the inflated artefact prevalence in the Nonacus dataset.

Finally, the gene panel content could influence artefact prevalence as some genes can be more prone to artefacts owing to their sequence make-up; for example, an excess of repetitive sequences within a gene would make them prone to artefacts. Owing to limited AWMGS experience in sequencing the majority of genes within the Illumina and Nonacus panels, the laboratory has no prior knowledge of the artefact prevalence in the panel gene sets, therefore the degree to which artefact-prone genes may be influencing the results of this study is unknown.

Artefact rate was shown to be higher in the datasets excluding synonymous variants than in the datasets including synonymous variants, however, there is no biological reason for such a difference as artefacts randomly occur across the genome regardless of their impact at the amino acid level (i.e. synonymous vs non-synonymous). For this reason, a potential improvement to the artefact algorithm used could have been to take an average of the predicted artefact frequencies across the two panel datasets (+/- synonymous variants). This would have had very little impact on the Illumina dataset as this would have been an average of 2% and 3%, but in the Nonacus dataset a predicted artefact level of 34% (average of 42% and 25%) could have had an impact on the TMB estimations and TMB scores generated.

4.2.2.3 Statistical vs clinical significance of differences in TMB scores

In three out of the four analysis conditions used (+/-synonymous +artefacts; +synonymous -artefacts), the Nonacus and Illumina patient cohorts gave TMB estimations that were different, but not significantly different according to the t-tests performed. In the other analysis group (-synonymous -artefacts) a t-test statistical difference ($t = 2.50$, $p = 0.028$) was observed between the TMB scores generated by the Nonacus and Illumina panels. Under these conditions (-synonymous -artefacts), the

large difference in mean TMB scores between the panel cohorts (Illumina mean TMB: 144 variants/Mb; Nonacus mean: 85 variants/Mb) can be explained by the combined impact of the reduction of the Nonacus TMB scores by the removal of suspected artefacts (predicted artefact frequency of 25-42% determined within this study), as well as the exclusion of synonymous variants which were much more prevalent in Nonacus targeted sequencing compared to Illumina targeted sequencing.

The implications for identifying a statistical significance in the difference between TMB estimations within even a single dataset, is that this work highlights the variation that is possible in TMB scores dependent on how the scores are generated, with both NGS panel choice and TMB analysis conditions (in terms of variants counted in the TMB estimation) impacting on the score. Having said this, although the differences in TMB estimations were deemed to only be statistically significant in one out of the four analyses performed, the fact that the scores generated were different dependent on the panel used could be clinically significant within a TMB clinical service setting. Indeed, within this study when ROC-curve generated TMB high thresholds were used (discussed in more detail later), four patients were determined to be TMB high by one panel and TMB low by another panel (across all four analysis parameter combinations; table 18). This total discrepancy in TMB status would have huge consequences if TMB was being used as a biomarker in the clinical setting, as would mean the difference between accessing immunotherapy or having other standard of care treatment. Variation in patient TMB status dependent on the NGS panel used is a result of the variation in TMB estimations between panels and the TMB high threshold used. As the true TMB status of the samples in this study is not known (no TMB testing previously performed), determining which the most accurate analysis parameters are, in terms of the accuracy of the TMB status assigned to each of these four discrepant samples, is not possible by such a comparison. However, accuracy of the TMB status can be determined by correlation of TMB status to drug response. The Nonacus panel-generated immunotherapy response predictions correlated with the RECIST-1 classification in 3/4 of these samples (table 14) indicating a potential improved utility of this panel over the Illumina panel within this cohort, although notably the sample numbers are extremely low. The sensitivity and specificity of the Nonacus and Illumina panels for response prediction across the whole cohort is discussed later (section 4.2.3.2); this extended analysis also supports the improved utility of the Nonacus panel.

The differences in all TMB values across the cohort of 15 patients when different NGS panels were used for genome interrogation, as well as the statistical significance of this difference in one analysis cohort, suggests that the panel selected to interrogate the genome is critical in the determination of TMB scores. This is a statement that echoes conclusions made in other TMB studies (Chalmers *et al.* 2017; Buchhalter *et al.* 2019; Campesato *et al.* 2015), and also reflects the findings of the 2021 EMQN/IQNPath pilot EQA scheme (Abate *et al.* 2022), which noted the variation in TMB scores when different NGS panels were used.

The TMB calculations performed are also critical to TMB status, as shown by the fact that the TMB status assigned to 3/13 patients in this study differed dependent on both the panel and analysis parameters (table 18). The exclusion/inclusion of artefacts in the TMB calculation did not impact on TMB status as much as the exclusion/inclusion of synonymous variants, with only 1/13 patients in the study changing TMB status

dependent on the inclusion of artefacts (panel and synonymous variant analysis stable), whilst the TMB status of 3/13 patients altered dependent on synonymous variant investigation. This finding suggests that the additional artefact removal step investigated within this thesis may be an unnecessary step in terms of having limited impact on patient outcomes (only 1/13 patients impacted). This potential lack of utility, reflects the absence of such an artefact removal step being discussed widely in TMB-focussed literature.

4.2.3 Study aim 3: Evaluation of the utility of TMB quantification in combination with PD-L1 expression analysis for immunotherapy response prediction

4.2.3.1 Association between TMB score and immunotherapy response in the PD-L1 expression high cohort

Spearman correlation coefficient (r_s) analysis showed that there was a weak positive correlation between TMB estimation and immunotherapy response (measured in terms of survival time post PD-L1 test) within this Welsh patient cohort of 17 patients of high PD-L1 expressor status. This positive correlation was present for both the Illumina and Nonacus NGS panels regardless of whether TMB estimates included or excluded synonymous variants and artefacts. This association supports the potential use of TMB in combination with PD-L1 expression status as an immunotherapy response predictor, showing that as TMB value increases, patient survival increases, endorsing similar conclusions made from internationally published data for non-Welsh patient cohorts (Rizvi *et al.* 2015; Carbone *et al.* 2017; Hellmann *et al.* 2018b).

The association between TMB estimation and immunotherapy response was weak in this Welsh study, with the scatter plots of these two attributes showing a wide spread of datapoints with little obvious correlation to the naked eye (figure 16). The correlation of the data will have been impacted by the small study size (n=17), with any outliers impacting heavily on the overall association; some apparent outliers are noticeable in figure 16, whereby long patient survival is associated with a lower TMB score, and vice versa, with higher TMB scores being found in patients with a short survival time. In addition to the limited size of this pilot study, there are a number of variables that were not controlled for within this patient cohort but which could impact on the patient response to immunotherapy and therefore survival, potentially masking correlation between TMB score and survival; these variables are discussed in section 4.3 in relation to the limitations of this study.

The highest r_s values across the eight panel/analysis datasets (four datasets per panel representing +/- synonymous variants and +/-artefacts) and consequently the most positive correlations between TMB estimation and immunotherapy response, were seen when the Nonacus panel was used for determination of TMB rather than when the Illumina panel was used. Specifically, the maximum r_s value of 0.154 was obtained when the Nonacus panel -synonymous +artefacts was used for TMB estimation. This r_s statistic alone suggests that this Nonacus panel/analysis combination for TMB estimation represents the optimal framework for TMB measurement within this small patient cohort. However, other factors requiring consideration in the evaluation of clinical utility of a specific panel/analysis combination are yet to be discussed.

4.2.3.2 Evaluation of the use of TMB high thresholds to define a TMB high patient group for accurate immunotherapy response patient stratification

Given the large variation in TMB estimations that have been reported both in this study and in some recent publications (Heeke *et al.* 2020; Ramos-Paradas *et al.* 2021; Vega *et al.* 2021; Abate *et al.* 2022) when different NGS panels and TMB calculations have been used within the same cohort, to have utility as a biomarker in a clinical environment a threshold of high TMB would be required to identify the cohort of patients who are likely to benefit from immunotherapy therefore guiding the use of this cancer treatment. Simply reporting a TMB value with no context regarding what the value means in clinical terms would not be clinically appropriate, and would not allow results between laboratories using different panel/analysis combinations to be compared; a TMB high/low status would have to be reported within a clinical setting, which is also a conclusion made within the TMB-focussed review by Sha *et al.* (2020). Interestingly, the majority of the laboratories participating in the EMQN/IQNPath 2021 pilot EQA scheme (Abate *et al.* 2022) disclosed that they were reporting only TMB estimation without the use of a TMB cut-off, which echoes the findings from the IQN Path 2019 survey showing that 61% of labs did not use a TMB threshold (Fenizia *et al.* 2021). Fenizia *et al.* (2021) noted that in line with the absence of a threshold in most labs, the majority of labs were not providing TMB-led patient stratification as were performing TMB assessment in only a research context.

10 variants/Mb TMB high threshold:

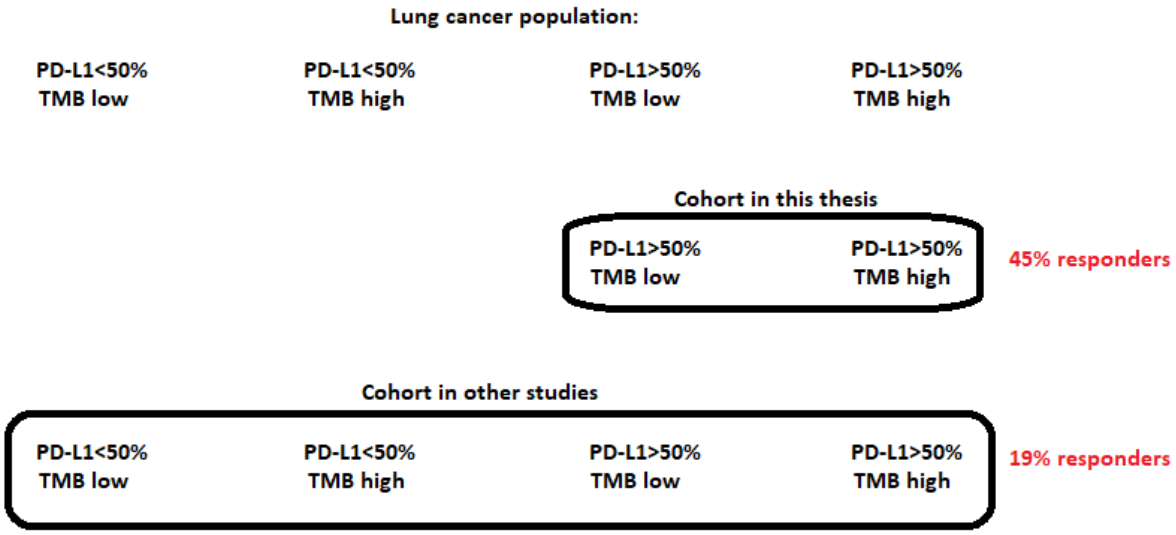
The 10 variants/Mb TMB high threshold determined by Ramalingam *et al.* (2018) and validated in three other research publications (Hellmann *et al.* 2018a; Pestinger *et al.* 2020; Marabelle *et al.* 2020) was not appropriate for use in this small Welsh patient cohort (total patient number = 17) where TMB estimations were >47 variants/Mb for all patients using all panel/analysis combinations.

The lack of utility of the 10 variants/Mb TMB high threshold within this thesis can be related to the fact that the four NSCLC-based publications (Ramalingam *et al.* 2018; Hellmann *et al.* 2018a; Pestinger *et al.* 2020; Marabelle *et al.* 2020) in which this threshold did show utility in accurately separating immunotherapy responders and non-responders, were based on NSCLC patient cohorts that were not pre-selected for PD-L1 high expression (figure 22). The Welsh cohort in this thesis, is composed of 17 patients who all had PD-L1 positive IHC results. In spite of the known inaccuracies of the PD-L1 biomarker in identifying immunotherapy responders, the patients in this pre-selected Welsh cohort are predicted to show a better response to immunotherapy than an unselected group of patients (Garon *et al.* 2015; Sul *et al.* 2016). Following on from this therefore, this PD-L1 positive cohort would be more likely to be TMB high, as high TMB correlates with improved immunotherapy response (Campesato *et al.* 2015; Johnson *et al.* 2016; Kowanetz *et al.* 2017; Hellmann *et al.* 2018a; Marabelle *et al.* 2020). This would skew the dataset within this thesis towards higher TMB values than in the unselected patient cohorts of Ramalingam *et al.* (2018), Hellmann *et al.* (2018a), Pestinger *et al.* (2020) and Marabelle *et al.* 2020. Based on data from Garon *et al.* (2015) and Sul *et al.* (2016), unselected patient cohorts would be predicted to contain a larger number of non-responders than the pre-selected PD-L1 high cohort of the thesis, so effectively reducing the TMB scores of an unselected patient cohort. This evaluation aligns with the fact that, in the Pestinger *et al.* (2020) study, which used the

same Illumina panel as utilised in this thesis and included both synonymous and non-synonymous variants in the TMB scores, the range of TMB values generated in the cohort of 22 lung patients was 1-30 variants/Mb. This is a much lower range than the TMB range in this Welsh study; when considering all analysis combinations (+/-synonymous variants and +/-artefacts) as a whole, the TMB estimations for the Illumina patient cohort (n=15) and for the Nonacus cohort (n=15) varied within a very similar range: 49-479 variants/Mb for Illumina, and 47-478 variants/Mb respectively. Importantly, no comparable dataset to this Welsh study (i.e. pre-selected PD-L1 high patients) has been identified in the literature, as the studies that demonstrated combined utility of TMB and PD-L1 assessment performed these evaluations independently of one another (Carbone *et al.* 2017; Castellanos *et al.* 2019); importantly, this was a fact that had been overlooked when this study was designed, so the implications of analysing a pre-selected PD-L1 patient cohort had not been pre-empted in terms of impact on TMB value. There is therefore no data to investigate whether the higher TMB values seen in the patients in the PD-L1 high cohort of this thesis are comparable to other reports.

The increase in TMB estimations in the PD-L1 selected NSCLC patient cohort in this thesis compared to 'unselected' NSCLC patients, shows the importance of specifying the remit of a TMB clinical service in terms of the patient group that will be evaluated. This decision regarding target patient group would have to be made prior to validation of a panel/analysis method to ensure that appropriate patient samples were used within the validation, ensuring the appropriate TMB threshold was selected within the validation and taken forward into live clinical service.

Figure 22: Impact of NSCLC patient cohort evaluated on TMB estimation. The lung cancer population as a whole can be considered to be split into four categories based on PD-L1 expression status and TMB status. The pre-selected PD-L1 high (>50%) cohort evaluated in this thesis is indicated. Sul *et al.* (2016) found that 45% of PD-L1 high (>50%) expressors responded to pembrolizumab; therefore, this is the estimated responder level within the thesis cohort. Also highlighted is the unselected patient cohorts evaluated by Ramalingam *et al.* (2018), Hellmann *et al.* (2018a), Pestinger *et al.* (2020), and Marabelle *et al.* (2020). Garon *et al.* (2015) determined the response to pembrolizumab within an unselected NSCLC cohort to be 19%.



ROC-curve generated TMB high thresholds:

Optimal TMB thresholds for maximising sensitivity (maximising true positive rate) and maximising specificity (minimising false positive rate) of the immunotherapy response predictor functionality of the TMB biomarker were determined by ROC curve analysis. Maximising true positives is important as, in a clinical TMB service, patients with false negative TMB results (i.e. TMB low reported instead of TMB high) would not receive immunotherapy which they could respond to. Equally, minimising false positives is important in a clinical TMB service as patients with false positive TMB calls would receive expensive immunotherapy that is unlikely to benefit them and from which they could experience side-effects, rather than a more appropriate treatment regime from which they could benefit and which could impact on their long-term survival.

Within this study, the process involved in generating a ROC curve and identifying the appropriate TMB high threshold was deemed to be straight forward, and could certainly be performed as part of a UK Genomics laboratory validation of a TMB service. In a review of the challenges of TMB quantification, Fancello *et al.* (2019) advocated the use of a statistical approach to TMB threshold setting to ensure the robustness of thresholds. The determination of these ROC curve thresholds relies upon the accuracy of both the RECIST-1 classification of patients (Eisenhauer *et al.* 2009), and the accuracy of TMB estimations, both of which are highlighted as limitations of this study (section 4.3). The presence of any outliers within the small cohort evaluated will also impact on the ROC-curve generated thresholds by impacting on the false

positive/false negative rate which is critical in the ROC curve threshold setting process. A larger cohort size would improve the accuracy of threshold setting.

A single TMB high threshold, providing optimal sensitivity and specificity for separation of patient responders and non-responders across the Illumina and Nonacus panels, was not identified in this study. This is perhaps not surprising given that the use of different TMB high thresholds in different targeted panel settings is well published (Campesato *et al.* 2015; Chalmers *et al.* 2017; Hellmann *et al.* 2018a; Heeke *et al.* 2018; Fenizia *et al.* 2021; table 4). Such variation amongst TMB high thresholds was also recently noted in the results from an IQN Path survey (Fenizia *et al.* 2021), which found that a variety of TMB thresholds were used in the 69 international labs performing TMB analysis, and that these were based on either literature, internal validation or validation against the FoundationOne CDx test. However, the fact that the TMB high threshold of 10 variants/Mb has shown utility across multiple panels (Ramalingam *et al.* 2018; Hellmann *et al.* 2018a; Pestinger *et al.* 2020; Marabelle *et al.* 2020) demonstrates that a single TMB threshold can show utility in different settings.

Although the variation in TMB threshold between the Nonacus and Illumina-targeted cohorts could have been predicted based on the existence of publications highlighting threshold variation between panels (Campesato *et al.* 2015; Chalmers *et al.* 2017; Hellmann *et al.* 2018a; Fenizia *et al.* 2021), variation in optimal TMB thresholds within the same panel as a result of variations in TMB analysis is unpublished. Collectively, these pieces of evidence lead to the conclusion that as a result of panel- and analysis-dependent TMB variation, TMB thresholds require validation on an appropriate scale to ensure utility in the panel/analysis combination being employed, which can result in panel/analysis-specific thresholds being identified.

The ROC-curve generated thresholds determined in this study show that the Nonacus panel provides a sensitivity and specificity for separation of immunotherapy responders/non-responders more reflective of the requirements of a clinical service compared to the Illumina-generated results. The highest sensitivity (100%) and specificity (78%) was achieved using the Nonacus panel +synonymous +/-artefacts (i.e. artefact inclusion/exclusion had no impact on these measurables). Focussing on the +/- artefact element of this panel/analysis combination, this information taken together with the (already discussed) observations that: the Nonacus panel -synonymous +artefacts generated the strongest correlation (r_s) between TMB score and survival, and that the removal of artefacts had limited utility in terms of the fact that TMB status was generally unaltered in this study by artefact inclusion/removal), does not support an additional artefact removal step prior to TMB estimation. There is no observed clinical benefit in this study in removing artefacts from TMB calculations. The utility of synonymous variants within TMB estimations is discussed in the next section.

4.2.3.3 Association between TMB high status and immunotherapy response in the PD-L1 expression high cohort

Across the four different TMB analysis methods (+/- synonymous variants +/- artefacts) in both the Nonacus and Illumina datasets (n = 15 for both the Illumina and Nonacus NGS panel cohorts), all eight panel/analysis datasets showed an increase in average patient survival in the high TMB group compared to the low TMB group (using ROC-curve generated TMB high thresholds; figure 19). This increased survival of the TMB

high group reflects the correlation (r_s) between increasing TMB score and longer patient survival observed in this study cohort (section 4.2.3.1).

The difference in survival noted in this thesis between high and low TMB groups did not meet statistical significance (χ^2=0.258, p=0.05) within the single panel/analysis cohort evaluated (Nonacus -synonymous +artefacts), but the outliers (TMB high patient with short survival time and TMB low patient with long survival time) noted within this small dataset (figure 16) will have a large impact and will have contributed to this lack of statistical significance, as well as other study limitations impacting on these statistics (section 4.3).

The improved survival of TMB high patients in this study mirrors the findings of other research studies, including the Checkmate-026 trial (Carbone *et al.* 2017; Peters *et al.* 2017; Seiwert *et al.* 2018). As with the observed correlation between TMB estimation and patient survival, this correlation supports the potential utility of a combined TMB and PD-L1 expression status biomarker for anti-PD-L1 immunotherapy treatment response in this Welsh lung cancer patient cohort. The magnitude of this potential benefit in comparison to the use of PD-L1 expression analysis alone in predicting immunotherapy response is discussed in the next section.

There was generally a greater difference in survival between the high and low TMB groups when TMB was calculated using the Nonacus panel rather than the Illumina panel, irrespective of the analysis parameters used. This finding is reflective of the observation that TMB estimations generated from the Nonacus panel more strongly correlated with immunotherapy response than those from the Illumina panel, and again suggests a potential role for the Nonacus panel within a clinical TMB service.

The panel/analysis combination with the largest difference in mean survival between the TMB high and low groups was Nonacus +synonymous +/- artefacts. The fact that the largest difference in survival observed occurs independently of artefact removal further supports a lack of utility of artefact removal in TMB estimation as previously noted (section 4.2.3.2). In terms of this panel/analysis combination including synonymous variants in the TMB calculations, this echoes the data from Rizvi *et al.* (2015) whom investigated the impact of synonymous variants on TMB values and noted that the inclusion of synonymous variants in TMB calculations strengthened the association between TMB and immunotherapy response. However, in this Welsh study, the cohort with the smallest difference between survival in the high and low TMB groups also included synonymous variants in the calculations (Illumina +synonymous +/- artefacts); therefore, this dataset contradicts the Rizvi *et al.* (2015) findings. The Rizvi *et al.* (2015) study involved a small cohort of 18 NSCLC patients, which is a similar study size as this Welsh study, therefore both are subject to the constraints of small datasets and no firm conclusions can be drawn from this information alone.

The exclusion of synonymous variants was required to produce the most positive correlation (r_s) between TMB and immunotherapy response (Nonacus -synonymous +/- artefacts). However, the greatest difference in mean survival between the high TMB and low TMB groups, and the greatest sensitivity and specificity for responder/non-responder classification was identified when synonymous variants were included in the TMB estimations (Nonacus +synonymous +/-artefacts). It is not

clear from this conflicting data whether there is utility in removing synonymous variants from TMB calculations. This lack of clarity could be related to the large number of synonymous variants detected in the Nonacus dataset which could be skewing correlations.

In the recent evaluation of 11 commercially available panels performed by Sha *et al.* (2020), it was noted that the panel-specific bioinformatics pipelines designed specifically for each panel by the relevant commercial company, all measured non-synonymous variants only. This dataset included the two FDA-approved TMB panels, MSK-IMPACT and FoundationOne CDx®, which have proven utility in analysing TMB for accurate prediction of immunotherapy response (Fda.gov, 2017; Fda.gov, 2020b). Therefore, there appears to be a new trend towards calculating TMB without the addition of synonymous variants, which could be a strategy followed in future research.

4.2.3.4 Combined benefit of TMB and PD-L1 expression status over PD-L1 expression status alone in the prediction of immunotherapy response

Within the total patient cohort (n=17), RECIST-1 classification of the PD-L1 high (>50%) patient tumours showed that only 41% of these patients responded to immunotherapy (RECIST 1.1 categories CR, PR and SD; Eisenhauer *et al.* 2009). This figure nearly perfectly matches that demonstrated by Sul *et al.* (2016) in a NSCLC trial in which only 45% of patients responded to pembrolizumab despite the tumours being PD-L1 high (>50%) expressors. The low sensitivity for accurate responder identification using PD-L1 expression status alone demonstrated in this thesis cohort (and by Sul *et al.* 2016) supports the need for an alternative immunotherapy response biomarker, which is a concept that drove this research study. Thesis data demonstrates that the sensitivity of TMB estimation for accurate responder/non-responder RECIST 1.1 classification within this PD-L1 expressor patient cohort ranges from 78% to 100%, dependent on the panel/analysis method used to determine the patient TMB scores (discussed in previous section). This suggests that performing TMB analysis alongside PD-L1 assessment within this patient cohort, using any of the panel/analysis combinations investigated within this thesis, would have greatly improved the sensitivity of immunotherapy responder detection compared to using the PD-L1 biomarker alone to predict response. This improved sensitivity is important clinically as ensures the correct patients receive immunotherapy, so maximising the benefit of this treatment in the patient population. This finding supports the thesis hypothesis by demonstrating combined utility of TMB and PD-L1 expression analysis and echoing published findings (Carbone *et al.* 2017; Peters *et al.* 2017; Seiwert *et al.* 2018).

No statistically significant difference in patient survival was identified between the PD-L1 high group (n=15) and the combined TMB high + PD-L1 high group (n=10) ($H = 0.077$, $p = 0.782$) in the Nonacus -synonymous variants +artefacts dataset statistically analysed. Despite this lack of statistical significance, which will be influenced by some of the limitations of this study (section 4.3), the median survival times of these groups was 636 days and 964 days respectively, equating to a 52% increase in survival in the TMB high + PD-L1 high group, which reflects the improved responses demonstrated in the Checkmate-026 trial when this combined biomarker is used (Carbone *et al.* 2017). This is suggestive of a combined benefit of PD-L1 and TMB assessment for immunotherapy response prediction using the Nonacus -synonymous variants +artefacts parameters.

4.2.4 Study aim 4: Evaluation of the feasibility of a TMB service within the NHS

4.2.4.1 Cost

If NGS-based TMB evaluation was performed alongside PD-L1 assessment for NSCLC sample analysis in a clinical setting, then this NGS cost (table 20) would be in addition to the existing £100 cost of PD-L1 IHC analysis (costings obtained via personal communication with University Hospital of Wales Cellular Pathology department). The need for TMB assessment prior to immunotherapy stratification would also likely delay the initiation of treatment whilst NGS testing took place; NSCLC samples analysed by NGS at AWMGS are currently reported within 14 calendar days of sample receipt which extends beyond the PD-L1 turnaround time. Whether such an additional testing cost and time delay could be tolerated within the NHS would be dependent on the added clinical benefit that TMB analysis provides in this clinical scenario. The potential benefit of TMB assessment in combination with PD-L1 assessment has been demonstrated in this thesis and in other studies (Carbone *et al.* 2017; Castellanos *et al.* 2019), although no statistically significant trial-based data regarding improved survival using TMB + PD-L1 status currently exists. Future NSCLC-focussed clinical trials proving the utility of a combined biomarker approach for prediction of immunotherapy response, could justify the use of TMB assessment for NSCLC patients in a clinical setting in terms of added patient benefit; specifically, avoiding unnecessary side-effects from inappropriate targeting, and ensuring patients receive the most appropriate treatment in a timely manner. As well as having patient benefit, the introduction of NGS-based TMB testing in the NHS could also be economically beneficial owing to the improved targeting of this expensive therapy (NICE 2016a) to the most appropriate patient group. The cost to the NHS of TMB analysis could be off-set to some degree by savings in inappropriately administered immunotherapy; importantly, AWMGS would still have to apply to Welsh Government commissioners for funding of the genomic test.

Having stated a case for the NHS absorbing the additional testing costs associated with an NGS-based TMB service, this thesis has shown that TMB testing has the potential to be delivered at no additional cost to the Genomics lab or to the NHS. Both the Illumina and Nonacus panels have utility in the delivery of genetic testing for the existing UK-funded NSCLC service in terms of the genes targeted by the panels overlapping with the genes relevant to NICE-approved stratified treatments in NSCLC (NICE, 2022). Therefore, TMB analysis could be provided by one of these panels at no additional cost, provided that the panel was used for existing standard of care testing too.

With the realisation that a TMB service could be cost neutral to the Genomics laboratory, and cost saving to the NHS in terms of the improved targeting of costly immunotherapy, a further cost saving approach to TMB service delivery could be considered within the NHS by using TMB alone as a biomarker of immunotherapy response. Such a TMB-only NSCLC clinical service would remove the requirement for the £100/patient assessment of PD-L1 status for immunotherapy stratification, as well as reducing the sample requirements by removing the need for FFPE material for PD-L1 assessment. Owing to the fact that the patient cohort in this study was pre-selected for PD-L1 high status, it is not possible to evaluate the utility of TMB alone in predicting immunotherapy response, although the recent KEYNOTE-158 trial has demonstrated this utility (Marabelle *et al.* 2020).

Co-analysis of TMB and NSCLC clinically relevant genes:

The use of an NGS panel for TMB assessment within NSCLC patients and for the detection of specific clinically relevant genetic variants would require both of these elements to be evaluated within a validation setting in a clinical laboratory. In the context of this study, the results indicate that the Nonacus panel may not be suitable for the detection of clinically relevant genetic variants in NSCLC patients owing to the low sequencing coverage achieved for the samples analysed, as well as the apparent large number of artefacts in the Nonacus dataset. Both of these factors could limit the utility of the Nonacus panel for specific NSCLC-relevant gene analysis by decreasing the likelihood of identifying low level (5%) variants and increasing the complexity of variant analysis via the interrogation of potential non-genuine variants (i.e. artefacts). Conversely, sequencing coverage was better, and estimated artefact prevalence was lower on the Illumina panel. However, there are limitations of this study that could be impacting on the Nonacus data (section 4.3), and this study certainly does not represent the scale of analysis that would be performed within a clinical service validation. This study therefore does not rule out either panel as having utility in TMB analysis across the NSCLC population as a whole, and no conclusion regarding the panel utility for NSCLC-relevant gene analysis can be made. If either panel was going to be considered for TMB service delivery within AWMGS, an appropriate validation including optimisation of the target enrichment pipeline would have to be performed alongside any additional validation work to evaluate detection of clinically relevant NSCLC gene variants.

Since this thesis was conceived, the Illumina panel investigated in this study has been validated and successfully implemented within the AWMGS for the delivery of NGS of all solid tumours including lung cancer samples for NICE-approved treatment stratification-based testing. Based on this new service development in AWMGS, the Illumina panel would be an ideal choice for the assessment of TMB in AWMGS.

Validation strategies for an NHS-based TMB service:

If the clinical utility of TMB as a biomarker was linked to a NICE-approved drug, then UK Genomics laboratories, including AWMGS, would be required to implement a TMB service within 60 days of this approval to facilitate patient access to this drug. Genomic laboratories in the UK utilise a number of different targeted panels for the analysis of solid tumour samples. Potentially, the most practical and cost-effective solution to delivery of a TMB service within an NHS Genomic laboratory would be for each lab to validate their existing solid tumour panel for its ability to evaluate TMB.

An alternative to a lab-by-lab approach to TMB service validation and implementation, standardisation of TMB assessment could dictate that all labs have to use a specific panel for TMB assessment. The benefit of this approach within the NHS would be that a central validation could be initiated, reducing validation resources required across all sites. However, given the existing variation in panels used for solid tumour analysis across the UK Genomic labs, a dictated panel would be unlikely to be a feasible option as would require:

- Either: the simultaneous use of an existing standard of care NGS panel and a TMB-focussed panel, which would have cost and resource implications. This strategy would also increase sample requirements for this service, which is an issue for lung cancer patients where tumour samples are often small.
- Or: complete re-validation of existing solid tumour NGS services at each UK laboratory onto the new TMB dictated panel to ensure an ongoing cost neutral position for TMB testing (obviously the validation would have costs associated with it so not initially cost neutral). This would again be cost prohibitive.

One other option for TMB service delivery would be the centralisation of TMB testing in the UK; however, this would have complexities in terms of:

- the movement of FFPE samples or DNA between labs and the associated time for this sample transfer which would delay the patient care pathway.
- the transfer of data and results between labs to ensure that analysis and reporting could be distributed across all UK laboratories, which would need to be carefully controlled and managed.
- the large sample numbers involved and the associated capacity requirements of the delivering laboratory.

Based on the above limitations, centralised TMB testing in the UK is unlikely to be feasible, and the most likely UK delivery model is the use of existing NGS panels within each UK laboratory for TMB assessment following appropriate validations.

4.2.4.2 Availability of EQA schemes for TMB assessment

The variation in TMB value dependent on the panel and analysis methods used as noted in this and other studies (Heeke *et al.* 2020; Ramos-Paradas *et al.* 2021; Vega *et al.* 2021; Abate *et al.* 2022) shows that TMB estimations generated from different panels cannot be compared directly. This presents an interesting conundrum for the establishment of a TMB-focussed EQA scheme, designed to evaluate the consistency and accuracy of TMB quantification across scheme participants who will undoubtedly use a range of different NGS methods to assess TMB. Considering the lack of utility of cross-panel comparison, it is interesting that the 2021 EMQN/IQNPath pilot EQA scheme (Abate *et al.* 2022) chose to compare all TMB values generated from the 24 participating laboratories to the FoundationOne CDx TMB estimations, which is one of the FDA-approved panels (Fda.gov, 2020b) that had been used to validate the samples within the EQA scheme. This EQA process is flawed as does not have utility owing to the proven differences in TMB estimation when different panels and analysis are used (Heeke *et al.* 2020; Ramos-Paradas *et al.* 2021; Vega *et al.* 2021). The EQA scheme was however able to perform more appropriate comparisons for two targeted NGS panels which were used by multiple participating labs, and so comparison of data from these labs was more informative in terms of evaluating lab performance in TMB assessment as these were like-for-like comparisons. Although, having said this, the bioinformatics pipelines used for evaluation of TMB in each participating lab were not scrutinised in this EQA pilot, therefore it is unknown if there are any perfect like-for-like comparisons in terms of the same panel and analysis being performed. The existence of a TMB pilot EQA scheme is an enabler of future TMB clinical services.

4.2.4.3 Variation in TMB status dependent on panel/analysis/threshold used

Within this study, only 6/13 (46%) samples were assigned the same TMB status (TMB high or TMB low) independent of the NGS panel and analysis parameters used. Stenzinger *et al.* (2020) similarly found that TMB statuses (TMB high/low) of 20 samples analysed in 15 laboratories using six different targeted NGS panels were inconsistent with the WES-determined TMB statuses of these samples in 25% of cases. This lack of consistency between the TMB status aligned to a patient dependent on the methodology used is a concern within a clinical setting including within the NHS where, unless UK guidance stipulated a specific panel had to be used for TMB estimation, there would be variability in the NGS panels and bioinformatic pipelines used in different laboratories (based on resources, local preferences etc.), which would ultimately mean that the treatment received would be determined by where a patient's sample was analysed. A factor to bear in mind here is that the panels used in this thesis have only been evaluated using 17 patients, therefore this pilot study is far from a true validation of these technologies, where many more samples would be evaluated, which would potentially reduce the discrepancies in TMB statuses assigned by each panel. This likely reduction in TMB status variation using appropriately validated panels is supported by the lower level of variation in TMB status assignment seen in the Stenzinger *et al.* (2020) publication that compared data from labs already delivering TMB services.

The Stenzinger *et al.* (2020) cross-laboratory evaluation, suggested the use of a three-tier TMB classification system, assigning samples a TMB high, TMB low, or a TMB intermediate status. When this system was used by Stenzinger *et al.* (2020) there was a reduction in the percentage of samples that were completely mis-classified (i.e. reported as TMB high rather than TMB low, or vice versa) from 25% using the two-tier high/low system to 1.5% using the three-tier system. This thesis data also supports the utility of an intermediate TMB status, as the results showed (table 18) that, of the seven patients in the study with variable TMB status dependent on panel and/or analysis used, five of these had TMB estimations that were within 20 variants of the TMB high threshold. If an intermediate status had been assigned to all patients whose TMB scores were within 20 variants of the high threshold, then the percentage of totally mis-classified samples (high instead of low/low instead of high) in this study would have decreased from 54% (7/13) to 23% (3/13). Notably, this intermediate threshold assessment has been performed looking at all of the analysis parameters in combination using only the seven discordant samples. The utility of an intermediate zone within TMB evaluation does warrant more investigation in future studies with a larger cohort size. The use of an intermediate zone is not novel in the genomics field as these are already used in other services where defined cut-offs are not appropriate.

The utility of an intermediate zone in minimising false positive and false negative rates is described well by Mattocks *et al.* (2010), where its utility is noted in relation in particular to situations in which both sensitivity and specificity are of critical importance. In a TMB clinical service, maximising sensitivity and specificity would both be important, as the results of the TMB analysis would dictate the patient's cancer treatment. How an intermediate TMB would be acted on by a clinician would be an important consideration to discuss with Oncologists as part of a TMB service validation, prior to service launch. Mattocks *et al.* (2010) suggest that samples in the intermediate zone could be reported as test failures; if this approach was taken within

a clinical TMB service then the test failure rate would have to be considered within the validation and this would want to be minimised to ensure the clinical utility of the service was maximised.

Validation of a TMB clinical service would involve determination of the appropriate TMB high threshold based on the panel and bioinformatics pipeline used for TMB assessment. The potential utility of ROC-curve generated thresholds has been demonstrated in this study in terms of ease of use within a clinical setting, although the analysis of a greater number of samples than used within this study, which would be performed within a clinical service validation, would be expected to result in a more clinically useful threshold based on the reduced impact of any false positive or false negatives on the ROC curve assessment. If a three-tier TMB classification system was to be used in a clinical TMB service, appropriate validation of the size of the intermediate zone would also have to be performed. Within this study, the intermediate zone was arbitrarily set as within 20 variants of the TMB high threshold, but within a clinical service this intermediate threshold would have to be appropriately considered and validated using an appropriate number of samples.

An area that is outside the scope of this thesis but which is important to mention in relation to the feasibility of a TMB service within the NHS, is the fact that published studies have shown that TMB threshold is likely to vary based on the tumour type being analysed (Sha *et al.* 2020). This would mean that individual threshold validations would have to be performed within laboratories delivering clinical TMB services for each tumour type analysed, which would have resource and cost implications.

4.3 Limitations: Elements influencing determination of TMB clinical utility within this study

4.3.1 Patient cohort

A key limitation of this pilot study is the cohort size, which was limited owing to the funding available for this thesis and the cost of NGS analysis. The small cohort (n=17) will have impacted on the correlations observed, with any outliers impacting heavily on the overall associations, and will have contributed to the lack of statistical significance in the findings. An increase in size of patient cohort could improve statistical significance of a future TMB evaluation study.

As the true TMB status of each sample is not known in this study, the accuracy of TMB scoring by each panel/analysis combination could not be determined directly. The utility of TMB estimation was instead assessed by comparing the predicted immunotherapy response (based on high/low TMB) against the RECIST-1 criteria (Eisenhauer *et al.* 2009) and/or the patient survival time. The use of samples of known TMB status would have been useful to assess in this study, and would certainly be required within a clinical validation of a TMB service; unfortunately, such samples were not available from the WCB. Cell lines, EQA samples, and/or cancer samples from the 100,000 Genomes Project of known TMB status could be used within any TMB research or TMB service validations of the future to provide a benchmark with which to assess panel/analysis utility. Another potential source of samples, for which TMB

status could be determined, are lung cancer samples already analysed within standard of care genomic testing at AWMGS using the 523-gene Illumina TruSight™ Oncology 500 panel, for which NGS data would be available for re-analysis in a research setting if research approval was sought. Importantly, WES-derived TMB status would be the preferred method used for sample benchmarking as this is still considered the 'gold standard' of TMB quantification. It is important to consider that certain biological and social factors, e.g. smoker status (Alexandrov *et al.* 2016; Sha *et al.* 2020) and stage of disease (Zhang *et al.* 2018), have been associated with higher TMB scores within published NSCLC datasets (table 3). In order to be as representative of the total NSCLC population as possible, the patient cohort was not controlled for any of these variables that could impact on TMB score. Patients were only selected from the WCB on the basis of having PD-L1 high expressing NSCLC tumours and having received pembrolizumab treatment in the first line setting. Variation in these other elements within the patient cohort could influence the results observed, potentially masking correlations between TMB score and patient survival. Perhaps if some elements of the patient cohort had been controlled, for example selection of patients at the same stage of disease, then an improved correlation between TMB score and survival would have been seen. However, as already noted, such variation is expected to be representative of the lung cancer population and would generally not be controlled within a clinical TMB service. The only element that could be controlled within a clinical setting would be the stage of disease as the TMB analysis would be aligned to a particular point in the patient care pathway.

4.3.2 Artefact removal algorithm

Reducing false positive variant calls, via the removal of sequencing artefacts, would obviously be ideal in terms of improving accuracy of the TMB score but the issue is, if the method of identification of these false positive variant calls is flawed, the TMB score generated could be an under- or over-representation of the genuine TMB of the sample. Importantly the number of artefacts removed from each dataset in this study was based on the predicted percentage of artefacts within each panel/analysis combination. Owing to resource limitations, this prediction was based on IGV interrogation of a small selection of shared variants (section 2.6), with the premise being that shared variants are likely to represent SNPs, hotspot variants or artefacts. The uneven interrogation of shared variants across the two panel cohorts (less Illumina-generated variants interrogated) could introduce bias into the predicted artefact frequencies.

The number of variants removed from the datasets based on the predicted artefact prevalence in each panel/analysis dataset is likely to be too high, as the predicted artefact number is likely to be an over-estimation. The reason behind this over-estimation is that the variants interrogated in IGV were the most frequently occurring variants within the datasets (having been identified in >4 patients) therefore have the highest likelihood of being an artefact. However, this predicted artefact prevalence was then used to remove potential artefacts across all shared variants, which included variants shared by just 2 patients.

Based on the idea that the method of artefact removal employed in this study will potentially result in an excess of variants being removed from the datasets, the TMB estimations generated when artefacts are removed would be expected to be a lower representation of the true values. A more accurate algorithm for prediction of

artefacts within each panel dataset could have been obtained by analysing a proportion of variants for likely artefact status from variants shared at both a high and low frequency across each dataset. Such improvements to this algorithm would have been beneficial in future studies if artefact removal had been identified as a critical component of accurate TMB estimation.

4.3.3 Responder classification

The classification of patients into responder and non-responder cohorts is critical in evaluating the clinical utility of TMB, as any mis-classification could skew the data in favour of/against the utility of TMB as a biomarker. In this study, patients who were CR, PR, or who showed SD according to RECIST 1.1 criteria (Eisenhauer *et al.* 2009) were classified as responders as these categories encompass patients in which no tumour-progression has occurred. Arguably, patients with SD could be omitted from the responder cohort as the tumours in these patients have not decreased in size according to RECIST 1.1 criteria (Eisenhauer *et al.* 2009) therefore perhaps 'response' has not been demonstrated in this cohort.

The response categorisation (CR, PR, SD, PD, IR) of this Welsh cohort is an action driven by the clinical team. Personal correspondence with local Welsh oncologists identified that the RECIST 1.1 criteria are not routinely used in cancer patient NHS care pathways, therefore perhaps suggesting that there is scope to mis-classify patients across the categories. The mis-classification risks noted can be deemed to be consistent across the cohort in terms of impacting on all patients to the same degree.

4.3.4 Survival data

Owing to the availability of samples within the WCB, patients in the cohort had a PD-L1 IHC test date of between August 2017 and May 2019, rather than all having analysis within 2017 which was the original plan. In line with the <10% 5-year survival rate (NICE 2021) and the 59% 3-year survival rate of lung cancer patients (Albano, Bilfinger, and Nemesure 2018), 71% (12/17) of the patients had sadly died at the time of medical record review in April 2022. For these patients, survival could be calculated from the date of the PD-L1 test to the date of death. For the five patients that were alive in April 2022, the survival time was calculated from the PD-L1 test date to 11/04/2022 (the date of the final medical record review). This resulted in these 5 patients having an artificially shortened survival time. The survival data for this Welsh cohort will therefore be an under-representation of true survival.

Owing to the fact that in this study, the PD-L1 tests were all performed some years ago, over an approximate 2-year time-frame, even though the survival times of 5 patients were under-estimated, these surviving patients still had the longest survival times in the study (1112-1699 days) by some margin, with the exception of a single patient (20M70076) who died with a survival time post-PD-L1 test of 1292 days. Unsurprisingly, given the survival time of patient 20M70076, this patient had stable response according to RECIST 1.1 criteria. The long survival times of the patients still alive at the end of this study means that there is minimal data skewing caused by a short survival time of one of these surviving patients.

Patient selection could have been improved by ensuring that all patients within the cohort were already deceased, thus ensuring accurate survival times for the cohort. The potential issue with this strategy, however, is that this could mean that the PD-L1 analysis was performed >5 years ago (pre-2017), which could impact on the quality of

the DNA obtained from these older FFPE samples. Alternatively, an 'overall survival rate' metric could have been used to assess the percentage of patients alive at a certain time point e.g. 18 months after diagnosis; this mirrors the approach used in other TMB studies, such as the Hellmann *et al* (2018a) publication in which progression-free survival at one-year was the metric used for evaluation of therapy response. This approach would have meant that all patients had complete and accurate data for this 'overall survival' metric. An alternative measure of response such as 'time to response' could have been used as a metric in this study if further interrogation of the medical records had been performed. This metric represents a different way of measuring response and would only have a value for those patients that respond to therapy, but is a metric typically used as a secondary endpoint in clinical trials.

Chapter 5: Conclusions

This study demonstrates the improved sensitivity of anti-PD-L1 immunotherapy stratification in a Welsh NSCLC patient cohort using a combined TMB and PD-L1 biomarker, compared to PD-L1-based stratification. This supports the study hypothesis and supports the potential clinical utility of this combined biomarker in this patient group; although, the small study size contributes to the lack of statistical significance in the survival difference between patients of high PD-L1 expression status versus patients with combined TMB high and PD-L1 high expression status, and limits the ability to generalise these findings across the Welsh lung cancer population as a whole. The potential cost neutral status of delivery of a TMB service within the NHS, as determined in this study, and the existence of a pilot EQA scheme for TMB quantification (Abate *et al.* 2022), extends the proposed clinical utility of a combined TMB and PD-L1 biomarker to encompass a demonstrated feasibility of TMB service provision in the NHS.

Analysis of the same patient cohort on different NGS panels highlights the key complexity of TMB assessment, namely that TMB estimation is dependent on both the targeted NGS panel used to assess variant number, whereby size and/or gene content of panel may impact on the TMB score, and the analysis performed in respect of the variants included within the TMB calculation. TMB estimation can also vary based on attributes of the tumour sample, e.g. pre-selection of samples in this study for PD-L1 high samples raised the TMB scores in comparison to other unselected study cohorts. The demonstration of variability in TMB estimation based on panel and analysis methodology highlights the essential requirement for panel/analysis-specific TMB thresholds determined through validation, as well as reporting of TMB status rather than TMB estimation, to deliver clinical utility and enable cross-comparison of TMB data between laboratories. This may result in the same threshold having utility across different panels, e.g. the 10 variants/Mb threshold determined by Ramalingam *et al.* (2018). The differences in TMB estimation as a result of the use of different panel and analysis combinations, can be clinically significant by altering the TMB (high/low) status of a patient dependent on the TMB threshold used, which would result in patients aligning to different treatment options within a TMB biomarker-based service. This represents a barrier to the clinical utility of TMB as a biomarker, although the clinical impact of this variation in TMB status can be reduced by the use of a three-tier threshold system, specifying TMB high, TMB low, and TMB intermediate categories.

There are many elements within this study that have identified a preference for the Nonacus panel over the Illumina panel in terms of added benefits that the Nonacus panel was observed to bring in relation to: an improved correlation between TMB estimation and immunotherapy response, and greater difference in survival between the Nonacus TMB high and low groups. However, some areas of the Nonacus panel performance are concerning in terms of the poor sequencing coverage achieved and the poorer quality of the sequencing compared to Illumina, although these could have been affected by lack of experience in the methodology. This is a pilot study evaluating a total cohort of only 17 patients, so in the context of this, neither the Illumina nor Nonacus panel can be ruled out as a potential targeted panel for use within a TMB clinical service. In general terms, in relation to optimal panel selection, study findings support a minimum panel size of 1.6Mb to provide a level of assurance of the accuracy

of the TMB estimations in relation to the gold-standard WGS/WES approaches. In terms of optimal analysis, an additional artefact removal step based on evaluation of potential artefacts in IGV was demonstrated to have minimal impact on TMB status, and including this step within future TMB studies/validations is not recommended. No clear conclusion regarding the utility of synonymous variants in TMB calculations could be made, as the optimal analysis parameters in relation to the exclusion/inclusion of synonymous variants varied within this study dependent on the area of utility being considered, e.g. correlation between TMB and immunotherapy response, or the sensitivity and specificity of responder/non-responder classification.

Based on the findings of this study there are a number of recommendations for NHS Genomics laboratories when embarking on a TMB service validation, many of which can be applied to the design of future research studies (table 21).

Table 21: Recommendations for TMB quantification within the NHS. The validation steps (recommendations 4 and 5) can be performed across laboratories if the panel/analysis combination is consistent. Importantly, if a validation of the same panel/analysis on the same patient population is already published then only an internal verification of the procedure (including evaluation of the TMB threshold within the publication) would be required.

Recommendations	
1 Service	Based on proven utility of TMB assessment for patient stratification (Marabelle *et al.* 2020) and associated cost benefits of a stratification service based only on TMB quantification, an independent biomarker model is recommended within the NHS.
2 Panel	The use of targeted NGS panels (>1.6Mb) already established within NHS Genomics labs are recommended for TMB estimation. Capability of the panel for TMB evaluation could be confirmed by reviewing the relevant commercial website; such confirmation would not be possible for custom-design panels or panels developed in-house.
3 Analysis	- An additional artefact removal step is not required within the TMB calculation. - Refer to commercial website regarding recommendations for inclusion/exclusion of synonymous variants for selected panel. If no guidance exists, exclusion of synonymous variants could be performed based on a growing preference towards these conditions (Sha *et al.* 2020). - Use of the Institut Curie TMB tool (Github, 2022ahttps://github.com/bioinfo-pf-curie/TMB) for TMB calculation.
4 Validation	- Validate panel/analysis combination using an appropriate number of TMB known status samples (ideally TMB determined by WES, or an FDA-approved panel). - The validation of the panel for TMB quantification should establish the minimum coverage of the panel required to ensure the accuracy of TMB scoring, ensuring that TMB is not under-estimated owing to poor coverage.
5 Threshold	- Determine the appropriate TMB thresholds using ROC curve analysis; the specification of an intermediate zone is also recommended. - Note: Different threshold validations may be required for different tumour types where TMB estimations can vary.
6 Report	- Clinical reports should record the TMB status as well as the TMB score. - Consideration into how to report intermediate TMB scores via consultation with referring Oncologists.

Chapter 6: Future work

The thesis findings and the published data regarding combined TMB and PD-L1 biomarker utility lack statistical significance; further research studies and trials would be required to demonstrate such utility prior to a clinical service based on a combined biomarker being considered. Given the demonstrated clinical utility of TMB assessment alone as a biomarker (Marabelle *et al.* 2020), and the benefits of this single biomarker strategy within the NHS (and potentially within other international clinical services) research into a combined biomarker may be less urgent.

A larger study to demonstrate utility of a TMB biomarker within a Welsh cohort could be initiated to observe if a stronger correlation between TMB score and immunotherapy response could be determined. This could be driven by AWMGS who, as the Illumina TruSight™ Oncology 500 panel was launched in July 2021 for the delivery of the routine NSCLC genomic service, have access to Illumina-panel generated sequencing data from hundreds of NSCLC patients. A research application for retrospective TMB evaluation of these samples could be submitted, but the issue with such a study design is that the study, as for this thesis, would be focussed on the utility of a combined TMB + PD-L1 biomarker, as immunotherapy response data would only be available on PD-L1 high samples, and, as concluded, this may not be the desired strategy for an NHS service.

Further research could be performed to answer the question posed in this study regarding whether concordance between targeted panel-generated and WGS/WES-determined TMB quantifications aligns with clinical utility of the targeted panel for TMB evaluation. A comparison of the recommended analysis parameters outlined in this study against the gold-standard WES would be an informative future study to demonstrate utility of these recommendations. Other areas highlighted within this study worthy of further evaluation include the use of a three-tier threshold system for improved clinical utility of TMB as a biomarker, and the inclusion of synonymous variants within TMB estimations, which remains an unresolved question in terms of a lack of consensus in this area.

The utility of the Agilent panel could be investigated further as the reason for failure of this panel was not identified. Technical support could be sought prior to use of this panel in the AWMGS laboratory, or a collaborative study with Agilent could enable the NGS set-up to be performed at a site external to AWMGS, with only data interrogation occurring at AWMGS; this would eliminate potential technical errors owing to lack of familiarity of AWMGS staff with the Agilent protocol, which is a possible explanation for this panel failure.

The ongoing provision of EQA schemes is important for quality assurance within NHS TMB services. The preferred format for EQAs would be the comparison of sample TMB status rather than TMB estimations, as these can be compared directly across labs using different TMB panel/analysis combinations. Differences in TMB status between participating labs and the validation labs could indicate a poor performing laboratory where perhaps the panel/analysis/threshold combination has not been appropriately validated. The issue with an EQA scheme based on comparison of TMB status is that within the IQNPath scheme (Abate *et al.* 2022) many participating labs did not assess TMB status, choosing instead to report TMB estimation only; therefore, TMB status results from all participating labs would not be available at the present time. However,

the number of labs assigning TMB status is likely to increase in the future if service delivery is moved from a research to a clinical setting based on further approvals being granted (e.g. NICE-approval) of drugs whose use is stratified based on TMB quantification. If TMB estimations were submitted by EQA participants then data on the TMB assessment parameters used would be required so that any caveats to the utility of comparing lab results from participants using the same panel were understood.

There is already an international TMB harmonisation project led by the Friends of Cancer Research, established to improve the utility of TMB by standardising TMB measurement and reporting across different NGS panels. The first phase of this collaborative project produced preliminary recommendations for the standardisation of TMB assessment, which included the use of a calibration curve generated using universal reference standards of varying TMB score (assessed by WES) (Merino *et al.* 2020). TMB scores for these reference standards generated at different laboratories using different panels could then be aligned to the calibration curve to produce conversion factors that could be applied to all TMB analysis at a given site, to aid comparison of values between laboratories. This is an interesting suggestion, and it will be fascinating to see if this idea is developed further in the next phase of this collaborative project to drive the utility of TMB assessment in clinical practice.

Appendices

Appendix 1:

A Units and C1 Credits for Appendix to DClinSci Thesis

Alliance Manchester Business School (AMBS) **A Units**		
Unit Title	**Credits**	**Assignment Word Count**
A1: Professionalism and Professional Development in the Healthcare Environment	30	Practice Paper – 2000 words A1 – Assignment 1 – 1500 words A1 – Assignment 2 – 4000 words
A2: Theoretical Foundations of Leadership	20	A2 – Assignment 1 – 3000 words A2 – Assignment 2 – 3000 words
A3: Personal and Professional Development to Enhance Performance	30	A3 – Assignment 1 – 1500 words A3 – Assignment 2 – 4000 words
A4: Leadership and Quality Improvement in the Clinical and Scientific Environment	20	A4 – Assignment 1 – 3000 words A4 – Assignment 2 – 3000 words
A5: Research and Innovation in Health and Social care	20	A5 – Assignment 1 – 3000 words A5 – Assignment 2 – 3000 words

Life Sciences – Section C

C1: Innovation Project – Credits = 70

Assignment: Literature Review & Lay Presentation

Appendix 2:

DClinSci section C1 submission: Innovation proposal and business case

Utility of Tumour Mutational Burden as a biomarker for immunotherapy stratification in Welsh lung cancer patients

1 Executive summary

Revenue requirement for this proposal: £21,436.56 (appendix 11.2); Capital requirement: £0

The existing PD-L1 expression biomarker used for prediction of immunotherapy response in lung cancer patients is imperfect, with only 45% of patients showing response to the anti-PD-L1 immunotherapy pembrolizumab after stratification of treatment has occurred (Garon *et al.* 2015). There is, therefore, a clinical and financial need for a more accurate predictor of immunotherapy response to minimise the use of unnecessary and ineffective therapy.

Tumour Mutational Burden (TMB) can be simply defined as the number of mutations found within a tumour (Alexandrov *et al.* 2013). Recent clinical trials have highlighted the potential clinical utility of TMB quantification using Next Generation Sequencing (NGS) as a predictor of immunotherapy response in lung cancer patients (Carbone *et al.* 2017; Hellmann *et al.* 2018). Despite the large interest in this field, a number of questions regarding the technical utility of TMB remain unanswered, and the clinical utility of TMB –based immunotherapy stratification to prolong overall survival is unproven; these factors hinder the use of TMB in the clinical setting.

Aims:

This proposal aims to drive forward TMB research using diagnostic-grade samples by targeting some of these unanswered questions, whilst at the same time ensuring the All Wales Medical Genomics Service (AWMGS) gains expertise in this novel area. The primary aim and objective of this research is to assess the clinical utility of TMB as a biomarker for anti-PD-L1 immunotherapy treatment response in a Welsh lung cancer patient cohort, either alone or in combination with PD-L1 expression analysis. This proposal aims to assess the utility of TMB by performing TMB measurements on a cohort of 24 patients using 3 targeted NGS panels and, in doing so, compare the accuracy of TMB estimation across these methodologies. TMB will be estimated using different analysis criteria and will be performed with a range of TMB high thresholds. The secondary objective will be achieved by the identification of a set of optimal conditions for TMB evaluation that effectively differentiate between immunotherapy responders and non-responders.

2 Background

Immunotherapy is of increasing importance in cancer patient care pathways. Owing to biological and genetic differences in tumours, cancer patients do not all respond to immunotherapy (Garon *et al.* 2015). One such difference between tumours is the

expression level of the PD-L1 protein. Research into differences in PD-L1 expression levels across lung tumours, resulted in high PD-L1 expression being linked to good immunotherapy response in this patient group (Taube *et al.* 2014). These findings led PD-L1 expression to be the first biomarker used for the stratification of immunotherapy use in lung cancer patients (NICE 2016a). However, even with the use of this biomarker, around half of the patients still do not respond to the immunotherapy (Garon *et al.* 2015).

Another difference between tumours is the number of genetic mutations within the tumour, described as TMB. The identification of tumours with high TMB by NGS, either by analysis of the whole exome (Whole Exome Sequencing, WES) or a proportion of the exome (targeted NGS), has been shown to be predictive of immunotherapy response in lung cancer patients (Rizvi *et al.* 2015; Hellmann *et al.* 2018). Importantly, the combined use of TMB level and PD-L1 status has been shown to have benefits in the prediction of immunotherapy response (Carbone *et al.* 2017). TMB therefore has the potential to be an alternative biomarker for immunotherapy response.

3 The current state of TMB research

Within a clinical trial setting, high TMB has been shown to be associated with improved progression-free survival in lung cancer patients (Hellmann *et al.* 2018). This trial data unfortunately did not achieve statistical significance in terms of overall survival of these patients (Hellmann *et al.* 2018), and until such clinical utility has been demonstrated, TMB will not be a viable biomarker in this patient group. Other barriers to the use of TMB as a biomarker in immunotherapy stratification are the lack of consensus within the literature regarding the mutations to count within a TMB estimate, and the TMB threshold to use to accurately differentiate between immunotherapy responders and non-responders (Rizvi *et al.* 2015; Pestinger *et al.* 2020). Importantly a high TMB threshold of 10 mutations/Mb has been used within 3 separate research studies using different targeted NGS panels and has been shown to be an effective threshold for predicting immunotherapy response (Pestinger *et al.* 2020; Hellmann *et al.* 2018; Ramalingam *et al.* 2018). Ramalingam *et al.* 2018 originally defined this 10 mutations/Mb threshold using a ROC (Receiver Operating Characteristic) curve, the use of which has been supported in other publications (Fancello *et al.* 2019). Within the literature, the key differences in the TMB calculations performed surround the inclusion/exclusion of synonymous mutations and insertion-deletions in the estimation (Fancello *et al.* 2019; Budczies *et al.* 2019).

The method required for accurate TMB quantification is highly debated within the literature. WES has traditionally been seen as the gold-standard for TMB measurement, but its use is often prohibited by the staffing and IT resources required to handle, interpret and store the large amounts of sequencing data generated. As such, targeted NGS panels have been increasingly used in TMB research, and have shown consistency in TMB estimations when compared to WES-generated levels (Rizvi *et al.* 2018; Chalmers *et al.* 2017). There is, however, debate regarding the ideal size and gene content of a targeted NGS panel to enable accurate TMB estimation (Bucchalter *et al.* 2018; Budczies *et al.* 2019). A minimum panel size of 1.5Mb has been quoted by Buchhalter *et al.* (2019) to be necessary for TMB quantification, but panels smaller than this have been proven to provide TMB estimates that are concordant to WES-determined values (Chalmers *et al.* 2017; Rizvi *et al.* 2018).

4 The value of further research into TMB utility as a biomarker for immunotherapy stratification

Addressing the TMB research gaps to identify a set of optimal analysis criteria for TMB quantification using targeted NGS panels will enhance the potential for TMB to be used as a biomarker within cancer care pathways. Focusing on a Welsh lung patient cohort will provide novel research into the clinical utility of TMB as an immunotherapy response biomarker in this population.

This proposal ensures that the AWMGS is well positioned to deliver a NGS-based TMB analysis should a service be required in the future. It could also drive further research studies within the AWMGS of TMB assessment in other tumour types or in circulating tumour DNA samples.

In terms of cost/benefit analysis, the current immunohistochemical (IHC) test performed to determine PD-L1 expression levels in lung cancer patients prior to immunotherapy stratification is cheaper than a targeted NGS test (£50 vs £350). However, if NGS has improved utility over expression analysis as a predictor of immunotherapy response, then it would be economically beneficial to introduce such NGS testing within a clinical setting owing to the high cost of immunotherapy treatment (NICE 2016a). As well as having monetary benefits, use of a more accurate predictor of response would be beneficial to cancer patients whom could avoid unnecessary treatments. Importantly, targeted NGS panels that can be used for TMB quantification are used currently within Genomics laboratories as part of cancer care pathways to identify specific genetic mutations associated with response to certain NICE-approved drugs. Therefore, the cost of the NGS panel test is already factored into a lung patient's pathway so effectively making TMB quantification cost neutral.

5 Option appraisal

Option 1: Do nothing

Not taking forward this proposal would mean that the research questions posed regarding the utility of TMB as an immunotherapy response biomarker within the Welsh population, and the optimal method of TMB quantification would remain unanswered. This puts the onus on other researchers to take forward these questions surrounding the clinical and technical utility of TMB, which are currently acting as barriers to the use of TMB as a biomarker. Should these questions be answered in other research studies and NICE approval for TMB-based immunotherapy stratification be granted, the response rate of AWMGS to deliver a clinical service for TMB quantification would be hindered, having not gained experience in this area through this proposal.

Option 2: Evaluate the clinical and technical utility of TMB within the AWMGS

This proposal has been designed to target unanswered questions that are hindering the use of TMB as a biomarker within the healthcare setting. It also provides the opportunity to generate TMB data specifically from Welsh lung cancer patients (a need which has not been met to date), and to develop TMB quantification experience within the AWMGS laboratory. Such expertise would be of great utility to the AWMGS and to the cancer patients of Wales should there be a need in the future for a TMB service to guide cancer care. This proposal will use diagnostic-grade samples therefore has real

utility in assessing TMB quantification capabilities of the NGS panels chosen within a diagnostic setting.

Alternative approaches within option 2 would be to either use WES or targeted NGS panels.

a) Use of WES: The AWMGS has no experience of WES of tumour samples and does not have the IT support to make this a feasible option. Using WES would mean that the scope of this proposal would be reduced to exclude the evaluation of the ideal size and gene content of targeted sequencing panels for TMB detection.

b) Use of a targeted sequencing panel: This approach fits best with the aims of this proposal and with the skill mix within the AWMGS laboratory, where targeted NGS panels have been used for genetic analysis of tumour samples since 2015.

6 Preliminary data in support of proposal

In support of this proposal and specifically the use of targeted NGS panels to define the technical requirements of TMB quantification (option 2b), a small TMB pilot study was performed at AWMGS, in collaboration with Qiagen, to assess whether TMB measurements could be successfully obtained by targeted NGS analysis of diagnostic-grade samples. NGS of 12 formalin-fixed paraffin-embedded (FFPE)-derived DNAs from a range of tumour types was performed at AWMGS using the Qiagen Life Sciences TMB panel (1.3Mb in size), with TMB quantification then being performed by analysts at Qiagen. Of these 12 samples, 11 gave TMB measurements (included lung, bladder and melanoma samples), whilst one melanoma sample did not successfully produce sequencing data (appendix 11.1).

This preliminary work included 6 lung samples with starting DNA concentrations of between 8 and 11ng/ul, which is highly reflective of the samples received within diagnostic service. The ability to successful sequence and estimate TMB levels within all 6 of these lung samples provides support for the use of targeted NGS panels within this proposal.

6 Costs

The cost of the chosen option 2b, covering sample sourcing (£2438), NGS consumables (£18,498.56), staffing (£0), and publication costs (£500) is £21,436.56 (appendix 11.2).

7 Stakeholder engagement

The proposal is fully supported by the Head of the AWMGS laboratory. The benefits that this research would bring have also been recognised by the Wales Cancer Research Centre, whom awarded a Translational Research Award (ASTRA) to this proposal following a successful application process. This ASTRA application was a joint collaboration with two local Lung Oncologists whom were fully engaged and enthused by this area of TMB research. The use of ASTRA funding can be supported by the following lay summary.

Lay summary:

Cancers arise from uncontrolled cell growth caused by the accumulation of genetic mutations in a single cell from which the tumour develops. The number of mutations in a tumour can be quantified, and this number, for some types of cancer including

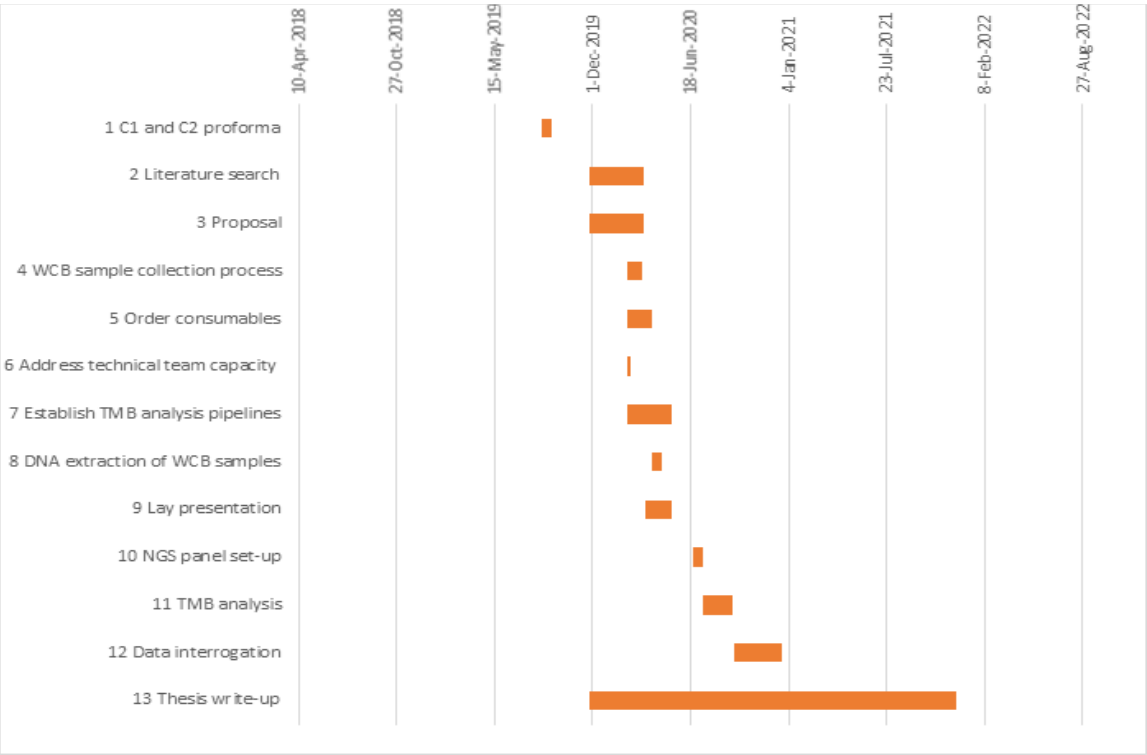
lung, has been shown to be associated with how well a patient responds to a particular cancer treatment called immunotherapy. This proposal aims to investigate how best to measure the number of mutations within a lung tumour, and will evaluate what the critical number of mutations is that determines therapy response. These are both areas lacking consensus and, once defined, could allow this mutation counting technique to be used within cancer care pathways to identify patients most likely to respond to immunotherapy. Such tailored cancer care ensures that patients receive the most effective therapies, inappropriate drug usage is reduced, and cancer patient survival is improved.

8 Risks and challenges (appendix 11.3)

Description of Risk	Risk Impact Score	Risk Likeli-hood Score	Risk Rating Score	Response / Mitigation Action
Risk 1: Insufficient DNA extracted from FFPE samples meaning that there are limited numbers of samples that can be analysed on all 3 panels, so limiting the ability to compare panels for TMB utility.	4	3	12	Panels with variable DNA input amounts selected to try to ensure that there will be enough DNA for at least 2 panels.
Risk 2: FFPEs are challenging to work with in terms of potentially providing poor quality DNA, which could influence the quality of the NGS data obtained.	3	3	9	Preliminary work performed to give an indication of the expected failure rate.
Risk 3: NGS costs are high and DNA samples are precious therefore cannot afford any errors in NGS set-up.	4	2	8	2/3 NGS panels have already been used in the lab, therefore familiarity with the technical process. Training provided to lab for remaining panel.
Risk 4: Establishing a bioinformatic analysis pipeline for TMB calculation.	4	1	4	Experienced bioinformatics team on hand in AWMGS; literature review identified approaches to performing TMB calculations.
Risk 5: Insufficient samples available within the Welsh Cancer Bank to make valuable conclusions.	4	3	12	The proposal was designed based on minimal samples required to produce statistically significant data.

9 Proposal timeline

Gannt chart detailing estimated timings of this proposal. This timeline may be subject to change owing to conflicting NHS service pressures within the laboratory or other unforeseen circumstances.



10 Method

A panel of 24 lung cancer patients already stratified for immunotherapy treatment based on PDL-1 IHC results will be sought from the Welsh Cancer Bank (WCB). The WCB has ethical approval from Wales Research Ethics Committee for cancer related research meaning that individual studies do not need separate ethical approval. DNA will be extracted from the identified patient tumour samples, which will be in the form of formalin-fixed paraffin-embedded (FFPE) tissue samples. 3 NGS panels, from Illumina, Agilent and Nonacus, will be used to generate NGS data.

NGS data will be analysed using an in-house bioinformatics solution. Multiple TMB calculations will be performed on the data, centring on the inclusion/exclusion of insertion-deletions and synonymous mutations. Scatter plots will be used to visualise the relationship between TMB and immunotherapy response (response plotted as post-PD-L1 assessment survival in months). Spearman rank correlation coefficient calculation and linear regression analysis will be used to investigate any association between TMB level and response, using 2 different TMB high thresholds (10 mutations/Mb and a Receiver Operating Characteristic curve generated threshold). A paired t-test will be used to compare the TMB values obtained from samples run on >1 panel. The potential benefit of using PD-L1 expression in combination with TMB for immunotherapy response prediction will be investigated using ANOVA to compare the mean response of the TMB high group, the PDL-1 >50% group and the TMD high + PDL-1 >50% group.

Sample numbers: The number of samples being processed on each of the 3 NGS panels will be dependent on the volume of DNA obtained from each sample (see risks section 8).

- A minimum number of 8 samples will be analysed on each of the 3 NGS panels; 8 represents the minimum number of observations required to perform the Spearman rank correlation test.
- A minimum of 12 samples will be analysed on >1 NGS panel to provide a statistically significant dataset. The expected difference between TMB levels measured by different panels is unknown so power calculation is not possible.

Selection of targeted NGS panels: The 3 NGS panels selected for use in this proposal was based on a number of factors:

a) Experience of AWMGS in using the NGS panel: 2/3 of the panels had been used in the laboratory before therefore staff were appropriately trained.
b) Gene content of the panels: All 3 panels target oncogenes and tumour suppressor genes but differ in their shared gene content, allowing evaluation of optimal gene content for TMB estimation.
c) Utility for other tumour analysis: All panels had potential utility to be used for current standard of care genetic analysis of lung tumours, therefore ensuring a cost neutral position for TMB should a service for required in the future.
d) Size of the panels: All panels are over the critical 1.5Mb size predicted by Bucchalter *et al.* (2019) to be essential for TMB estimation. One panel is <1.6Mb which represents the panel size limit suggested by the work of Hatakeyama *et al.* (2018).
e) The Illumina panel has proven utility in TMB measurement by comparison to whole genome sequencing (Pestinger *et al.* 2020); therefore, Illumina TMB measurements will be used as a guide by which the TMB levels from the other panels can be compared.

10 References

1) Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S.A.J.R., Behjati, S., Biankin, A.V., Bignell, G.R., Bolli, N., Borg, A., Børresen-Dale, A.-L., Boyault, S., Burkhardt, B., Butler, A.P., Caldas, C., Davies, H.R., Desmedt, C., Eils, R., Eyfjörd, J.E., Foekens, J.A., Greaves, M., Hosoda, F., Hutter, B., Ilicic, T., Imbeaud, S., Imielinsk, M., Jäger, N., Jones, D.T.W., Jones, D., Knappskog, S., Kool, M., Lakhani, S.R., López-Otín, C., Martin, S., Munshi, N.C., Nakamura, H., Northcott, P.A., Pajic, M., Papaemmanuil, E., Paradiso, A., Pearson, J.V., Puente, X.S., Raine, K., Ramakrishna, M., Richardson, A.L., Richter, J., Rosenstiel, P., Schlesner, M., Schumacher, T.N., Span, P.N., Teague, J.W., Totoki, Y., Tutt, A.N.J., Valdés-Mas, R., van Buuren, M.M., van 't Veer, L., Vincent-Salomon, A., Waddell, N., Yates, L.R. Australian Pancreatic Cancer Genome Initiative, ICGC Breast Cancer Consortium, ICGC MMML-Seq Consortium, ICGC PedBrain, Zucman-Rossi, J., Futreal, P.A., McDermott, U., Lichter, P., Meyerson, M., Grimmond, S.M., Siebert, R., Campo, E., Shibata, T., Pfister, S.M., Campbell, P.J., and Stratton, M.R. (2013) Signatures of mutational processes in human cancer. *Nature,* 500:415-21.

2) Buchhalter, I., Rempel, E., Endris, V., Allgäuer, M., Neumann, O., Volckmar, A.-L., Kirchner, M., Leichsenring, J., Lier, A., von Winterfeld, M., Penzel, R., Christopoulos, P., Thomas, M., Fröhling, S., Schirmacher, P., Budczies, J., and Stenzinger, A. (2019) Size matters: Dissecting key parameters for panel-based tumor mutational burden analysis. *Int J Can* 144(4):848-858.
3) Budczies, J., Allgäuer, M., Litchfield, K., Rempel, E., Christopoulos, P., Kazdal, D., Endris, V., Thomas. M., Frohling, S., Peters, S., Swanton, C., Schirmacher, P, and Stenzinger, A. (2019) Optimizing panel-based tumor mutational burden (TMB) measurement. *Ann Oncol* 30(9):1496-1506.
4) Carbone, D.,P., Reck, M., Paz-Ares, L., Creelan, B., Horn, L., Steins, M., Felip, E., van den Heuvel, M.M., Ciuleanu, T.-E., Badin, F., Ready, N., Hiltermann, T.J.N., Nair, S., Juergens, R., Peters, S., Minenza, E., Wrangle, J.M., Rodriguez-Abreu, D., Borghaei, H., Blumenschein, G.R., Villaruz, L.C., Havel, L.M., Krejci, J., Jaime, J.C., Chang, H., Geese, W.J., Bhagavatheeswaran, P., Chen, A.C., and Socinski, M.A. (2017) First-Line Nivolumab in Stage IV or Recurrent Non–Small-Cell Lung Cancer. *N Engl J Med* 376:2415-2426
5) Chalmers, Z.R., Connelly, C.F., Fabrizio, D., Gay, L., Ali, S.M., Ennis, R., Schrock, A., Campbell, B., Shlien, A., Chmielecki, J., Huang, F., He, Y., Sun, J., Tabori, U., Kennedy, M., Lieber, D.S., Roels, S., White, J., Otto, G.A., Ross, J.S., Garraway, L., Miller, V.A., Stephens, P.J., and Frampton, G.M. (2017) Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med* **9,** 34.
6) Fancello, L., Gandini, S., Pelicci, P.G. and Mazzarella, L. (2019) Tumour mutational burden quantification from targeted gene panels: major advancements and challenges. *J for Immunotherapy of Cancer* 7:180.
7) Garon, E.B., Rizvi, N.A., Hui, R., Leighl, N., Balmanoukian, A.,S., Eder, J.P., Patnaik, A., Aggarwal, C., Gubens, M., Horn, L., Carcereny, E., Ahn, M., Felip, E., Lee, J.S., Hellmann, M.D., Hamid, O., Gldman, J.W., Soria, J.C., Dolled-Filhart, M.D., Rutledge, R.Z., Zhang, J., Lunceford, J.K., Rangwala. R., Lubiniecki, G.M., Roach, C., Emancipator, K., and Gandhi, L. (2015) Pembrolizumab for the Treatment of Non–Small-Cell Lung Cancer. *N Engl J Med* 372:2018-2028.
8) Hatakeyama, K., Nagashima, T., Urakami, K., Ohshima, K., Serizawa, M., Ohnami, S., Shimoda, Y., Ohnami, S., Maruyama, K., Naruoka, A., Akiyama, Y., Kusuhara, M., Mochizuki, T., and Yamaguchi, K. (2018) Tumor mutational burden analysis of 2,000 Japanese cancer genomes using whole exome and targeted gene panel sequencing. *Biomed Res.* 39(3):159-167.
9) Hellmann, M.D., Ciuleanu, T.E., Pluzanski, A., Lee, J.S., Otterson, G.A., Audigier-Valette, C., Minenza, E., Linardou, H., Burgers, S., Salman, P., Borghaei, H., Ramalingam, S.S., Brahmer, J., Reck, M., O'Byrne, K.J., Geese, W.G., Green, G., Chang, H., Szustakowski, J., Bhagavatheeswaran, P., Healey, D., Fu, Y., Nathan, F., and Paz-Ares, L. (2018) Nivolumab plus Ipilimumab in lung cancer with a high tumor mutational burden. *N Engl J Med*. 378(22):2093-2104.
10) NICE (2016a) *Pembrolizumab for treating PD-L1-positive non-small-cell lung cancer after chemotherapy.* National Institute for Health and Care Excellence (NICE). [Online] [Accessed 18th September 2020] https://www.nice.org.uk/guidance/ta428/documents/final-appraisal-determination-document
12) Pestinger, V., Smith, M., Sillo, T., Findlay, T.M., Laes, J.-F., Martin, G., Middleton, G., Taniere, P., and Beggs, A.D. (2020) Use of an Integrated Pan-

Cancer Oncology Enrichment Next-Generation Sequencing Assay to Measure Tumour Mutational Burden and Detect Clinically Actionable Variants. *Mol Diagn Ther.* 24(3):339-349.

13) Ramalingam, S.S., Hellmann, M.D., Awad, M.M., Borghaei, H., Gainor, J., Brahmer, J., Spigel, D.R., Reck, M., O'Byrne, K.J., Paz-Ares, L., Zerba, K., Li, X., Geese, W.J., Green, G., Lestini, B., Szustakowski, J.D., Chang, H., and Ready, N. (2018) Tumor mutation burden (TMB) as a biomarker for clinical benefit from dual immune checkpoint blockade with nivolumab (nivo) + ipilimumab (ipi) in first-line (1L) non-small cell lung cancer (NSCLC): identification of TMB cutoff from CheckMate 568. *Am Ass Can Res* 78 (13).
14) Rizvi, N.A., Hellmann, M.D., Snyder, A., Kvistborg, P., Makarov, V., Havel, J.J., Lee, W., Yuan, J., Wong, P., Ho, T.S., Miller, M.L., Rekhtman, N., Moreira, A.L., Ibrahim, F., Bruggeman, C., Gasmi, B., Zappasodi, R., Maeda, Y., Sander, C., Garon, E.B., Merghoub, T., Wolchok, J.D., Schumacher, T.N., Chan, T.A. (2015) Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* 348(6230):124–128.
15) Rizvi, H., Sanchez-Vega, F., La, K., Chatila, W., Jonsson, P., Halpenny, D., Plodkowski, A., Long, N., Sauter, J.L., Rekhtman, N., Hollmann, T., Schalper, K.A., Gainor, J.F., Shen, R., Ni, A., Arbour, K.C., Merghoub, T., Wolchok, J., Snyder, A., Chaft, J.E., Kris, M.G., Rudin, C.M., Socci, N.D., Berger, M.F., Taylor, B.S., Zehir, A., Solit, D.B., Arcila, M.E., Ladanyi, M., Riely, G.J., Schultz, N., and Hellmann, M.D. (2018) Molecular Determinants of Response to Anti–Programmed Cell Death (PD)-1 and Anti–Programmed Death-Ligand 1 (PD-L1) Blockade in Patients With Non–Small-Cell Lung Cancer Profiled With Targeted Next-Generation Sequencing. *J Clin Oncol* 36(7): 633–641.
16) Taube, J.M., Klein, A., Brahmer, J.R., Xu, H., Pan, X., Kim, J.H., Chen, L., Pardoll, D.M., Topalian, S.L., and Anders, R.A. (2014) Association of PD-1, PD-1 Ligands, and Other Features of the Tumor Immune Microenvironment with Response to Anti–PD-1 Therapy. *Clinical Cancer Research*. 20(19):5064–5074.

11 Appendix

11.1 TMB data from 12 samples: preliminary Qiagen study

Sample	All mutations / Mb	Somatic mutations / Mb	Intergenic mutations / Mb	Synonymous mutations / Mb	Non-synonymous mutations / Mb (TMB score)
1	562.93	13.75	0	4.58	9.17
2	496.4	15.02	1.5	1.5	12.02
3	515.64	4.5	0.75	0.75	3
4	516.13	12.75	0.75	3.75	8.25
5	582.63	10.52	0	2.26	8.27
6	523.1	6.06	0	0	6.06
7	531.33	9.76	0.75	4.5	4.5
8	496.36	10.51	0.75	2.25	7.51
9	344.68	15.79	1.75	3.51	10.52
10	406.04	14.06	1.76	3.52	8.79
11	0	0	0	0	0
12	336.38	52.97	1.77	12.36	38.85

11.2 Estimated costings of the proposal

* Note: The number of samples that can be sequenced is restricted by the NovaSeq costs combined with the budget for this research work.

Item	Additional information	Estimated cost
WCB sample sourcing	24 samples from Welsh Cancer Bank (£50/sample + £450 admin fee + £500 patient data); Maxwell extraction costs for 24 samples (£12/sample)	£2,438
Library preparation: Illumina - TSO500	TSO500 DNA kit for 48* samples	£8758.20
Library preparation: Agilent Custom SureSelect CancerCore v1	Free kit provided as Agilent panel to be trialled for its variant detection capabilities in a parallel project	free
Library preparation: Nonacus Cell3 Target pan cancer panel	Nonacus Pan-Cancer (524) for 16* samples	£1389
Sequencing: NovaSeq	NovaSeq Xp 2-Lane Manifold Pack (£1071.60), NovaSeq Xp 2-Lane Kit (£461.70), NovaSeq 6000 SP Reagent Kit (300 cycle) (£5371.56)	£6904.86

Other consumables	Target Pure NGS clean-up beads, 10ml (£157.50), UK Delivery, Dry Ice (£38), Beckman - AMPure XP DNA cleanup kit (60ml) (£975), Oxford Gene Technology-Dynabeads™ M270 Streptavidin, 2ml (£276)	£1446.50
Staffing	1x Band 6: member of staff already funded by Welsh government for development work. 1x Band 8a: HSST funded post with dedicated development time	free
Publication	MAP (Molecular Analysis for Personalised Therapy) conference poster presentation 2021: cost of poster, meeting fee and transport to meeting venue	£500
TOTAL		£21,436.56

11.3 Risk calculations

Risk Rating = Impact x Likelihood

Impact		Likelihood	
Level	**Description**	**Level**	**Description**
1	Insignificant	5	Almost Certain
2	Minor	4	Likely
3	Moderate	3	Possible
4	Major	2	Unlikely
5	Catastrophic	1	Rare

Risk rating Key

Score 1-3	*Low risk*	This level of risk is considered acceptable, no action is required over and above existing procedures
Score 4-6	*Moderate risk*	Action Required- review
Score 8-12	*Significant risk*	Action required-review
Score 15-25	*Critical risk*	Immediate action required

Appendix 3:

DClinSci section C1: Oral presentation report

Guidelines on the Assessment of the Oral Presentation

The assessment panel will assess the presentation based on the following criteria:

- Quality and clarity of explanation of the project and proposal for a lay audience (awareness of the use of jargon, scientific language and acronyms)
- Synthesis of relevant scientific evidence for a lay audience
- Ability to persuade a lay audience of the merits (or otherwise) of the project and its potential role in healthcare science services
- Style of presentation (slides, delivery; body language, eye contact, voice, confidence) and appropriateness for a lay audience
- Demonstration of values, attitudes and behaviours expected of a leader in clinical science

The assessment panel should also consider the presentation against the following criteria, in line with expectations for doctoral degrees:

- The creation and interpretation of new knowledge, through original research or other advanced scholarship. This new knowledge must be of a quality to satisfy peer review, extend the forefront of the discipline and merit publication.
- Significant contribution towards the development of novel and innovative research.
- A systematic acquisition and understanding of a substantial body of knowledge that is at the forefront of an academic discipline or area of professional practice.
- The general ability to conceptualise, design and implement a project for the generation of new knowledge, applications or understanding at the forefront of the discipline, and to adjust the project design in the light of unforeseen problems.
- A detailed understanding of applicable techniques for research and advanced academic enquiry

Following deliberation, the assessment will be given a pass/fail outcome and written feedback will be available afterwards from the panel.

Should the student fail the presentation one further opportunity to undertake the oral presentation will be offered.

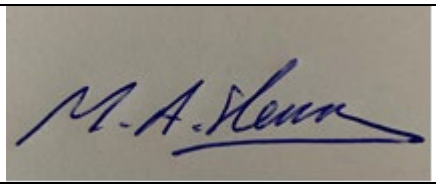
Section 1: Student

Name of Student:	Helen Roberts
Title of Project:	Evaluation of a new method of predicting treatment response in lung cancer patients

Section 2: Report and Recommendation

Please provide a report on the explanation of the work and response to questions.
Good background, interesting news focussed and highlighted on immunotherapy. Good lay conversion. Predicting response-stratifying, precision medicine. Surface protein what is it? Number of genetic mutations= better way to predict is high TMB also= more aggressive or worse prognosis for the cancer? Very nice description, novel and interesting research/clinical question. Can you get DNA from the blood? -secretions linked to genetic mutations Can you combine with other therapies? Very good response to all questions.
Overall Recommendation **Please tick**
✶ Pass ☐ Fail
Feedback *In the event of a fail being awarded, please provide detailed feedback to the student on what is required in order to attain a pass mark.*

Section 5: Signatures

Name of Panel Chair	Mark Slevin	Signature of Panel Chair	
Date:	25/3/21		

Please provide a list of panel attendees

Fiona Wilkinson

Jane Lynch

John McCormack

Appendix 4:

Royal College of Pathologists (FRCPath) part 1 examination certificate

The Royal College of Pathologists

By these letters make it known that

Helen Marie Roberts

having undertaken the required training and
after having passed the Part One examination in

Molecular Genetics

has been awarded

Associateship of
The Royal College of Pathologists

In witness whereof the Seal of the College and the signatures of the proper Officers have been affixed this 22nd day of January in the year of our Lord 2015

LONDON

President *Registrar* *Member of Council*

Appendix 5:

Gene lists of evaluated targeted NGS panels

Gene content of 3 panels

Clinically relevant NSCLC genes are highlighted in red.

Agilent	Illumina	Nonacus
AKT1	ABL1	ABCB1
AKT2	ABL2	ABCC2
AKT3	ACVR1	ABL1
ALK	ACVR1B	ABL2
AMER1	AKT1	ACTB
APC	AKT2	ACVR1B
APLNR	AKT3	ADH1B
AR	ALK	AIP
ARAF	ALOX12B	AKT1
ARID1A	ANKRD11	AKT2
ARID1B	ANKRD26	AKT3
ARID2	APC	ALDH2
ASXL1	AR	ALK
ATM	ARAF	AMER1
ATR	ARFRP1	AP3B1
ATRX	ARID1A	APC
AURKA	ARID1B	AR
AXL	ARID2	ARAF
B2M	ARID5B	ARFRP1
BAP1	ASXL1	ARID1A
BCL2	ASXL2	ARID2
BLM	ATM	ARID5B
BRAF	ATR	ASXL1
BRCA1	ATRX	ATM
BRCA2	AURKA	ATR
CBL	AURKB	ATRX
CCND1	AXIN1	AURKA
CCND2	AXIN2	AURKB
CCND3	AXL	AXIN1
CCNE1	B2M	AXL
CD274	BAP1	B2M
CD58	BARD1	BAP1
CDK12	BBC3	BARD1
CDK2	BCL10	BCL2
CDK4	BCL2	BCL2L1
CDK6	BCL2L1	BCL2L11
CDKN1A	BCL2L11	BCL2L2
CDKN1B	BCL2L2	BCL6
CDKN2A	BCL6	BCOR
CDKN2B	BCOR	BCORL1
CHEK2	BCORL1	BCR
CIITS	BCR	BIRC3
CREBBP	BIRC3	BLM
CTCF	BLM	BMPR1A
CTNNB1	BMPR1A	BRAF
DAXX	BRAF	BRCA1
DICER1	BRCA1	BRCA2
DNMT3A	BRCA2	BRD4
EGFR	BRD4	BRIP1
EP300	BRIP1	BTG1
EPHA3	BTG1	BTG2
ERBB2	BTK	GTK
ERBB3	C11orf30	BTLA
ERBB4	CALR	BUB1B
ERG	CARD11	C11orf30
ESR1	CASP8	CALR
ETV6	CBFB	CARD11
EZH2	CBL	CBFB
FAS	CCND1	CBL
FBXW7	CCND2	CCND1
FGF19	CCND3	CCND2
FGFR1	CCNE1	CCND3
FGFR2	CD274	CCNE1
FGFR3	CD276	CCT6B
FGFR4	CD74	CD22
GATA3	CD79A	CD274
GNA11	CD79B	CD58
GNAQ	CDC73	CD70
GNAS	CDH1	CD79A
H3F3A	CDK12	CD79B
H3F3B	CDK4	CDA
HGF	CDK6	CDC73
HIST1H3B	CDK8	CDH1
HIST1H3C	CDKN1A	CDK10
HIST2H3C	CDKN1B	CDK12
HLA-A	CDKN2A	CDK4
HLA-B	CDKN2B	CDK6
HLA-C	CDKN2C	CDK8
HNF1A	CEBPA	CDKN1A
HRAS	CENPA	CDKN1B
IDH1	CHD2	CDKN1C
IDH2	CHD4	CDKN2A
IGF1R	CHEK1	CDKN2B
JAK1	CHEK2	CDKN2C
JAK2	CIC	CEBPA
JAK3	CREBBP	CEP57

Agilent	Illumina	Nonacus
JUN	CRKL	CHD2
KDR	CRLF2	CHD4
KIT	CSF1R	CHD7
KLF4	CSF3R	CHEK1
KMT2A	CSNK1A1	CHEK2
KRAS	CTCF	CHIC2
MAP2K1	CTLA4	CIC
MAP2K2	CTNNA1	CKS1B
MAP2K4	CTNNB1	CREBBP
MAP3K1	CUL3	CRKL
MAPK1	CUX1	CRLF2
MAX	CXCR4	CSF1R
MCL1	CYLD	CSF3R
MDM2	DAXX	CTCF
MED12	DCUN1D1	CTLA4
MEN1	DDR2	CTNNA1
MET	DDX41	CTNNB1
MLH1	DHX15	CUL3
MSH2	DICER1	CUX1
MSH6	DIS3	CXCR4
MTOR	DNAJB1	CYLD
MUTYH	DNMT1	CYP19A1
MYB	DNMT3A	CYP2A6
MYC	DNMT3B	CYP2B6
MYCN	DOT1L	CYP2C19
NBN	E2F3	CYP2C9
NF1	EED	CYP2D6
NF2	EGFL7	CYP3A4
NFE2L2	EGFR	CYPA5
NOTCH1	EIF1AX	DAXX
NOTCH2	EIF4A2	DDR1
NOTCH3	EIF4E	DDR2
NOTCH4	EML4	DDX3X
NPM1	EP300	DHFR
NRAS	EPCAM	DICER1
NTRK1	EPHA3	DLG2
PALB2	EPHA5	DMNT3A
PBRM1	EPHA7	DNM2
PDCD1LG2	EPHB1	DNMT3A
PDGFRA	ERBB2	DOT1L
PDGFRB	ERBB3	DPYP
PHF6	ERBB4	DUSP2
PIK3CA	ERCC1	EBF1
PIK3CB	ERCC2	ECT2L
PIK3R1	ERCC3	EED

PMS2	ERCC4	EGFR
POLE	ERCC5	EGFR1
POLQ	ERG	EP300
PPP2R1A	ERRFI1	EPCAM
PTCH1	ESR1	EPHA3
PTEN	ETS1	EPHA5
PTPN11	ETV1	EPHA7
RAC1	ETV4	EPHB1
RAD21	ETV5	ERBB2
RAD50	ETV6	ERBB3
RAF1	EWSR1	ERBB4
RB1	EZH2	ERCC1
RET	FAM123B	ERCC2
RHOA	FAM175A	ERCC3
RNF43	FAM46C	ERCC4
ROS1	FANCA	ERCC5
RPL5	FANCC	ERG
RUNX1	FANCD2	ERRFI1
SETBP1	FANCE	ESR1
SETD2	FANCF	ESR2
SF3B1	FANCG	ETV1
SMAD4	FANCI	ETV4
SMARCA4	FANCL	ETV5
SMARCB1	FAS	EWSR1
SMO	FAT1	EXOC2
SOCS1	FBXW7	EXT2
SPOP	FGF1	EZH2
SRC	FGF10	FAM123B
STAG1	FGF14	FAM46C
STAG2	FGF19	FANCA
STAT3	FGF2	FANCB
STAT5B	FGF23	FANCC
STK11	FGF3	FANCD2
SYK	FGF4	FANCE
TERT	FGF5	FANCF
TGFBR2	FGF6	FANCG
TP53	FGF7	FANCL
TSC1	FGF8	FAS
TSC2	FGF9	FAT1
U2AF1	FGFR1	FBXO11
VHL	FGFR2	FBX032
WT1	FGFR3	FCGR2B
YAP1	FGFR4	FGF10
	FH	FGF14
	FLCN	FGF14

Agilent	Illumina	Nonacus
	FLI1	FGF19
	FLT1	FGF23
	FLT3	FGF3
	FLT4	FGF6
	FOXA1	FGFR1
	FOXL2	FGFR2
	FOXO1	FGFR3
	FOXP1	FGFR4
	FRS2	FH
	FUBP1	FIP1L1
	FYN	FLCN
	GABRA6	CLT1
	GATA1	FLT3
	GATA2	FLT4
	GATA3	FOXL2
	GATA4	FOXP1
	GATA6	FRS2
	GEN1	FSTL5
	GID4	FUBP1
	GLI1	GABRA6
	GNA11	GADD45B
	GNA13	GATA1
	GNAQ	GATA2
	GNAS	GATA3
	GPR124	GATA4
	GPS2	GATA6
	GREM1	GLI1
	GRIN2A	GNA11
	GRM3	GNA13
	GSK3B	GNAQ
	H3F3A	GNAS
	H3F3B	GPR124
	H3F3C	GRIN2A
	HGF	GRM3
	HIST1H1C	GSK3B
	HIST1H2BD	GSTM1
	HIST1H3A	GSTP1
	HIST1H3B	GSTT1
	HIST1H3C	H3F3A
	HIST1H3D	HBA1
	HIST1H3E	HBA2
	HIST1H3F	HBB
	HIST1H3G	HDAC1
	HIST1H3H	HDAC2
	HIST1H3I	HDAC4

	HIST1H3J	HDAC7
	HIST2H3A	HGF
	HIST2H3C	HNF1A
	HIST2H3D	HNF1B
	HIST3H3	HRAS
	HLA-A	HSD3B1
	HLA-B	HSP90AA1
	HLA-C	ID3
	HNF1A	IDH1
	HNRNPK	IDH2
	HOXB13	IGF1R
	HRAS	IGF2
	HSD3B1	IKBKE
	HSP90AA1	IKZF1
	ICOSLG	IKZF2
	ID3	IKZF3
	IDH1	IL2RA
	IDH2	IL2RB
	IFNGR1	IL2RG
	IGF1	IL7R
	IGF1R	INHBA
	IGF2	INPP4B
	IKBKE	INPP5D
	IKZF1	IRAK4
	IL10	IRF1
	IL7R	IRF2
	INHA	IRF4
	INHBA	IRF8
	INPP4A	IRS2
	INPP4B	ITCH
	INSR	JAK1
	IRF2	JAK2
	IRF4	JAK3
	IRS1	JARID2
	IRS2	JUN
	JAK1	KDM2B
	JAK2	KDM5A
	JAK3	KDM5C
	JUN	KDM6A
	KAT6A	KDR
	KDM5A	KEAP1
	KDM5C	KEL
	KDM6A	KIT
	KDR	KLHL6
	KEAP1	KMT2A

Agilent	Illumina	Nonacus
	KEL	KMT2B
	KIF5B	KMT2C
	KIT	KRAS
	KLF4	LAMA2
	KLHL6	LCK
	KMT2B	LEF1
	KMT2C	LMO1
	KMT2D	LRP1B
	KRAS	LTK
	LAMP1	LYN
	LATS1	LYST
	LATS2	LZTR1
	LMO1	MAGI2
	LRP1B	MAP2K1
	LYN	MAP2K2
	LZTR1	MAP2K4
	MAGI2	MAP3K1
	MALT1	MAPK1
	MAP2K1	MCL1
	MAP2K2	MDM2
	MAP2K4	MDM4
	MAP3K1	MECOM
	MAP3K13	MED12
	MAP3K14	MED13
	MAP3K4	MEF2B
	MAPK1	MEN1
	MAPK3	MET
	MAX	MGMT
	MCL1	MITF
	MDC1	MLH1
	MDM2	MLL
	MDM4	MLL2
	MED12	MLLT10
	MEF2B	MPL
	MEN1	MRE11A
	MET	MSH2
	MGA	MSH3
	MITF	MSH6
	MLH1	MST1R
	MLL	MTHFR
	MLLT3	MTOR
	MPL	MUC16
	MRE11A	MUTYH
	MSH2	MYC
	MSH3	MYCL1

	MSH6	MYCN
	MST1	MYD88
	MST1R	MYST3
	MTOR	NBN
	MUTYH	NCOR1
	MYB	NCSTN
	MYC	NEK2
	MYCL1	NELL2
	MYCN	NF1
	MYD88	NF2
	MYOD1	NFE2L2
	NAB2	NFKBIA
	NBN	NKX2-1
	NCOA3	MOTCH1
	NCOR1	NOTCH2
	NEGR1	NPM1
	NF1	NQO1
	NF2	NRAS
	NFE2L2	NRG1
	NFKBIA	NSD1
	NKX2-1	NT5C2
	NKX3-1	NTRK1
	NOTCH1	NTRK2
	NOTCH2	NTRK3
	NOTCH3	NUP93
	NOTCH4	PAG1
	NPM1	PAK3
	NRAS	PALB2
	NRG1	PARK2
	NSD1	PAX5
	NTRK1	PBRM1
	NTRK2	PC
	NTRK3	PCGF2
	NUP93	PCCD1
	NUTM1	PDCD1
	PAK1	PDCD1LG2
	PAK3	PDGFB
	PAK7	PDGFRA
	PALB2	PDGFRB
	PARK2	PDK1
	PARP1	PHF6
	PAX3	PHOX2B
	PAX5	PICK3R1
	PAX7	PIK3C2B
	PAX8	PIK3C3

Agilent	Illumina	Nonacus
	PBRM1	PIK3CA
	PDCD1	PIK3CB
	PDCD1LG2	PIK3CD
	PDGFRA	PIK3CG
	PDGFRB	PIK3R1
	PDK1	PIK3R2
	PDPK1	PLCG2
	PGR	PLK1
	PHF6	PMS1
	PHOX2B	PMS2
	PIK3C2B	POLD1
	PIK3C2G	POT1
	PIK3C3	PPM1L
	PIK3CA	PPP2R1A
	PIK3CB	PRDM1
	PIK3CD	PREX2
	PIK3CG	PRF1
	PIK3R1	PRKAR1A
	PIK3R2	PRKCI
	PIK3R3	PRKDC
	PIM1	PSMB1
	PLCG2	PSMB2
	PLK2	PSMB5
	PMAIP1	PSMD1
	PMS1	PSMD2
	PMS2	PTCH1
	PNRC1	PTEN
	POLD1	PTGFR
	POLE	PTPN11
	PPARG	PTPN2
	PPM1D	PTPN6
	PPP2R1A	PTPRO
	PPP2R2A	QKI
	PPP6C	RAC1
	PRDM1	RAD21
	PREX2	RAD50
	PRKAR1A	RAD51
	PRKCI	RAF1
	PRKDC	RANBP2
	PRSS8	RARA
	PTCH1	RARB
	PTEN	RARG
	PTPN11	RASGEF1A
	PTPRD	RB1
	PTPRS	RBM10

	PTPRT	RECQL4
	QKI	RELN
	RAB35	RET
	RAC1	RHOA
	RAD21	RICTOR
	RAD50	RNF43
	RAD51	ROS1
	RAD51B	RPS6KB1
	RAD51C	RPTOR
	RAD51D	RRM1
	RAD52	RUNX1
	RAD54L	RUNX1T1
	RAF1	RXRA
	RANBP2	RXRB
	RARA	RXRG
	RASA1	SBDS
	RB1	SDHA
	RBM10	SDHB
	RECQL4	SDHC
	REL	SDHD
	RET	SEPT.9
	RFWD2	SERP2
	RHEB	SETBP1
	RHOA	SETD2
	RICTOR	SF3B1
	RIT1	SGK1
	RNF43	SH2D1A
	ROS1	SHH
	RPS6KA4	SHOC2
	RPS6KB1	SLC22A1
	RPS6KB2	SLC22A2
	RPTOR	SLC31A1
	RUNX1	SLC34A2
	RUNX1T1	SLC45A3
	RYBP	SLCO1B1
	SDHA	SLIT2
	SDHAF2	SMAD2
	SDHB	SMAD3
	SDHC	SMAD4
	SDHD	SMAD7
	SETBP1	SMARCA4
	SETD2	SMARCB1
	SF3B1	SMC1A
	SH2B3	SMC3
	SH2D1A	SMO

Agilent	Illumina	Nonacus
	SHQ1	SNCAIP
	SLIT2	SOCS1
	SLX4	SOS1
	SMAD2	SOX10
	SMAD3	SOX2
	SMAD4	SOX9
	SMARCA4	SPEN
	SMARCB1	SPOP
	SMARCD1	SPRED1
	SMC1A	SPTA1
	SMC3	SRC
	SMO	SRSF2
	SNCAIP	STAG2
	SOCS1	STAT3
	SOX10	STAT4
	SOX17	STAT5A
	SOX2	STAT5B
	SOX9	STIL
	SPEN	STK11
	SPOP	STMN1
	SPTA1	STX11
	SRC	STXBP2
	SRSF2	SUFU
	STAG1	SUZ12
	STAG2	SYK
	STAT3	TAF1
	STAT4	TAS2R38
	STAT5A	TEK
	STAT5B	TEKT4
	STK11	TERC
	STK40	TERT
	SUFU	TET2
	SUZ12	TGFBR2
	SYK	TLE1
	TAF1	TLE4
	TBX3	TMPRSS2
	TCEB1	TNFAIP3
	TCF3	TNFRSF14
	TCF7L2	TNFRSF17
	TERC	TNFRSF19
	TERT	TOP1
	TET1	TOP2A
	TET2	TP53
	TFE3	TP63
	TFRC	TPMT

	TGFBR1	TRAF2
	TGFBR2	TRAF3
	TMEM127	TRAF5
	TMPRSS2	TRRAP
	TNFAIP3	TSC1
	TNFRSF14	TSC2
	TOP1	TSHR
	TOP2A	TTF1
	TP53	TUBB3
	TP63	TYK2
	TRAF2	TYMS
	TRAF7	U2AF1
	TSC1	UGT1A1
	TSC2	UNC13D
	TSHR	VEGFA
	U2AF1	VHL
	VEGFA	WEE1
	VHL	WISP3
	VTCN1	WRN
	WISP3	WT1
	WT1	XIAP
	XIAP	XPC
	XPO1	XPO1
	XRCC2	XRCC1
	YAP1	YAP1
	YES1	YES1
	ZBTB2	ZAP70
	ZBTB7A	ZBED4
	ZFHX3	ZBTB2
	ZNF217	ZMYM3
	ZNF703	ZNF217
	ZRSR2	ZNF703
		ZRSR2

Appendix 6:

Maxwell® DNA extraction supplier protocol

Maxwell® RSC DNA FFPE Kit

Instructions for Use of Products AS1450 and ASB1450.

Promega

Quick Protocol

Preprocessing FFPE Section Samples

1. Place the FFPE tissue section into a 1.5ml microcentrifuge tube. If using slide-mounted tissue sections, scrape the section off the slide using a clean razor blade. Tap or centrifuge tube briefly to collect the sample at the bottom of the tube.
   Note: 5–10 micron-thick tissue sections ranging in size from 20mm^2 to 200mm^2 for a total of up to 2.0mm^3 of tissue can be used.
2. Add 300µl of Mineral Oil to the sample tubes. Vortex for 10 seconds.
3. Heat the samples at 80°C for 2 minutes, then place samples at room temperature while the master mix is prepared.
4. Prepare a master mix of the Lysis Buffer, Proteinase K Solution and Blue Dye as shown below:

Reagent	Amount per Reaction	Reactions (number + 2)	Total
Lysis Buffer	224µl	n + 2	224 × (n + 2)µl
Proteinase K	25µl	n + 2	25 × (n + 2)µl
Blue Dye	1µl	n + 2	1 × (n + 2)µl

For fewer than six samples, prepare enough master mix for n + 1 samples.
Note: Use the master mix within 1 hour of preparation. Master mix cannot be stored for later use.

6. Add 250µl of master mix to each sample tube, and vortex for 5 seconds.
7. Centrifuge sample tubes at 10,000 × *g* for 20 seconds to separate layers. If a pellet is present in the aqueous layer (lower, blue layer), gently mix the aqueous phase with a pipette to resuspend the pellet.
8. Transfer the sample tubes to a 56°C heat block and incubate for 30 minutes.
9. Transfer the sample tubes to an 80°C heat block and incubate for 4 hours.
10. Remove the sample tubes from the heat block, and allow the samples to cool to room temperature for 5 minutes.
11. Add 10µl RNase A Solution to the aqueous (blue) phase in each sample tube. Mix by pipetting.
12. Incubate for 5 minutes at room temperature (15–30°C). During the incubation, begin cartridge preparation.
13. Centrifuge the sample tubes at full speed in a microcentrifuge for 5 minutes.
14. Immediately transfer the blue, aqueous phase containing the DNA to well #1 of a Maxwell® FFPE Cartridge.

Additional protocol information in Technical Manual #TM437, available online at: **www.promega.com**

Method Setup and Cartridge Preparation

Maxwell® RSC Method Setup

Before using the Maxwell® RSC DNA FFPE Kit for the first time, the FFPE DNA method must be installed on your instrument. The method is available at: **www.promega.com/resources/software-firmware/**

See the *Maxwell® RSC Methods Installation Technical Manual* #TM435 for instructions.

Cartridge Preparation

1. Place the cartridges to be used in the deck tray with well #1 (the largest well in the cartridge) farthest away from the Elution Tubes. Press down on the cartridge to snap it into position. Ensure both cartridge ends are fully seated in the deck tray. Carefully peel back the seal so that the entire seal is removed from the top of the cartridge. Ensure that all sealing tape and any residual adhesive are removed from the cartridges. **Caution:** Handle cartridges with care. Seal edges may be sharp.
2. Place one plunger into well #8 of each cartridge. Well #8 is the well closest to the Elution Tube.
3. Place an empty Elution Tube into the Elution Tube position for each cartridge in the deck tray.
4. Add 50µl of Nuclease-Free Water to the bottom of each Elution Tube. The Elution Tubes must stay open during the RNA purification.
   Note: Use only the CSC/RSC Plungers, Elution Tubes and Nuclease-Free-Water supplied with the Maxwell® RSC DNA FFPE Kit. Plungers for Maxwell® 16 LEV kits are not compatible with Maxwell® RSC Instruments. Other elution tubes may not be compatible with Maxwell® RSC Instruments and may affect performance. Use of other elution buffers may impact DNA purification performance or downstream use.
5. Proceed to the next section, Instrument Run on Maxwell® RSC Instruments.

Setup and configuration of deck trays. Nuclease-Free Water is added to the elution tubes as shown. Plungers are in well #8 of the cartridge.

Instrument Run on Maxwell® RSC Instruments (Cat.# AS4500, AS8500)

1. Follow the instrument run instructions in the *Maxwell® RSC DNA FFPE Kit Technical Manual* #TM437.
2. Refer to the *Maxwell® RSC Instrument Operating Manual* #TM411 or *Maxwell® RSC 48 Instrument Operating Manual* #TM510 for detailed information.

Additional protocol information in Technical Manual #TM437, available online at: **www.promega.com**

PROMEGA CORPORATION • 2800 WOODS HOLLOW ROAD • MADISON, WI 53711-5399 USA • TELEPHONE 608-274-4330
www.promega.com

Appendix 7:

Qubit™ Flex fluorometer DNA quantification supplier protocol

Extract below is taken from the Qubit™ Flex fluorometer supplier protocol (pages 36-37, 42) available at: https://assets.thermofisher.com/TFS-Assets/LSG/manuals/MAN0018186_Qubit_Flex_Fluorometer_UG.pdf

Read samples

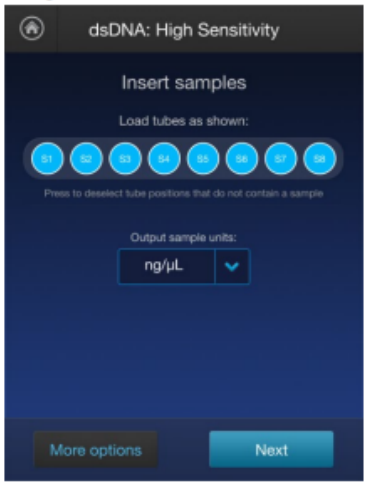
Before you begin

- Calibrate the Qubit™ Flex Fluorometer as described on page 25. (Run the appropriate standards or accept the values from the previous calibration.)
- Prepare the samples. Refer to the instructions provided with the assay.

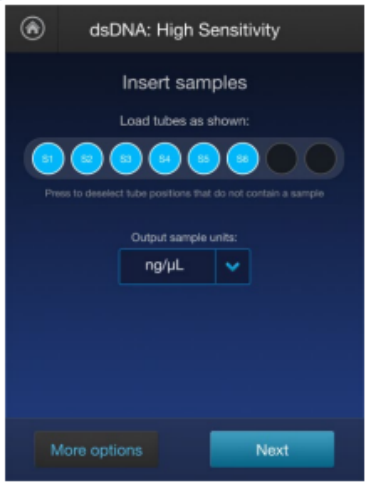
Note: Incubate the samples for the appropriate amount of time after mixing them with the working solution (2 minutes for the Qubit™ DNA and RNA assays, 15 minutes for the Qubit™ protein assay).

Insert samples

1. When prompted, load the tube strip containing the samples as shown in the **Insert samples** screen. If you have fewer than 8 samples, press to deselect the tube positions that do not contain a sample.

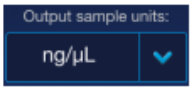


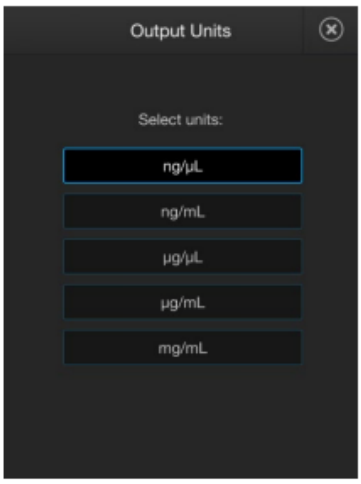
All 8 tubes contain samples



No sample in positions S7 and S8

2. Press **Output sample units** to open the **Output Units** screen, then select the desired units.

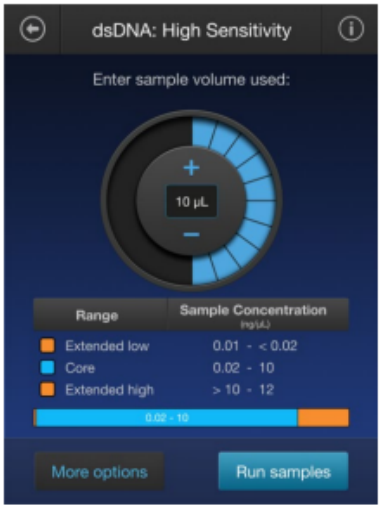




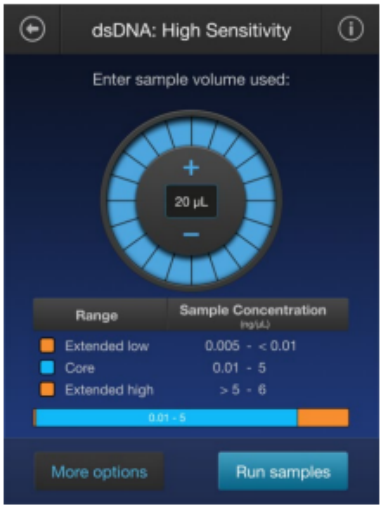
3. Press **Next** to go to the Sample volume screen.
4. In the **Sample volume** screen, enter the **sample volume** added to the assay tube (between 1 and 20 µL).

   You can enter the volume directly in the sample volume text box, use the + and – buttons, or adjust the sample volume wheel.

   When you enter the sample volume, the assay range information on the screen automatically changes to reflect the new core and extended accuracy ranges based on the sample volume.



dsDNA HS Assay range
for 10 µL sample volume



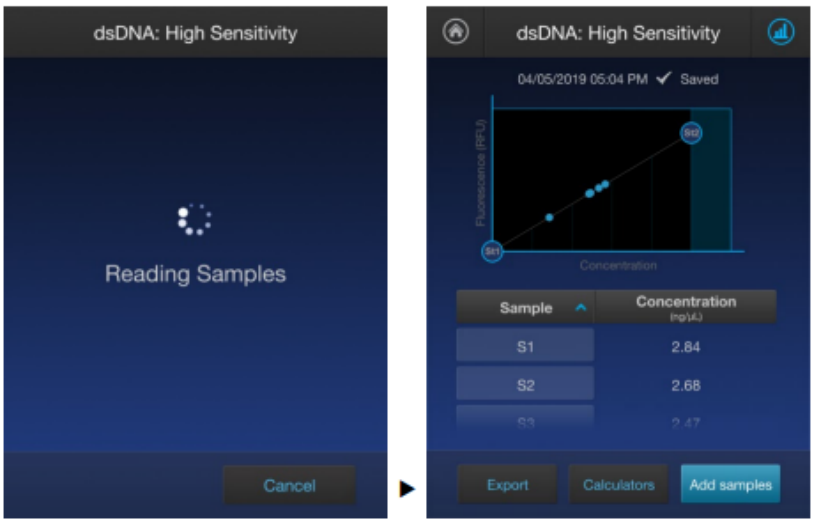
dsDNA HS Assay range
for 20 µL sample volume

Note: The sample volume used (1–20 µL) changes the assay accuracy range. For highest accuracy, use the maximum sample volume that would keep the concentration measurements within the core range.

If the sample concentration is outside of what the assay can accurately quantify, a different sample volume or assay may be required.

Run Samples

1. Press **Run samples**. The reading takes approximately 3 seconds and the results are displayed in graph view in the Results screen (see "Results", page 43).



2. To display the results in list view, press the **Graph** button to unselect it.

   The Results screen lists the concentration of each original sample using the output units selected at the beginning of the assay.

Note: By default, the Results screen displays the measurements in graph view. However, the graph settings are "sticky", so that if you close the graph, the next time anyone runs an assay, the graph view is hidden and the results are shown in list form.

3. To run more samples, press **Add samples**, and repeat the procedure.

Appendix 8:

Illumina TruSight™ Oncology 500 panel library preparation supplier protocol

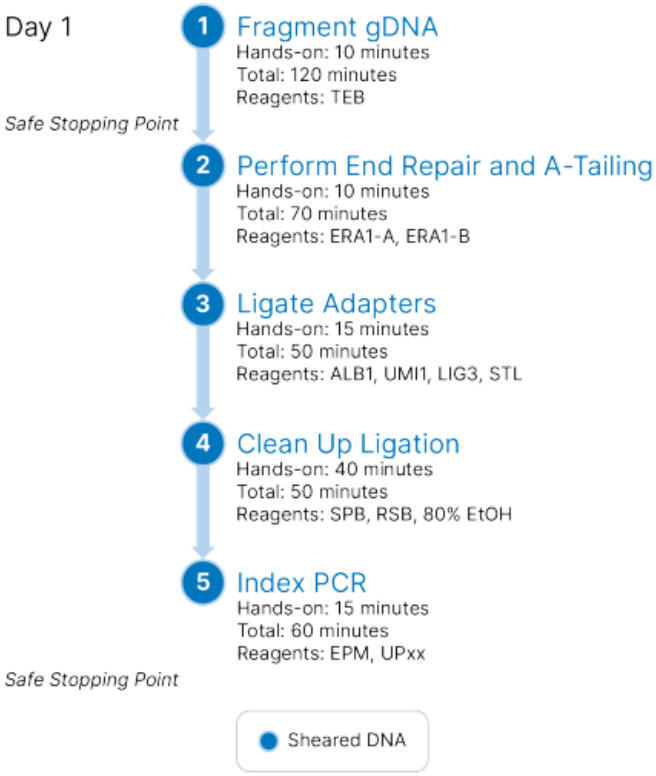
Overview of Illumina TruSight™ Oncology 500 panel library preparation protocol is shown below (pages 7 and 8); full protocol available online: https://support.illumina.com/downloads/trusight-oncology-500-reference-guide-1000000067621.html

Library Prep DNA Only Workflow

The following diagram illustrates the recommended DNA only library preparation workflow using a TruSight Oncology 500 Kit. Safe stopping points are marked between steps. Hands-on and total times are approximate and based on eight DNA samples. Times include degassing the Covaris ultrasonicator.

RNA and DNA libraries may be prepared simultaneously. Illumina recommends performing the TruSight Oncology 500 Kit assay workflow according to the following schedule:

Figure 1 TruSight Oncology 500 Kit DNA Only Workflow (Part 1)



Enrichment DNA Only Workflow

The following diagram illustrates the recommended DNA only enrichment workflow using a TruSight Oncology 500 Kit. Safe stopping points are marked between steps. Hands-on and total times are approximate and based on eight DNA samples.

Figure 2 TruSight Oncology 500 Kit DNA Only Workflow (Part 2)


Day 1 (continued)
1
Set Up First Hybridization
Hands-on: 15 minutes
Total: overnight
Reagents: TCA1, TCB1, OPD2
Overnight Hybridization
Day 2
2
Capture Targets One
Hands-on: 60 minutes
Total: 100 minutes
Reagents: SMB, EEW, EE2, HP3, ET2
3
Set Up Second Hybridization
Hands-on: 10 minutes
Total: 1.5-4 hours
Reagents: TCA1, TCB1, OPD2
4
Capture Targets Two
Hands-on: 25 minutes
Total: 60 minutes
Reagents: SMB, RSB, EE2, HP3, ET2
Safe Stopping Point
5
Amplify Enriched Library
Hands-on: 5 minutes
Total: 60 minutes
Reagents: PPC3, EPM
6
Clean Up Amplified Enriched Library
Hands-on: 30 minutes
Total: 40 minutes
Reagents: SPB, RSB, 80% EtOH
Safe Stopping Point
7
Quantify Libraries (Optional)
8
Normalize Libraries
Hands-on: 40 minutes
Total: 50 minutes
Reagents: LNA1, LNB1, LNW1, HP3, LNS1, EE2
Safe Stopping Point
Enrichment

Appendix 9:

Agilent SureSelect Community Design Glasgow Cancer Core panel library preparation supplier protocol

DAY 1 Stage 1: Enzymatic Fragmentation Est time ~ 40 mins

1) Thaw 5x SureSelect Fragmentation Buffer from the Low Input enzymatic fragmentation kit box and the first 16 samples to be processed at room temperature. Gently and briefly vortex samples and spin down before placing on ice. Gently and briefly vortex 5x SureSelect Fragmentation Buffer and spin down before placing on ice.

2) Thaw the End Repair A-Tailing Buffer from the SSEL XT HS Reagent Kit. Visually inspect the solution for any precipitate and if necessary vortex until it dissolves. Vortex and spin down before placing on ice.

3) Thaw the Ligation Buffer from the SSEL XT HS Reagent Kit. Note the ligation buffer may need to be briefly warmed in your hand to ensure all the white precipitate is re-dissolved. Vortex for 20 seconds at high speed to mix and then spin down.

4) Prepare the fragmentation master mix as follows:

Reagent	Volume for 16 Reactions
5X SureSelect Fragmentation Buffer	36 µl
SureSelect Fragmentation Enzyme	18 µl
Total	**54 µl**

5) Mix by pipetting up and down, then briefly spin down.

6) Add 3µl of the fragmentation master mix to each well containing 7µl of input gDNA. Mix well by pipetting up and down. Seal the wells and briefly spin down the samples. Keep the plate chilled as far as possible while you are doing this.

7) Place the sample plate in the thermocycler and run the Enz fragment program as below:

Step	Temperature	Time
Step 1	37°C	15 Minutes
Step 2	65°C	5 Minutes
Step 3	4°C	∞ Hold

*Use a reaction volume of 10µl and Lid temp of 105°C

8) Take a ≥2.2ml aliquot of AMPureXP beads out of the fridge to come to room temperature for later in the workflow.

9) When the Enz fragment program reaches the 4°C hold step, remove the sample plate from the thermocycler. Add 40µl of nuclease-free water to each sample well, seal, spin down and place the sample plate on ice. If short-term storage is required store the samples at 4°C.

Note: This is not a stopping point, proceed to end repair and A-tailing as soon as possible.

Stage 2: Prepare Ligation Master Mix Est time ~ 5 mins

10) Remove the T4 DNA Ligase (Blue Cap) from the SSEL XT HS Reagent Kit from storage at -20°C, flick to mix and spin down.

11) Prepare the ligation master mix in a 1.5 ml Eppendorf as follows:

Reagent	Volume for 16 reactions
Ligation Buffer	405 µl
T4 DNA Ligase (Blue Cap)	35.2 µl
Total	**440.2 µl**

12) Mix well by pipetting up and down 15-20 times, pipette slowly as the buffer is very viscous. Spin down briefly. The ligation master mix needs to equilibrate at room temperature for at least 30 mins. While this is happening, carry out the end repair and A-tailing.

Stage 3: End Repair and A-tailing Est time ~ 45 mins

13) Remove the End Repair A-Tailing Enzyme Mix (Orange Cap) from the SSEL XT HS Reagent Kit from storage at -20°C, flick to mix and spin down.

14) Prepare the end repair A-tailing master mix in a 1.5 ml Eppendorf as follows:

Reagent	Volume for 16 reactions
End Repair A-Tailing Buffer	281.5 µl
End Repair A-Tailing Enzyme Mix (Orange Cap)	70.5 µl
Total	**352 µl**

15) Mix well by pipetting up and down 15-20 times, pipette slowly as the buffer is very viscous. Spin down briefly.

16) Add 20µl of end repair-A tailing master mix to each sample well containing 50µl of fragmented gDNA. Mix by pipetting up and down 15-20 times using a pipette set to 60µl.

17) Seal the wells, briefly spin down the samples, then transfer the plate to the thermocycler and run the ER-ATail program as below:

Step	Temperature	Time
Step 1	20°C	15 Minutes
Step 2	72°C	15 Minutes
Step 3	4°C	∞ Hold

*Use a reaction volume of 70µl and Lid temp of 105°C

While the program is running, remove the Adaptor Oligo Mix (white cap) from the SSEL XT HS Reagent Kit from storage at -20°C and allow to thaw. Gently and briefly vortex and spin down. Reagent Volume for 16 reactions Ligation Buffer 405 µl T4 DNA Ligase (Blue Cap) 35.2 µl Total 440.2 µl Reagent Volume for 16 reactions End Repair A-Tailing Buffer 281.5 µl End Repair A-Tailing Enzyme Mix (Orange Cap) 70.5 µl Total 352 µl

Stage 4: Ligate Adaptors Est time ~ 50 mins (incl. 25 mins hands-off time)

18) Once the ER-ATail program has finished, allow the plate to come to room temperature, then add 25µl of ligation master mix to each sample well. Pipette up and down 15 times to mix the sample with a pipette set to 85µl. Once all samples are mixed seal the plate and briefly spin down.

19) Add 5µl of the Adaptor Oligo Mix (white cap) to each sample. Pipette up and down 15 times to mix the sample with a pipette set to 85µl. Once all samples are mixed seal the plate and briefly spin down.

20) Transfer the plate to the thermocycler and run the Ligation program as below:

Step	Temperature	Time
Step 1	20°C	30 Minutes
Step 2	4°C	∞ Hold

*Use a reaction volume of 100µl and Lid temp of 105°C

While the program is running, prepare a fresh dilution of 70% Ethanol in a 50ml Falcon tube by combining 7ml 100% ethanol with 3ml Nuclease free water.

Stage 5: Post-Ligation Clean Up Est time ~ 75 mins

21) Once the ligation program is complete, move the sample plate to the bench at room temperature and remove the strip caps.

22) Vortex the aliquot of AMPureXP beads (see stage 1: 8) above) for 30 secs, until the mixture is homogeneous.

23) Add 80µl of homogeneous AMPure beads to each sample well, pipetting up and down 15 times to mix. (Keep the remaining beads at room temperature for later.)

24) Incubate the sample and bead mixture for 5 mins at room temperature.

25) Place the sample plate on the magnetic separation device and wait for the solution to clear, this can take 5 – 10 mins.

26) Keep the plate on the magnetic separator; with a pipette set for 200µl carefully remove and discard the cleared liquid from the wells. Do not touch the beads while removing the solution.

27) Carry out the following steps twice:

• Keep the plate on the magnetic separator; Add 200µl of freshly prepared 70% ethanol to each sample well.

• Keep the plate on the magnetic separator; Wait 1 min to allow any disturbed beads to settle, then remove and discard the ethanol.

28) Seal the wells of the sample plate and briefly centrifuge the sample plate to collect any residual liquid.

29) Return the plate to the magnetic separator and wait for 1 min. Remove any residual ethanol using a P10 pipette.

30) Run the 37C program on the thermocycler (37C for 3 minutes) and place the sample plate on it for no more than 60 secs, to completely dry the beads.

Note: over-drying will lead to loss of yield.

31) Add 35µl of nuclease-free water to each sample well. Seal the sample plate then mix well on the vortex.

32) Briefly spin the plate to collect the liquid and leave the plate to incubate at room temperature for 2 mins.

33) Return the plate to the magnetic separator and leave the solution to clear for 5 mins.

34) Keep the plate on the magnetic separator; Remove the cleared supernatant (approx. 34.5µl) to a fresh PCR plate and keep on ice. Discard the beads.

Stage 6: Amplify the Adaptor-Ligated Library Est time ~ 60 mins (incl. 35 mins hands-off time)

35) Thaw the Index Primers (black-capped tubes) from the SSEL XT HS Reagent Kit. For the first set of 16 samples this will be index primers 1-16 (A01–H02), for the second set of samples this will be index primers 17-32 (A03-A04). Record which primer you are assigning to which sample.

36) Thaw the Forward Primer (brown cap), 5x Herculase II Reaction Buffer (clear cap) and 100mM dNTP Mix (green cap) from the SSEL XT HS Reagent Kit.

37) Once all reagents have thawed, gently and briefly vortex and spin down. Keep on ice until needed.

38) Remove the Herculase II Fusion DNA Polymerase enzyme (red cap) from the SSEL XT HS Reagent Kit from storage at -20°C, flick to mix and spin down. Keep on ice.

39) Prepare the pre-capture PCR master mix in a 1.5 ml Eppendorf as follows:

Reagent	Volume for 16 reactions
5× Herculase II Reaction Buffer (clear cap)	170µl
100 mM dNTP Mix (green cap)	8.5µl
Forward Primer (brown cap)	34µl
Herculase II Fusion DNA Polymerase (red cap)	17µl
Total	229.5µl

40) Mix well by pipetting up and down and spin down briefly. Keep on ice.

41) Add 13.5µl of PCR reaction mixture to each sample well containing adaptor ligated library.

42) Add 2µl of the appropriate SureSelect XT HS Index Primer to each reaction. Cap the wells and vortex at high speed for 5 seconds. Spin the reaction plate briefly to collect the liquid.

43) Start the thermocycler PreCap-PCR program as below:

Segment	Number of Cycles	Temperature	Time
1	1	98°C	2 Minutes
2	12	98°C	30 Seconds
		60°C	30 Seconds
		72°C	1 Minute
3	1	72°C	5 Minutes
4	1	4°C	∞ Hold

*Use a reaction volume of 50µl and Lid temp of 105°C

Once the lid has heated and the wells have just reached 98°C, immediately add the plate and close the lid.

Stage 7: Pre-Capture PCR Clean Up and Normalisation Est time ~ 60 mins

44) Once the PreCap-PCR program is complete, move the sample plate to the bench at room temperature and remove the strip caps.

45) Vortex the aliquot of AMPureXP beads (see stage 1: 8)) for 30 secs, until the mixture is homogeneous.

46) Add 50µl of homogeneous AMPure beads to each sample well, pipetting up and down 15 times to mix.

47) Incubate the sample and bead mixture for 5 mins at room temperature.

48) Place the sample plate on the magnetic separation device and wait for the solution to clear, this can take 5 – 10 mins.

49) Keep the plate on the magnetic separator; with a pipette set for 200µl carefully remove and discard the cleared liquid from the wells. Do not touch the beads while removing the solution.

50) Carry out the following steps twice: • Keep the plate on the magnetic separator; Add 200µl of freshly-prepared 70% ethanol to each sample well. • Keep the plate on the magnetic separator; Wait 1 min to allow any disturbed beads to settle, then remove and discard the ethanol.

51) Seal the wells of the sample plate and briefly centrifuge the sample plate to collect any residual liquid.

52) Return the plate to the magnetic separator and wait for 1 min. Remove any residual ethanol using a P10 pipette.

53) Run the 37C program on the thermocycler (37C for 3 minutes) and place the sample plate on it for no more than 60 secs, to completely dry the beads.

Note: over-drying will lead to loss of yield

54) Add 15µl of nuclease-free water to each sample well. Seal the sample plate then mix well on the vortex.

55) Briefly spin the plate to collect the liquid and leave the plate to incubate at room temperature for 2 mins.

56) Return the plate to the magnetic separator and leave the solution to clear for 3 mins.

57) Keep the plate on the magnetic separator; Remove the cleared supernatant (approx. 15µl) to a fresh PCR plate and keep on ice. The remaining beads can be discarded.

58) Use 1µl of the library and 9µl nuclease-free water to make a 1:10 dilution aliquot for QC. Determine the concentration of each library using this dilution and your preferred quantitation method (e.g. Bioanalyzer, TapeStation, Picogreen etc). Note that methods that visualise the library (e.g. Tapestation) allow additional QC of features such as insert size distribution and adapter contamination.

59) Dilute each library (or an aliquot of each library) to 125ng/µl using nuclease-free water, store in the fridge overnight.

DAY 2 Stage 8: Preparing pre-capture pools Est time ~ 10 mins

60) Pre-capture libraries must be pooled before Hybridisation with XT2 baits. Each hybridisation requires a pool containing 1500ng of library DNA in 12µl. To prepare one pool of 16 samples, twice the required volume is made up, in order to avoid micropipetting. Pool 16 samples into a 0.5ml Eppendorf as follows:

Number of libraries per pool	Individual library amount required (ng)	Amount of 125ng/µl normalisation required (µl)	Total DNA in pool (ng)	Total pool volume (µl)
16	187.5	1.5	3000	24

61) Close the tubes, briefly vortex, spin down and leave on ice.

Stage 9: Hybridisation Est time ~ 80 mins

62) Thaw the SureSelect XT HS and XT Low Input Blocker Mix (blue cap) and SureSelect RNase Block (purple cap) from the SSEL XT HS Reagent Kit. Gently and briefly vortex and spin down and keep on ice.

63) Thaw the Fast Hybridization Buffer from the SSEL XT HS Reagent Kit at room temperature.

64) Thaw the SureSelect XT2 Custom 0.5-2.9Mb baits (Baitset ID: 3184181) from -80°C storage, keeping it on ice the whole time. Note: the capture library is very sensitive to temperature. Once it is thawed, briefly spin down.

65) Transfer 12µl of the DNA pool to a 0.2ml Eppendorf and add 5µl of SureSelect XT HS and XT Low Input Blocker Mix (blue cap). Vortex at high speed for 5 seconds. Spin down to collect the liquid.

66) Start the fast hyb program (as below) and transfer the tube to the thermocycler.

Segment	Number of Cycles	Temperature	Time
1	1	95°C	5 minutes
2	1	65°C	10 minutes
3	1	65°C	1 minute **(PAUSE cycler here)**
4	60	65°C	1 minute
		37°C	3 seconds
5	1	65°C	∞ Hold

Use a reaction volume of 30µl and a heated lid (105°C)

Start a timer for 15 mins. Note the thermocycler must be paused during segment 3 of the fast hyb program (see section 3.3), to allow additional reagents to be added.

67) While the first 15 mins of the fast hyb program are running, prepare a 0.5ml Eppendorf of hybridisation mix at room temperature. Add the reagents in the order shown:

Reagent	Volume for 1 reaction
Nuclease-free water	4.5µl
SureSelect RNase Block (Purple cap)	0.5µl
SureSelect Fast Hybridisation Buffer	6µl
Capture Baits	2µl
Total	**13µl**

68) Briefly vortex the hybridisation mix and spin down.

69) When the fast hyb program reaches step 3, pause it but do not remove the tube. Complete the next few steps as quickly as possible to minimise evaporation. Note do not remove the tube from the thermocycler throughout this process, take care as the base and thermocycler lid will be hot.

70) Open the thermocycler and open the library pool tube. Add the full 13µl of hybridisation mix to the library tube and mix by pipetting up and down 10 times but taking care to not introduce bubbles.

71) Ensure the tube is properly re-sealed (Note failure to do this will result in excessive evaporation and hybridisation failure). Close the thermocycler and resume the fast hyb program.

72) While the fast hyb program is running, prepare the magnetic beads as follows.

73) Take the Dynabeads MyOne Streptavidin T1 magnetic beads from 4°C storage and vortex for 60 secs to homogenise.

74) Put 50µl beads into a 0.2ml Eppendorf.

75) Carry out the following steps three times:

• Add 200µl of SureSelect Binding Buffer to the bead-containing tube.

• Mix by pipetting up and down 20 times.

• Place the plate onto the magnetic separator

• Wait for 5 mins to allow the beads to settle, then remove and discard the supernatant.

• Remove the plate from the magnetic separator.

76) Resuspend the beads in 200µl of binding buffer and leave on the bench at room temperature until the fast hyb program is complete.

Stage 10: Capture Washes Est time ~ 90 mins

77) When the fast hyb program reaches the 65°C hold, remove the tube and immediately transfer the hybridisation reaction (~30µl) to a tube containing washed streptavidin beads, pipette up and down to mix.

Note a 0.2ml Eppendorf should comfortably fit 230µl and still allow the lid to close cleanly.

Take care if using other consumables.

78) Incubate the tube on a mixer, mixing vigorously at 1400-1800 rpm at room temperature for 30 mins. While this is incubating prepare wash buffer 2 as follows.

79) Make 6 aliquots of 200µl each of wash buffer 2 in 0.2ml Eppendorfs. Close the tubes and place them in the thermocycler and run the Wash program (70C hold with heated lid at 105C).

80) When the sample/bead mixture has finished its 30-min incubation, take it off the shaker and briefly spin down.

81) Place the tubes on the magnetic separator and leave for 5-10mins until the solution has cleared. Remove and discard the supernatant.

82) Resuspend the beads in 200µl of SureSelect Wash buffer 1. Mix by pipetting up and down 20 times until the beads are fully resuspended.

83) Place the tube on the magnetic separator and wait 1 minute for the solution to clear, then remove and discard the supernatant.

84) The bead captured pooled library must now be washed with prewarmed wash buffer 2. Start the wash program (70C hold with heated lid at 105C) on the second thermocycler block (if available) and carry out the following steps 6 times:

• Resuspend the beads in 200µl of 70°C pre-warmed wash buffer 2, pipetting up and down 15 times until the beads are resuspended.

• Close the tube then vortex at high speed for 8 seconds. Very briefly spin the tubes to collect the liquid without pelleting the beads. The beads must remain in suspension.

• Incubate for 5 mins at 70°C on the second thermocycler block. • After 5 mins move the tube to the magnetic separator at room temperature.

• Wait 1 minute for the solution to clear, then remove and discard the supernatant.

85) During the wash incubations complete the following tasks:

• Chill a 1.5mL Eppendorf tube of nuclease-free water on ice.

• Take the ≥55µl aliquot of AMPureXP beads out of the fridge to come to room temperature for later in the workflow.

• Thaw the 5x Herculase II Reaction Buffer (clear cap), 100mM dNTP Mix (green cap) and SureSelect Post-Capture Primer Mix (clear cap) from the SSEL XT HS Reagent Kit. Gently and briefly vortex each tube, spin down and keep on ice.

• Remove the Herculase II Fusion DNA Polymerase enzyme (red cap) from the SSEL XT HS Reagent Kit from storage at -20°C, flick to mix and spin down. Keep on ice.

86) Once the washing procedure has been completed, verify that all residual wash buffer has been removed. Add 25µl of nuclease free water to the captured sample and pipette up and down 10 times to resuspend the beads.

Stage 11: Post-Capture Amplification Est time ~ 45 mins (incl. 20 mins hands-off time)

87) Prepare the post-capture PCR master mix in a 0.2 ml Eppendorf as follows:

Reagent	Volume for 1 reaction
Nuclease-free water	12.5µl
5x Herculase II Reaction Buffer (clear cap)	10µl
100mM dNTP Mix (green cap)	0.5µl
SureSelect Post-Capture Primer Mix (clear cap)	1µl
Herculase Fusion DNA Polymerase	1µl
Total	**25µl**

88) Mix well by pipetting up and down and spin down briefly. Keep on ice.

89) Add the full 25µl of the master mix to the tube containing the bead-bound target-enriched DNA. Mix well by pipetting up and down 20 times, until the bead suspension is homogeneous.

90) Put the tube in the thermocycler and run the PostCapAmp program as below:

Segment	Number of Cycles	Temperature	Time
1	1	98°C	2 Minutes
2	11	98°C	30 Seconds
		60°C	30 Seconds
		72°C	1 Minute
3	1	72°C	5 Minutes
4	1	4°C	∞ Hold

Note when this step is complete samples should be moved to your post-PCR room.

91) While the thermocycler is running, prepare a fresh dilution of 70% Ethanol in a 1.5ml Eppendorf tube by combining 700µl 100% ethanol with 300µl Nuclease free water

92) When the thermocycler run is complete, remove and spin the tube briefly to collect the liquid. 93) Put the tube on the magnetic separator at room temperature and wait for 2 mins, for the sample to become clear. Transfer the supernatant to a fresh 0.2ml Eppendorf tube and discard the beads. Stage 12: Post-Capture PCR clean up Est time ~ 30 mins

94) Vortex the aliquot of AMPureXP beads (see stage 10: 85) for 30 secs, until the mixture is homogeneous.

95) Add 50µl of homogeneous AMPure beads to the sample tube, pipetting up and down 15 times to mix.

96) Incubate the sample and bead mixture for 5 mins at room temperature.

97) Place the tube on the magnetic separation device and wait for the solution to clear, this can take 5 – 10 mins.

98) Keep the tube on the magnetic separator; with a pipette set for 200µl carefully remove and discard the cleared liquid from the tubes. Do not touch the beads while removing the solution.

99) Carry out the following steps twice:

• Keep the tube on the magnetic separator; Add 200µl of freshly-prepared 70% ethanol.

• Keep the tubes on the magnetic separator; Wait 1 min to allow any disturbed beads to settle, then remove and discard the ethanol.

100) Briefly spin down the tube to collect any residual liquid.

101) Return the tube to the magnetic separator and wait for 1 min. Remove any residual ethanol using a P10 pipette.

102) Run the 37C program on the thermocycler (37C for 3 minutes) and place the sample tube on it for no more than 60 secs, to completely dry the beads. Note: over-drying will lead to loss of yield

103) Add 25µl of nuclease-free water to the tube, close and vortex to mix well.

104) Briefly spin down to collect the liquid and leave the tube to incubate at room temperature for 2 mins.

105) Return the tube to the magnetic separator and leave the solution to clear for 3 mins. 106) Keep the tube on the magnetic separator; Remove the cleared supernatant (approx. 25µl) to a fresh 0.5ml Eppendorf tube and keep on ice. Stage 13: Amplified library normalisation Est time ~ 20 mins

107) Use 1µl library and 9µl nuclease-free water to make a dilution for QC. Determine the concentration of your libraries in nM and check the size distribution of each pool using your preferred method (e.g. High Sensitivity assay on Tapestation or Bioanalyzer).

108) Normalise each library pool to 5nM in 10µl in a fresh tube using the following calculations for the volumes of library and water needed:

$$V_s = \frac{10\ \mathrm{x}5}{C_s}$$

$$V_b = 10 - V_s$$

Where Cs is the library concentration in nM, Vs is the volume of amplified library needed and Vb is the volume of nuclease-free water needed. Dispense the nuclease-free water first,

pipetting directly into the bottom of a clean 0.5ml Eppendorf. Dispense the amplified library into the Nuclease-free water, pipetting up and down 5 times to flush the pipette tip and mix the dilution.

109) Vortex the normalised libraries briefly and spin down. Store at -20°C until ready for sequencing.

DAYS 3 & 4

110) Repeat days 1 and 2 for the second set of 16 samples. Do not store the pooled normalised libraries at -20°C at the end of the process, instead retrieve the pool of the first 16 libraries from the freezer.

Stage 14: Pooling and sequencing Est time ~ 20 mins

111) Prepare a final pool for sequencing containing all 32 samples and PhiX control according to the following table:

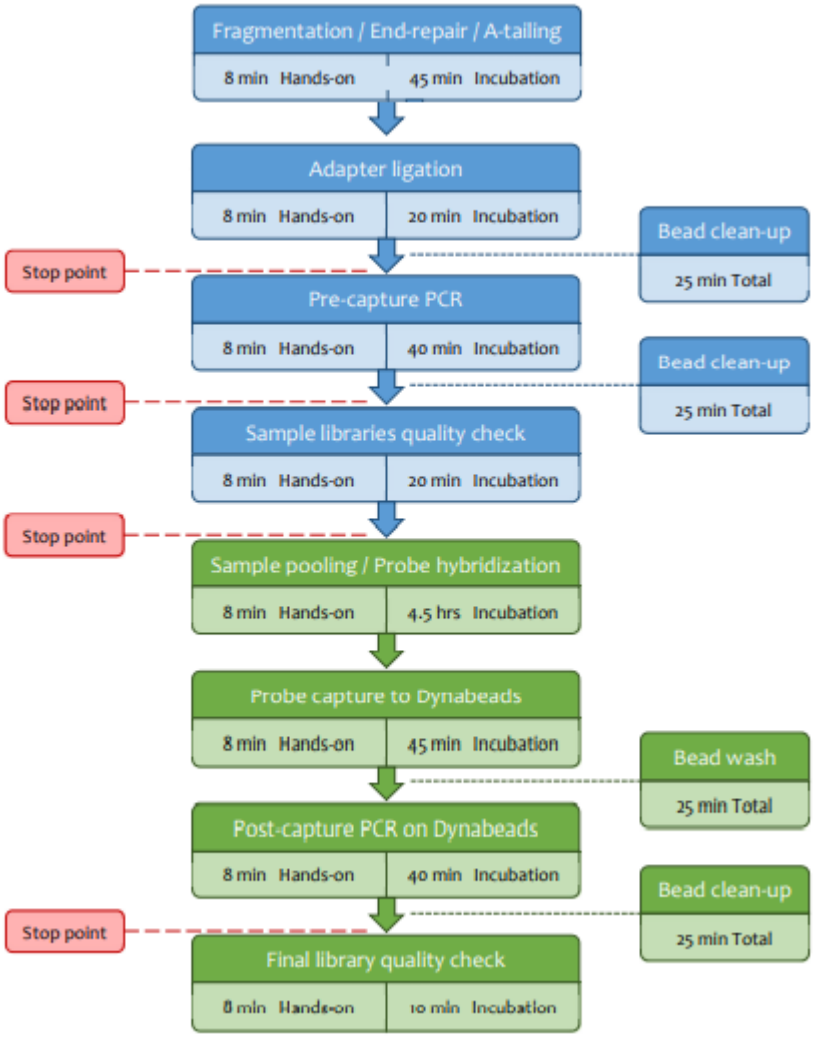
Reagent	Volume
Pool 1	7.5µl
Pool 2	7.5µl
1nM PhiX	1µl
Nuclease-free water	4µl
Total	20µl

112) Proceed to sequencing on a single high-output NextSeq flow cell following the appropriate Illumina protocol. The 8bp sample indexes must be sequenced but reading the unique molecular indexes is optional.

Appendix 10:

Nonacus Cell3™ Target: Pan Cancer panel library preparation supplier protocol

Overview of Nonacus Cell3™ Target: Pan Cancer panel library preparation protocol overview is shown below, with the full protocol available online: https://nonacus.com/wp-content/uploads/2020/05/Nonacus-Cell3-Target-protocol-v1-2-2-interactive.pdf



Appendix 11:

Tabulated variant numbers and TMB scores per patient dependent on panel and TMB calculations performed

Table below highlighting the number of variants per patient. Two datasets provided per panel based on +/-synonymous. 'Na' in table indicates where insufficient DNA meant that samples were not sequenced on a particular panel. The average variant number is calculated from the same 13 patients sequenced on these two NGS panels. Artefacts have not been excluded in this data. Red text indicates higher value in Nonacus panel compared to Illumina panel.

Patient ID	Illumina (-synonymous)	Illumina (+ synonymous)	Patient ID	Nonacus (-synonymous)	Nonacus (+ synonymous)
21M70071	40	79	21M70071	122	205
21M70072	134	158	21M70072	119	195
21M70073	176	196	21M70073	128	228
21M70074	148	164	21M70074	86	129
21M70075	173	213	21M70075	137	244
21M70076	192	245	21M70076	143	259
21M70077	na	na	21M70077	131	250
21M70078	82	144	21M70078	215	399
21M70079	64	116	21M70079	189	364
21M70081	49	72	21M70081	70	133
21M70084	231	243	21M70084	73	98
21M70086	269	385	21M70086	na	na
21M70087	na	na	21M70087	151	305
21M70088	93	105	21M70088	67	110
21M70089	145	177	21M70089	na	na
21M70090	72	87	21M70090	107	187
21M70091	80	98	21M70091	128	249
Average	**118**	**148**	**Average**	**122**	**215**

The following two tables show the TMB score (variants/Mb) per patient. Four datasets provided per panel based on +/-synonymous and +/-artefacts. 'Na' in table indicates where insufficient DNA meant that samples were not sequenced on a particular panel. The average TMB score is calculated from the same 13 patients sequenced on these two NGS panels. Red text indicates higher value in Nonacus panel compared to Illumina panel.

Illumina (-synonymous)

Patient ID	+artefacts	-artefacts
21M70071	50	49
21M70072	167	163
21M70073	219	214
21M70074	184	181
21M70075	215	211
21M70076	239	234
21M70077	na	na
21M70078	102	100
21M70079	80	78
21M70081	61	60
21M70084	288	282
21M70086	335	328
21M70087	na	na
21M70088	116	114
21M70089	181	177
21M70090	90	88
21M70091	100	98
Average	**147**	**144**

Illumina (+synonymous)

Patient ID	+artefacts	-artefacts
21M70071	98	95
21M70072	197	191
21M70073	244	237
21M70074	204	198
21M70075	265	257
21M70076	305	296
21M70077	na	na
21M70078	179	174
21M70079	144	140
21M70081	90	87
21M70084	303	294
21M70086	479	465
21M70087	na	na
21M70088	131	127
21M70089	220	214
21M70090	108	105
21M70091	122	118
Average	**184**	**178**

Nonacus (-synonymous)

Patient ID	+artefacts	-artefacts
21M70071	146	85
21M70072	142	83
21M70073	153	89
21M70074	103	60
21M70075	164	95
21M70076	171	99
21M70077	157	91
21M70078	257	149
21M70079	226	131
21M70081	84	49
21M70084	87	51
21M70086	na	na
21M70087	181	105
21M70088	80	47
21M70089	na	na
21M70090	128	74
21M70091	153	89
Average	**146**	**85**

Nonacus (+synonymous)

Patient ID	+artefacts	-artefacts
21M70071	245	184
21M70072	233	175
21M70073	273	205
21M70074	154	116
21M70075	292	219
21M70076	310	233
21M70077	299	224
21M70078	478	358
21M70079	436	327
21M70081	159	119
21M70084	117	88
21M70086	Na	Na
21M70087	365	274
21M70088	132	99
21M70089	Na	Na
21M70090	224	168
21M70091	298	224
Average	**258**	**193**

Appendix 12:

Process of selection of variants to interrogate in dbSNP and IGV

Table below describes the data interrogation identifying the number of unique and shared variants across the Illumina and Nonacus sequencing runs for both datasets (+/- synonymous). In each dataset, the variants most commonly shared between patients were interrogated in dbSNP and IGV (shaded cells in table).

	Illumina sequencing data		**Nonacus sequencing data**	
	Excluding synonymous variants	**Including synonymous variants**	**Excluding synonymous variants**	**Including synonymous variants**
Number of variants identified across all 15 patients	1948	2482	1866	3355
Number of unique variants identified across all 15 patients (as % of total variants)	1738 (89%)	2041 (82%)	613 (33%)	982 (29%)
Number of non-unique variants ie each shared by >1 patient (as % of total variants)	210 (11%)	441 (18%)	1253 (67%)	2373 (71%)
Number of variants shared by 2 patients	62	141	168	304
Number of variants shared by 3 patients	19	34	87	177
Number of variants shared by 4 patients	3	9	61	120
Number of variants shared by 5 patients	2	3	37	70
Number of variants	0	1	21	40

shared by 6 patients				
Number of variants shared by 7 patients	1	0	5	14
Number of variants shared by 8 patients	0	0	6	6
Number of variants shared by 9 patients	0	0	2	2
Number of variants shared by 10 patients	0	0	0	0

Appendix 13:

IGV and dbSNP evaluation of variants for identification of artefacts

IGV – Illumina -synonymous					
Reference sequence and variant (HGVS)	**No. of patients variant identified in**	**dbSNP reference SNP (rs) ID**	**Hg19 position**	**Reference/ alternate alleles and ALFA-generated frequencies in European population**	**Conclusion**
NM_170606.3: c.2512G>A	7	rs2479172	chr7: 151945007	C/T 0.73450/0.26550	SNP
NM_004327.4: c.3183-4G>A	5	rs180801	chr22: 23653880	G/A 0.61469/0.38531	SNP
NM_001321809.1: c.1036+9389_1036+9390del	5	rs528099287	chr14: 68944343	TTTTTTTTTTT/TTTTTTTTT 0.89415/0.00974	ARTEFACT
NM_001139.3: c.650+7C>G	4	rs2304908	chr17: 7983969	G/C 0.67102/0.32898	SNP
NM_001220777.1: c.-6+1395A>G	4	rs2395655	chr6: 36645696	A/G 0.584186/0.415814	SNP
NM_001305544.2: c.139A>G	4	rs3744093	chr17: 56492800	T/C 0.617880/0.382120	SNP

IGV – Illumina +synonymous					
Reference sequence and variant (HGVS)	**No. of patients variant identified in**	**dbSNP reference SNP (rs) ID**	**Hg19 position**	**Reference/ alternate alleles and ALFA-generated frequencies in European population**	**Conclusion**
NM_004327.4: c.3183-4G>A	6	rs180801	chr22: 23653880	G/A 0.62353/0.37647	SNP

NM_001362843.2: c.2164-15del	5	rs747710183	chr22: 41545024	TTTTTTTTTTTTTTT/TTTTTTTTTTTTT 0.84194/0.04254	ARTEFACT
NM_001321809.1: c.1036+9389_1036+9390del	5	rs528099287	chr14: 68944343	TTTTTTTTTTTTTT/TTTTTTTTTTTT 0.89415/0.00974	ARTEFACT
NM_001163034.2: c.2643C>T	5	rs1567962	chr17: 78919558	C/T 0.640046/0.359954	SNP
NM_001139.3: c.650+7C>G	4	rs2304908	chr17: 7983969	G/C 0.67102/0.32898	SNP
NM_001130823.3: c.1389A>G	4	rs2228611	chr19: 10267077	T/C 0.504528/0.495469	SNP
NM_001238.4: c.1215C>T	4	rs7257694	chr19: 30314666	C/T 0.60409/0.39508	SNP
NM_001304815.1: c.7260C>T	4	rs1052023	chr19: 42799049	C/T 0.623433/0.376567	SNP
NM_004958.4: c.5553C>T	4	rs2275527	chr1: 11190646	G/A 0.751726/0.24188	SNP
NM_001220777.1: c.-6+1395A>G	4	rs2395655	chr6: 36645696	A/G 0.584186/0.415814	SNP
NM_001305544.2:c.139A>G	4	rs3744093	chr17: 56492800	T/C 0.616860/0.383140	SNP
NM_001128226.3:c.887C>G	4	rs7332388	chr13: 73349359	G/C 0.58391/0.41609	SNP
NM_001305544.2:c.1252C>A	4	rs2526374	chr17: 56435885	G/T 0.638847/0.361153	SNP

IGV – Nonacus -synonymous					
Reference sequence and variant (HGVS)	**No. of patients variant identified in**	**dbSNP reference SNP (rs) ID**	**Hg19 position**	**Reference/alternate alleles and ALFA-generated frequencies in European population**	**Conclusion**
NM_001159995.3: c.338-4dup	9	rs750640301	chr8: 32472019	TTTTTTTTT/ TTTTTTTTTT 0.99295/0.00705	ARTEFACT

NM_001286559.2: c.724C>T	9	rs1052809	chr2: 95542476	C/T 0.99971/0.00029	VUS
NM_001303103.1: c.1419del	8	n/a	chr9: 84208131	n/a	ARTEFACT
NM_006267.5: c.1064-4dup	8	rs770828218	chr2: 109365363	TTTTTTTTT/TTTTTTTTTT 1.0000/0.0000	ARTEFACT
NM_001114121.2: c.676dup	8	rs757716680	chr11: 125505377	AAAAAAAAA/AAAAAAAAAA 0.99399/0.00601	ARTEFACT
NM_001243835.2: c.274-4dup	8	rs760220932	chr2: 191941054	AAAAAAAAAA/AAAAAAAAAAA 0.95286/0.04465	ARTEFACT
NM_001349370.2: c.5753C>T	8	rs4082155	chr3: 47125385	G/A 0.437944/0.562056	SNP
NM_004260.3: c.2296+1delC	8	rs1134207 7	chr8: 145738767	G/GG 0.99997/0.00003	VUS
NM_000965.4: c.787-9dup	7	rs750483330	chr3: 256349976	TTTTTTTTT/TTTTTTTTTT 0.99848/0.00114	ARTEFACT
NM_000553.6: c.3222G>T	7	rs1801195	chr8: 309992800	G/A 0.563891/0.436109	SNP
NM_001159995.3: c.1170-537dup	7	rs201198881	chr8: 326207220	TTTTTTTTT/TTTTTTTTTT 0.99383/0.00317	ARTEFACT
NM_199242.2: c.1992+5G>A	7	rs17581728	chr17: 73830996	C/T 0.763238/0.236762	SNP
NM_000057.4: c.1544dup	7	rs367543043	chr15: 91304138	AAAAAAAAA/AAAAAAAAAA 1.0000/0.0000	ARTEFACT

IGV – Nonacus +synonymous					
Reference sequence and variant (HGVS)	**No. of patients variant identified in**	**dbSNP reference SNP (rs) ID**	**Hg19 position**	**Reference/alternate alleles and ALFA-generated frequencies in European population**	**Conclusion**
NM_001286559.2: c.724C>T	9	rs1052809	chr2: 95542476	C/T 0.99971/0.00029	VUS
NM_001159995.3: c.338-4dup	9	rs750640301	chr8: 32472019	TTTTTTTTT/TTTTTTTTTT 0.99295/0.00705	ARTEFACT

NM_006267.5: c.1064-4dup	8	rs770828218	chr2: 109365363	TTTTTTTTT/TTTTTTTTTT 1.0000/0.0000	ARTEFACT
NM_001114121.2: c.676dup	8	rs757716680	chr11: 125505377	AAAAAAAAA/AAAAAAAAAA 0.99399/0.00601	ARTEFACT
NM_001243835.2: c.274-4dup	8	rs760220932	chr2: 191941054	AAAAAAAAAA/AAAAAAAAAAA 0.95286/0.04465	ARTEFACT
NM_004260.3: c.2296+1del	8	rs11342077	chr8: 145738767	G/GG 0.99997/0.00003	VUS
NM_001303103.1: c.1419del	8	n/a	chr9: 84208131	GGGGGG/GGGGG	ARTEFACT
NM_001349370.2: c.5753C>T	8	rs4082155	chr3: 471253845	G/A 0.437944/0.562056	SNP
NM_001202521.1: c.*25T>C	7	rs1049623	chr6: 30864829	T/C 0.614612/0.385388	SNP
NM_001159995.3: c.1170-537dup	7	rs201198881	chr8: 32620720	TTTTTTTTT/TTTTTTTTTT 0.99383/0.00317	ARTEFACT
NM_000965.4: c.787-9dup	7	rs750483330	chr3: 256349976	TTTTTTTTT/TTTTTTTTTT 0.99848/0.00114	ARTEFACT
NM_000553.6: c.3222G>T	7	rs1801195	chr8: 309992820	G/T 0.563891/0.436109	SNP
NM_199242.2: c.1992+5G>A	7	rs17581728	chr17: 73830996	C/T 0.763238/0.236762	SNP
NM_000057.4: c.1544dup	7	rs367543043	chr15: 91304138	AAAAAAAAA/AAAAAAAAAA 1.0000/0.0000	ARTEFACT
NM_001005360.2: c.2139T>C	7	rs2229920	chr19: 10939792	T/C 0.712232/0.287768	SNP
NM_001128844.2: c.1524T>C	7	rs7935	chr19: 11105608	T/C 0.648046/0.351954	SNP
NM_001349370.2: c.3333T>C	7	rs6767907	chr3: 471626661	A/G 0.408488/0.591512	SNP
NM_001318040.1: c.897T>C	7	rs7030167	chr9: 116791161	T/C 0.332259/0.667741	SNP
NM_017617.5: c.5094C>T	7	rs10521	chr9: 139397707	G/A 0.63613/0.36387	SNP

NM_000553.6: c.513C>T	7	rs1800389	chr8: 30924557	C/T 0.28882/0.71118	SNP
NM_005245.4: c.8904C>T	7	rs1280099	chr4: 187538330	G/A 0.545402/0.454598	SNP
NM_005245.4: c.9351T>C	7	rs2249917	chr4: 187534375	A/G 0.543105/0.456895	SNP

Appendix 14:

Determination of TMB high thresholds

The tables below show the range of TMB high thresholds evaluated in each panel/analysis dataset, and the associated true positive and true negative rates, and false positive and false negative rates across the patient cohort.

Illumina: +synonymous +artefacts

Illumina: high TMB score (variants/Mb)	True +ve (=responder)	False -ve	False +ve	True -ve (non-responder)	False +ve rate (spec)	True +ve rate (sens)
>10	6	0	9	0	1.00	1.00
>90	6	0	9	0	1.00	1.00
>100	5	1	8	1	0.89	0.83
>120	5	1	7	2	0.78	0.83
>130	**5**	**1**	**6**	**3**	**0.67**	**0.83**
>140	4	2	6	3	0.67	0.67
>170	4	2	5	4	0.56	0.67
>190	3	3	5	4	0.56	0.50
>200	3	3	4	5	0.44	0.50
>220	3	3	3	6	0.33	0.50
>240	3	3	2	7	0.22	0.50
>260	2	4	2	7	0.22	0.33
>300	1	5	2	7	0.22	0.17
>400	0	6	1	8	0.11	0.00

Nonacus: +synonymous +artefacts

Nonacus: high TMB score (variants/Mb)	True +ve (=responder)	False -ve	False +ve	True -ve (non-responder)	False +ve rate (spec)	True +ve rate (sens)
>10	6	0	9	0	1.00	1.00
>90	6	0	9	0	1.00	1.00
>100	6	0	9	0	1.00	1.00
>120	6	0	8	1	0.89	1.00
>130	6	0	7	2	0.78	1.00
>140	6	0	6	3	0.67	1.00
>170	6	0	4	5	0.44	1.00
>190	6	0	4	5	0.44	1.00
>200	6	0	4	5	0.44	1.00
>220	6	0	4	5	0.44	1.00
>240	**6**	**0**	**2**	**7**	**0.22**	**1.00**

>260	5	1	2	7	0.22	0.83
>290	4	2	2	7	0.22	0.67
>300	2	4	2	7	0.22	0.33
>400	0	6	2	7	0.22	0.00
>440	0	6	1	8	0.11	0.00

Illumina: -synonymous +artefacts

Illumina: high TMB score (variants/Mb)	True +ve (=responder)	False -ve	False +ve	True -ve (non-responder)	False +ve rate (spec)	True +ve rate (sens)
>10	6	0	9	0	1.00	1.00
>40	6	0	9	0	1.00	1.00
>70	5	1	8	1	0.89	0.83
>90	5	1	7	2	0.78	0.83
>100	**5**	**1**	**6**	**3**	**0.67**	**0.83**
>110	4	2	5	4	0.56	0.67
>130	3	3	5	4	0.56	0.50
>150	3	3	5	4	0.56	0.50
>160	3	3	5	4	0.56	0.50
>180	3	3	3	6	0.33	0.50
>190	3	3	2	7	0.22	0.50
>210	3	3	2	7	0.22	0.50
>220	1	5	2	7	0.22	0.17
>240	0	6	2	7	0.22	0.00
>290	0	6	1	8	0.11	0.00

Nonacus: -synonymous +artefacts

Nonacus: high TMB score (mutations/Mb)	True +ve (=responder)	False -ve	False +ve	True -ve (non-responder)	False +ve rate (spec)	True +ve rate (sens)
>10	6	0	9	0	1.00	1.00
>40	6	0	9	0	1.00	1.00
>70	6	0	9	0	1.00	1.00
>90	6	0	9	0	1.00	1.00
>100	6	0	6	3	0.67	1.00
>110	6	0	6	3	0.67	1.00
>130	**6**	**0**	**3**	**6**	**0.33**	**1.00**
>150	5	1	2	7	0.22	0.83
>160	3	3	2	7	0.22	0.50
>180	3	3	2	7	0.22	0.50
>190	0	6	2	7	0.22	0.00
>210	0	6	2	7	0.22	0.00

>220	0	6	2	7	0.22	0.00
>240	0	6	1	8	0.11	0.00
>290	0	6	0	9	0.00	0.00

Illumina: +synonymous -artefacts

Illumina: high TMB score (mutations/Mb)	True +ve (=responder)	False -ve	False +ve	True -ve (non-responder)	False +ve rate (spec)	True +ve rate (sens)
>10	6	0	9	0	1.00	1.00
>70	6	0	9	0	1.00	1.00
>90	6	0	8	1	0.89	1.00
>100	5	1	8	1	0.89	0.83
>110	5	1	7	2	0.78	0.83
>120	**5**	**1**	**6**	**3**	**0.67**	**0.83**
>140	4	2	6	3	0.67	0.67
>170	4	2	5	4	0.56	0.67
>190	3	3	5	4	0.56	0.50
>210	3	3	3	6	0.33	0.50
>230	3	3	2	7	0.22	0.50
>250	2	4	2	7	0.22	0.33
>290	1	5	1	8	0.11	0.17
>300	0	6	1	8	0.11	0.00

Nonacus: +synonymous -artefacts

Nonacus: high TMB score (mutations/Mb)	True +ve (=responder)	False -ve	False +ve	True -ve (non-responder)	False +ve rate (spec)	True +ve rate (sens)
>10	6	0	9	0	1.00	1.00
>70	6	0	9	0	1.00	1.00
>90	6	0	8	1	0.89	1.00
>110	6	0	7	2	0.78	1.00
>120	6	0	4	5	0.44	1.00
>140	6	0	4	5	0.44	1.00
>170	6	0	3	6	0.33	1.00
>180	**6**	**0**	**2**	**7**	**0.22**	**1.00**
>200	5	1	2	7	0.22	0.83
>210	4	2	2	7	0.22	0.67
>220	3	3	2	7	0.22	0.50
>230	2	4	2	7	0.22	0.33
>270	1	5	2	7	0.22	0.17
>300	0	6	2	7	0.22	0.00
>320	0	6	1	8	0.11	0.00

Illumina: -synonymous -artefacts

Illumina: high TMB score (mutations/Mb)	True +ve (=responder)	False -ve	False +ve	True -ve (non-responder)	False +ve rate (spec)	True +ve rate (sens)
>10	6	0	9	0	1.00	1.00
>40	6	0	9	0	1.00	1.00
>60	5	1	9	0	1.00	0.83
>70	5	1	8	1	0.89	0.83
>80	5	1	7	2	0.78	0.83
>90	5	1	6	3	0.67	0.83
>100	**5**	**1**	**5**	**4**	**0.56**	**0.83**
>110	4	2	5	4	0.56	0.67
>130	4	2	5	4	0.56	0.67
>160	3	3	5	4	0.56	0.50
>180	3	3	3	6	0.33	0.50
>190	3	3	3	6	0.33	0.50
>210	3	3	3	6	0.33	0.50
>220	1	5	2	7	0.22	0.17
>240	0	6	2	7	0.22	0.00
>290	0	6	1	8	0.11	0.00

Nonacus: -synonymous -artefacts

Nonacus: high TMB score (mutations/Mb)	True +ve (=responder)	False -ve	False +ve	True -ve (non-responder)	False +ve rate (spec)	True +ve rate (sens)
>10	6	0	9	0	1.00	1.00
>40	6	0	9	0	1.00	1.00
>50	6	0	7	2	0.78	1.00
>60	6	0	6	3	0.67	1.00
>70	6	0	5	4	0.56	1.00
>80	6	0	4	5	0.44	1.00
>90	4	2	3	6	0.33	0.71
>100	1	5	2	7	0.22	0.17
>110	0	6	2	7	0.22	0.00
>130	0	6	2	7	0.22	0.00
>140	0	6	1	8	0.11	0.00

References

Abate, R.E. (2020). *Update on the project Tumor Mutation Burden (TMB): from Recommendations for Testing to EQA Schemes.* [online] Available at: https://www.iqnpath.org/update-on-the-project [Accessed 13th November 2022].

Abate, R.E., Cheetham, M.H., Fairley, J.A., Pasquale, R., Sacco, A., Wolstenholme, N., Deans, Z.C., Patton, S.J., and Normanno, N. (2022) External quality assessment (EQA) for tumour mutations burden: results of an international IQN path feasibility pilot scheme. *Virchows Arch*. https://doi.org/10.1007/s00428-022-03444-y

Acgs.uk.com, (2022). *Best Practice Guidelines*. [online] Available at: https://www.acgs.uk.com/quality/best-practice-guidelines/ [Accessed 13th November 2022].

Albano, D., Bilfinger, T., and Nemesure, B. (2018) 1-, 3-, and 5-year survival among early-stage lung cancer patients treated with lobectomy vs SBRT. *Lung Cancer: Targets and Therapy* 9: 65-71.

Alexandrov, L.B., Nik-Zainal, S., Wedge, D.C., Aparicio, S.A.J.R., Behjati, S., Biankin, A.V., Bignell, G.R., Bolli, N., Borg, A., Børresen-Dale, A.-L., Boyault, S., Burkhardt, B., Butler, A.P., Caldas, C., Davies, H.R., Desmedt, C., Eils, R., Eyfjörd, J.E., Foekens, J.A., Greaves, M., Hosoda, F., Hutter, B., Ilicic, T., Imbeaud, S., Imielinsk, M., Jäger, N., Jones, D.T.W., Jones, D., Knappskog, S., Kool, M., Lakhani, S.R., López-Otín, C., Martin, S., Munshi, N.C., Nakamura, H., Northcott, P.A., Pajic, M., Papaemmanuil, E., Paradiso, A., Pearson, J.V., Puente, X.S., Raine, K., Ramakrishna, M., Richardson, A.L., Richter, J., Rosenstiel, P., Schlesner, M., Schumacher, T.N., Span, P.N., Teague, J.W., Totoki, Y., Tutt, A.N.J., Valdés-Mas, R., van Buuren, M.M., van 't Veer, L., Vincent-Salomon, A., Waddell, N., Yates, L.R. Australian Pancreatic Cancer Genome Initiative, ICGC Breast Cancer Consortium, ICGC MMML-Seq Consortium, ICGC PedBrain, Zucman-Rossi, J., Futreal, P.A., McDermott, U., Lichter, P., Meyerson, M., Grimmond, S.M., Siebert, R., Campo, E., Shibata, T., Pfister, S.M., Campbell, P.J., and Stratton, M.R. (2013) Signatures of mutational processes in human cancer. *Nature,* 500:415-421.

Alexandrov, L.B., Ju, Y.S., Haase, K., Loo, P.V., Martincorena, I., Nik-Zainal, S., Totoki, Y., Fujimoto, A., Nakagawa, H., Shibata, T., Campbell, P.J., Vineis, P., Phillips, D.H., and Stratton, M.R. (2016) Mutational signatures associated with tobacco smoking in human cancer. *Science* 354:618-622.

Bettegowda, C., Sausen, M., Leary, R.J., Kinde, I., Wang, Y., Agrawal, N., Bartlett, B.R., Wang, H., Luber, B., Alani, R.M., Antonarakis, E.S., Azad, N.S., Bardelli, A., Brem, H., Cameron, J.L., Lee, C.C., Fecher, L.A., Gallia, G.L., Gibbs, P., Le, D., Giuntoli, R.L., Goggins, M., Hogarty, M.D., Holdhoff, M., Hong, S.M., Jiao, Y., Juhl, H.H., Kim, J.J., Siravegna, G., Laheru, D.A., Lauricella, C., Lim, M., Lipson, E.J., Marie, S.K., Netto, G.J., Oliner, K.S., Olivi, A., Olsson, L., Riggins, G.J., Sartore-Bianchi, A., Schmidt, K., Shih, l.M., Oba-Shinjo, S.M., Siena, S., Theodorescu, D., Tie, J., Harkins, T.T., Veronese, S., Wang, T.L., Weingart, J.D., Wolfgang, C.L., Wood, L.D., Xing, D., Hruban, R.H., Wu, J., Allen, P.J., Schmidt, C.M., Choti, M.A., Velculescu, V.E., Kinzler, K.W., Vogelstein, B., Papadopoulos, and N., Diaz, L.A. Jr. (2014) Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Transl Med* 6:224.

Bewicke-Copley, F., Kumar, E.A., Palladino, G., Korfi, K., and Wang, J. (2019) Applications and analysis of targeted genomic sequencing in cancer studies. *Comp and Struct Biotech Journal* 17:1348-1359.

Bio-bwa-sourceforge.net, (2022). *Burrows-Wheeler Aligner*. [online] Available at: http://bio-bwa.sourceforge.net/ [Accessed 7th September 2021].

Bioinformatics.babraham.ac.uk, (2022). *FastQC*. [online] Available at: http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc [Accessed 25th November 2021].

Bristol-Myers Squibb, (2018). *U.S. Food and Drug Administration (FDA) accepts application for Opdivo plus low-dose Yervoy for treatment of first-line non-small cell lung cancer in patients with tumor mutational burden* >10 mut/Mb. [online] Available at: https://news.bms.com/news/corporate-financial/2018/US-Food-and-Drug-Administration-FDA-Accepts-Application-for-Opdivo-Plus-Low-Dose-Yervoy-for-Treatment-of-First-Line-Non-Small-Cell-Lung-Cancer-in-Patients-with-Tumor-Mutational-Burden-10-mutMb/default.aspx [Accessed on 18th September 2020].

Buchhalter, I., Rempel, E., Endris, V., Allgäuer, M., Neumann, O., Volckmar, A.-L., Kirchner, M., Leichsenring, J., Lier, A., von Winterfeld, M., Penzel, R., Christopoulos, P., Thomas, M., Fröhling, S., Schirmacher, P., Budczies, J., and Stenzinger, A. (2019) Size matters: Dissecting key parameters for panel-based tumor mutational burden analysis. *Int J Can* 144(4):848-858.

Budczies, J., Allgäuer, M., Litchfield, K., Rempel, E., Christopoulos, P., Kazdal, D., Endris, V., Thomas. M., Frohling, S., Peters, S., Swanton, C., Schirmacher, P, and Stenzinger, A. (2019) Optimizing panel-based tumor mutational burden (TMB) measurement. *Ann Oncol* 30(9):1496-1506.

Campesato, L.F., Barroso-Sousa, R., Jimenez, L., Correa, B.R., Sabbaga, J., Hoff, P.M., Reis, L.F.L., Galante, P.A.F., and Camargo, A.A. (2015) Comprehensive cancer-gene panels can be used to estimate mutational load and predict clinical benefit to PD-1 blockade in clinical practice. *Oncotarget* 6(33):34221–34227.

Carbone, D.,P., Reck, M., Paz-Ares, L., Creelan, B., Horn, L., Steins, M., Felip, E., van den Heuvel, M.M., Ciuleanu, T.-E., Badin, F., Ready, N., Hiltermann, T.J.N., Nair, S., Juergens, R., Peters, S., Minenza, E., Wrangle, J.M., Rodriguez-Abreu, D., Borghaei, H., Blumenschein, G.R., Villaruz, L.C., Havel, L.M., Krejci, J., Jaime, J.C., Chang, H., Geese, W.J., Bhagavatheeswaran, P., Chen, A.C., and Socinski, M.A. (2017) First-Line Nivolumab in Stage IV or Recurrent Non–Small-Cell Lung Cancer. *N Engl J Med* 376:2415-2426.

Castellanos, E., Snider, J., Mahamed Ali, S., Backenroth, D., Albacker, L.A., Murugesan, K., Li, G., Frampton, G.M., Alexander, B.M., and Carson, K.R. (2019) Tumor mutational burden (TMB) and PD-L1 expression as predictors of response to immunotherapy (IO) in NSCLC. *J Clin Oncol* 37(15):suppl 2630-2630.

Chalmers, Z.R., Connelly, C.F., Fabrizio, D., Gay, L., Ali, S.M., Ennis, R., Schrock, A., Campbell, B., Shlien, A., Chmielecki, J., Huang, F., He, Y., Sun, J., Tabori, U., Kennedy, M., Lieber, D.S., Roels, S., White, J., Otto, G.A., Ross, J.S., Garraway, L., Miller, V.A., Stephens, P.J., and Frampton, G.M. (2017) Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med* 9:34.

Chan, T.A., Yarchoan, M., Jaffee, E., Swanton, C., Quezada, S.A., Stenzinger, A., and Peters, S. (2019) Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. *Ann Oncol.* 30(1):44–56.

Chevallier, M., Borgeaud, M., Addeo, A., and Friedlaender, A. (2021) Oncogenic driver mutations in non-small cell lung cancer: Past, present and future. *World J Clin Oncol* 12(4):217-237.

Chiriva-Internati, M., and Bot, A. (2015) A New Era in Cancer Immunotherapy: Discovering Novel Targets and Reprogramming the Immune System. *International Reviews of Immunology* 34:101-103.

ClinicalTrials.gov (2022). *Search of: TMB | Recruiting, Not yet recruiting, Active, not recruiting, Enrolling by invitation Studies | Lung Cancer - Search Details.* [online] Available at: https://clinicaltrials.gov/ [Accessed on 9th June 2022].

Cottrell, T.R., and Taube, J.M. (2018) PD-L1 and Emerging Biomarkers in Immune Checkpoint Blockade Therapy. *Cancer J.* 24(1):41-46.

Cristescu, R., Mogg, R., Ayers, M., Albright, A., Murphy, E., Yearley, J., Sher, X., Liu, X., Nebozhyn, M., Lunceford, J.K., Joe, A.K., Cheng, J., Plimack, E.R., Ott, P.A., Seiwert, T.Y., McClanahan, T.K., Loboda, A., and Kaufman, D.R. (2017) Mutational load (ML) and T-cell-inflamed microenvironment as predictors of response to pembrolizumab. ASCO-SITC clinical immuno-oncology symposium. *J Clin Oncol* 35:S7S.

Davis, A.A., Kwang Chae, Y., Agte, S., Pan, A., Simon, N.I., Taxter, T.J., Behdad, A., Carneiro, B.A., Cristofanilli, M., and Giles, F.J. (2017) *J Clin Oncol* 35:15.

Do, H., and Dobrovic, A. (2015). Sequence Artifacts in DNA from Formalin-Fixed Tissues: Causes and Strategies for Minimization. *Clinical Chemistry* 61(1):64–71.

Eisenhauer, E.A., Therasse, P., Bogaerts, J.,, Schwartz, L.H., Sargent, D., Ford, R., Dancey, J., Arbuck, S., Gwyther, S., Mooney, M., Rubinstein, L., Shankar, L., Dodd, L., Kaplan, R., Lacombe, D., and Verweij, J. (2009) New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 45(2):228-247.

Emqn.org, (2022). *EMQN.* [online] Available at: https://www.emqn.org [Accessed 31st October 2022].

England.nhs.uk, (2022). *National Genomic Test Directory.* [online] Available at: https://www.england.nhs.uk/publication/national-genomic-test-directories/ [Accessed 10th September 2022].

Fancello, L., Gandini, S., Pelicci, P.G. and Mazzarella, L. (2019) Tumour mutational burden quantification from targeted gene panels: major advancements and challenges. *J for ImmunoTherapy of Cancer* 7:180.

Fda.gov, (2017). *Evaluation of Automatic Class III Designation for MSK-IMPACT (Integrated Mutation Profiling of Actionable Cancer Targets).* [online] Available at: https://www.accessdata.fda.gov/cdrh_docs/reviews/den170058.pdf [Accessed 18th November 2022].

Fda.gov, (2020a). *Hematology/Oncology (Cancer) Approvals & Safety Notifications.* [online] Available at: https://www.fda.gov/drugs/resources-information-approved-

drugs/hematologyoncology-cancer-approvals-safety-notifications [Accessed 18th September 2020].

Fda.gov, (2020b). *FoundationOne CDx FDA approval*. [online] Available at: https://www.accessdata.fda.goc/cdrh_docs/pdf17/P170019a.pdf. Accessed 18 Nov 2022.

Fenizia, F., Wolstenholme, N., Fairley, J.A., Rouleau, Et., Cheetham, M.H., Horan, M.P., Torlakovic, E., Besse, B., Al Dieri, R., Tiniakos, D.G., Deans, Z.C., Patton, S.J., and Normanno, N. (2021) Tumour mutation burden testing: a survey of the International Quality Network for Pathology (IQN Path). *Virchows Archiv* 479:1067-1072.

Fu, W., O'Connor, T.D., Jun, G., Kang, H.M., Abecasis, G., Leal, S.M., Gabriel, S., Rieder, M.J., Altshuler, D., SHendure, J., Nickerson, D.A., Bamshad, M.J., HNLBI Exome Sequencing Project, and Akey, J.M. (2012) Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. *Nature* 493:216-220.

Gandara, D.R., Paul, S.M., Kowanetz, M., Schleifman, E., Zou, W., Li, Y., Rittmeyer, A., Fehrenbacher, L., Otto, G., Malboeuf, C., Lieber, D.S., Lipson, D., Silterra, J., Amler, L., Riehl, T., Cummings, C.A., Hegde, P.S., Sandler, A., Ballinger, M., Fabrizio, D., Mok, T., and Shames, D.S. (2018) Blood-based tumor mutational burden as a predictor of clinical benefit in non-small-cell lung cancer patients treated with atezolizumab. *Nat Med* 24(9):1441-1448.

Garon, E.B., Rizvi, N.A., Hui, R., Leighl, N., Balmanoukian, A,S, Eder, J.P., Patnaik, A., Aggarwal, C., Gubens, M., Horn, L., Carcereny, E., Ahn, M., Felip, E., Lee, J.S., Hellmann, M.D., Hamid, O., Gldman, J.W., Soria, J.C., Dolled-Filhart, M.D., Rutledge, R.Z., Zhang, J., Lunceford, J.K., Rangwala. R., Lubiniecki, G.M., Roach, C., Emancipator, K., and Gandhi, L. (2015) Pembrolizumab for the Treatment of Non–Small-Cell Lung Cancer. *N Engl J Med* 372:2018-2028.

Gatk.broadinstitute.org, (2022a). *MarkDuplicates (Picard)*. [online] Available at: https://gatk.broadinstitute.org/hc/en-us/articles/360037052812-MarkDuplicates-Picard [Accessed 25th November 2021].

Gatk.broadinstitute.org, (2022b). *Mutect2.* [online] Available at: hhttps://gatk.broadinstitute.org/hc/en-us/articles/4409917447707-Mutect2 [Accessed 25th November 2021].

Github, (2022a). *Bioinfo-pf-curie/TMB.* [online] Available at: https://github.com/bioinfo-pf-curie/TMB [Accessed 16th December 2021].

Github (2022b). *Somatic Enrichment pipeline*. [online] Available at: https://github.com/AWGL/SomaticEnrichment [Accessed 25th November 2021].

Gnomad.broadinstitute.org, (2022) *GnomAD.* [online] Available at: https://gnomad.broadinstitute.org/ [Accessed 25th November 2021].

Hatakeyama, K., Nagashima, T., Urakami, K., Ohshima, K., Serizawa, M., Ohnami, S., Shimoda, Y., Ohnami, S., Maruyama, K., Naruoka, A., Akiyama, Y., Kusuhara, M., Mochizuki, T., and Yamaguchi, K. (2018) Tumor mutational burden analysis of 2,000 Japanese cancer genomes using whole exome and targeted gene panel sequencing. *Biomed Res.* 39(3):159-167.

Heeke, S., and Hofman, P. (2018) Tumour mutational burden assessment as a predictive biomarker for immunotherapy in lung cancer patients: getting ready for prime-time or not? *Transl Lung Cancer Res* 7:631-638.

Hellmann, M.D., Ciuleanu, T.E., Pluzanski, A., Lee, J.S., Otterson, G.A., Audigier-Valette, C., Minenza, E., Linardou, H., Burgers, S., Salman, P., Borghaei, H., Ramalingam, S.S., Brahmer, J., Reck, M., O'Byrne, K.J., Geese, W.G., Green, G., Chang, H., Szustakowski, J., Bhagavatheeswaran, P., Healey, D., Fu, Y., Nathan, F., and Paz-Ares, L. (2018a) Nivolumab plus Ipilimumab in lung cancer with a high tumor mutational burden. *N Engl J Med*. 378(22):2093-2104.

Hellmann, M.D., Callahan, M.K., Awad, M.M., Calvo, E., Ascierto, P.A., Atmaca, A., Rizvi, N.A., Hirsch, F.R., Selvaggi, G., Szustakowski, J.D., Sasson, A., Golhar, R., Vitazka, P., Chang, H., Geese, W.J., Antonia, S.J. (2018b) Tumor Mutational Burden and Efficacy of Nivolumab Monotherapy and in Combination with Ipilimumab in Small-Cell Lung Cancer. *Cancer Cell* 33(5):853-861.

Hirsch, F.R., McElhinny, A., Stanforth, D., Ranger-Moore,J.,Jansson, M., Kulangara, K., Richardson, W., Towne, P., Hanks, D., Vennapusa, B., Mistry, A., Kalamegham, R., Averbuch, S., Novotny, J., Rubin, E., Emancipator, K., McCaffery, I., Williams, J.A., Walker, J., Longshore, J., Tsao, M., and Kerr, K.M. (2017) PD-L1 Immunohistochemistry Assays for Lung Cancer: Results from Phase 1 of the Blueprint PD-L1 IHC Assay Comparison Project. *J Thorac Oncol.* 12:208–22.

Hugo, W., Zaretsky, J.M., Sun, L., Song, C., Moreno, B.H., Hu-Lieskovan, S., Berent-Maoz, B., Pang, J., Chmielowski, B., Cherry, G., Seja, E., Lomeli, S., Kong, X., Kelley, M.C., Sosman, J.A., Johnson, D.B., Ribas, A., Lo, R.S. (2016) Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. *Cell* 165(1):35-44.

Igv.org, (2022). *IGV: Integrative Genomics Viewer*. [online] Available at: https://igv.org [Accessed 10th April 2022].

Illumina.com, (2011). *Quality Scores for Next Generation Sequencing*. [online] Available at: https://www.illumina.com/Documents/products/technotes/technote_Q-Scores.pdf [Accessed 21st March 2022].

Incorvaia, L., Fanale, D., Badalamenti, G., Barraco, N., Bono, M., Corsini, L.R., Galvano, A., Gristina, V., Listì, A., Vieni, S., Gori, S., Bazan, V., and Russo, A. (2019) Programmed Death Ligand 1 (PD-L1) as a Predictive Biomarker for Pembrolizumab Therapy in Patients with Advanced Non-Small-Cell Lung Cancer (NSCLC). *Adv Ther.* 36(10):2600-2617.

Johnson, D.B., Frampton, G.M., Rioth, M.J., Yusko, E., Xu, Y., Guo, X., Ennis, R.C., Fabrizio, D., Chalmers, Z.R., Greenbowe, J., Ali, S.M., Balasubramanian, S., Sun, J.X., He, Y., Frederick, D.T., Puzanov, I., Balko, J.M., Cates, J.M., Ross, J.S., Sanders, C., Robins, H., Shyr, Y., Miller, V.A., Stephens, P.J., Sullivan, R.J., Sosman, J.A., and Lovly, C.M. (2016) Targeted Next Generation Sequencing Identifies Markers of Response to PD-1 Blockade. *Cancer Immunol Res.* 4(11):959-967.

Koboldt, D.C. (2020) Best practices for variant calling in clinical sequencing. *Genome Medicine* 12(91).

Kowanetz, M., Zou, W., Shames, D., Cummings, C., Rizvi, N., Spira, A., Frampton, G., Leveque, V., Flynn, S., Mocci, S., Shankar, G., Funke, R., Ballinger, M., Watrekamp, D., Chen, D., Sandler, A., Hampton, G., Amler, L., Hedge, P., and Hellmann, M. (2017) Tumor Mutation Burden (TMB) is associated with improved efficacy of atezolizumab in 1L and 2L+ NSCLC patients. *J Thorac Oncol* 12: S321–S322.

Lawrence, M.S., Stojanov, P., Polak, P., Kryukov, G.V., Cibulskis, K., Sivachenko, A., Carter, S.L., Stewart, C., Mermel, C.H., Roberts, S.A., Kiezun, A., Hammerman, P.S., McKenna, A., Drier, Y., Zou, L., Ramos, A.H., Pugh, T.J., Stransky, N., Helman, E., Kim, J., Sougnez, C., Ambrogio, L., Nickerson, E., Shefler, E., Cortés, M.L., Auclair, D., Saksena, G., Voet, D., Noble, M., DiCara, D., Lin, P., Lichtenstein, L., Heiman, D.I., Fennell, T., Imielinski, M., Hernandez, B., Hodis, E., Baca, S., Dulak, A.M., Lohr, J., Landau, D.A., Wu, C.J., Melendez-Zajgla, J., Hidalgo-Miranda, A., Koren, A., McCarroll, S.A., Mora, J., Crompton, B., Onofrio, R., Parkin, M., Winckler, W., Ardlie, K., Gabriel, S.B., Roberts, C.W.M., Biegel, J.A., Stegmaier, K., Bass, A.J., Garraway, L.A., Meyerson, M., Golub, T.R., Gordenin, D.A., Sunyaev, S., Lander, E.S., and Getz, G. (2013) Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* 499(7457):214-218.

Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., Tukiainen, T., Birnbaum, D.P., Kosmicki, J.A., Duncan, L.E., Estrada, K., Zhao, F., Zou, J., Pierce-Hoffman, E., Berghout, J., Cooper, D.N., Deflaux, N., DePristo, M., Do, R., Flannick, J., Fromer, M., Gauthier, L., Goldstein, J., Gupta, N., Howrigan, D., Kiezun, A., Kurki, M.I., Moonshine, A.L., Natarajan, P., Orozco, L., Peloso, G.M., Poplin, R., Rivas, M.A., Ruano-Rubio, V., Rose, S.A., Ruderfer, D.M., Shakir, K., Stenson, P.D., Stevens, C., Thomas, B.P., Tiao, G., Tusie-Luna, M.T., Weisburd, B., Won, H., Yu, D., Altshuler, D.M., Ardissino, D., Boehnke, M., Danesh, J., Donnelly, S., Elosua, R., Florez, J.C., Gabriel, S.B., Getz, G., Glatt, S.J., Hultman, C.M., Kathiresan, S., Laakso, M., McCarroll, S., McCarthy, M.I., McGovern, D., McPherson, R., Neale, B.M., Palotie, A., Purcell, S.M., Saleheen, D., Scharf, J.M., Sklar, P., Sullivan, P.F., Tuomilehto, J., Tsuang, M.T., Watkins, H.C., Wilson, J.G., Daly, M.J., MacArthur, D.G., and Exome Aggregation Consortium (2016) Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536:285–291.

Lim, C., Tsao, M.S., Le, L.W., Shepherd, F.A., Feld, R., Burkes, R.L., Liu, G., Kamel-Reid, S., Hwang, D., Tanguay, J., da Cunha Santos, G., and Leighl, N,B. (2015) Biomarker testing and time to treatment decision in patients with advanced non-small-cell lung cancer. *Ann Oncol* 26(7):1415-1421.

Marabelle, A., Fakih, M., Lopez, J., Shah, M., Shapira-Frommer, R., Nakagawa, K., Chung, H.C., Kindler, H.L., Lopez-Martin, J.A., Miller Jr, W.H., Italiano, A., Kao, S., Piha-Paul, S.A., Delord, J.P., McWilliams, R.R., Fabrizio, D.A., Aurora-Garg, D., Xu, L., Jin, F., Norwood, K., and Bang, Y.J. (2020) Association of tumour mutational burden with outcomes in patients with advanced solid tumours treated with pembrolizumab: prospective biomarker analysis of the multicohort, open-label, phase 2 KEYNOTE-158 study. *Lancet Oncol* 21:1353–1365.

Marcus, L., Fashoyin-Aje, L.A, Donoghue, M., Yuan, M., Rodriguez, L., Gallagher, P.S., Philip, R., Ghosh, S., Theoret, M.R., Beaver, J.A., Pazdur, R., and Lemery, S.J. (2021) FDA Approval Summary: Pembrolizumab for the Treatment of Tumor Mutational Burden–High Solid Tumors. *Clin Cancer Res* 27(17):4685–4689.

Mattocks, C.J., Morris, M.A., Matthijs, G., Swinnen, E., Corveleyn, A., Dequeker, E., Müller, C.R., Pratt, V., and Wallace, A.(2010) A standardized framework for the validation and verification of clinical molecular genetic tests. *Eur J Hum Genet* 18(12):1276-1288.

Melendez, B., Campenhout, C.V., Rorive, S., Remmelink, M., Salmon, I., and D-Haene, N. (2018) Methods of measurement for tumour mutations burden in tumour tissue. *Transl Lung Cancer Res* 7(6):661-667.

Mycancergenome.org (2022a). *Lung carcinoma*. [online] Available at: https://www.mycancergenome.org/content/disease/lung-carcinoma/ [Accessed on 13th November 2022].

Mycancergenome.org (2022b). *KMT2D*. [online] Available at: https://www.mycancergenome.org/content/gene/kmt2d/ [Accessed on 13th November 2022].

Ncbi.nlm.nih.gov, (2019). *National Library of Medicine: PubMed.gov*. [online] Available at: https://pubmed.ncbi.nlm.nih.gov/ [Accessed on 18th September 2020].

NHS, (2022). *Overview: Lung cancer.* [online] Available at: https://www.nhs.uk/conditions/lung-cancer/ [Accessed 13 October 2022].

NICE, (2016a). *Pembrolizumab for treating PD-L1-positive non-small-cell lung cancer after chemotherapy.* [online] National Institute for Health and Care Excellence (NICE). Available at: https://www.nice.org.uk/guidance/ta428/documents/final-appraisal-determination-document [Accessed 18th September 2020].

NICE, (2016b). *Nivolumab for treating advanced (unresectable or metastatic) melanoma.* [online] National Institute for Health and Care Excellence (NICE). Available at: https://www.nice.org.uk/guidance/ta384/documents/final-appraisal-determination-document [Accessed 18th September 2020].

NICE, (2021). *Lung and pleural cancers – recognition and referral: How common is it?* [online] National Institute for Health and Care Excellence (NICE). Available at: https://cks.nice.org.uk/topics/lung-pleural-cancers-recognition-referral/background-information/prevalence [Accessed 13th October 2022].

NICE, (2022). *Lung cancer: diagnosis and management.* [online] Available at: https://www.nice.org.uk/guidance/NG122 [Accessed 13th October 2022].

Pai-Scherf, L., Blumenthal, G.M., Li, H., Subramaniam, S., Mishra-Kalyani, P.S., Hong Zhao, K.H., Yu, J., Paciga, M., Goldberg, K.B., McKee, A.E., Keegan, P., and Pazdur, R. (2017) FDA Approval Summary: Pembrolizumab for Treatment of Metastatic Non-Small Cell Lung Cancer: First-Line Therapy and Beyond. *Oncologist,* 22(11):1392-1399.

Pcingola.github, (2021) *SnpEff and EnpSift*. [online] Available at: http://pcingola.github.io/SnpEff/ [Accessed 25th November 2021].

Pestinger, V., Smith, M., Sillo, T., Findlay, T.M., Laes, J.-F., Martin, G., Middleton, G., Taniere, P., and Beggs, A.D. (2020) Use of an Integrated Pan-Cancer Oncology Enrichment Next-Generation Sequencing Assay to Measure Tumour Mutational Burden and Detect Clinically Actionable Variants. *Mol Diagn Ther* 24(3):339-349.

Petrackova, A., Vasinek, M., Sedlarikova, L., Dyskova, T., Schneiderova, P., Novosad, T., Papajik, T., and Kriegova, E. (2019) Standardization of Sequencing Coverage Depth in NGS: Recommendation for Detection of Clonal and Subclonal Mutations in Cancer Diagnostics. *Front Oncol* 9:851.

Peters, S., Gettinger, S., Johnson, M.L., Jänne, P.A., Pasi A. Jänne

Garassino, M.C.,Marina C. Garassino

Christoph, D., Toh, C.K., Rizvi, N.A., Naiyer A. Rizvi

Chaft, J.E., Jamie E. Chaft

Costa, E.C., Enric Carcereny Costa

Patel, J.D., Chow, L.Q.M., Koczywas, M., Marianna Koczywas

Ho, C., Früh, M., van den Heuvel, M., Rothenstein, J., Reck, M., Paz-Ares, L., Shepherd, F.A., Kurata, T., Li, Z., Zhengrong Li

Qiu, J., Jiaheng Qiu

Kowanetz, M., Marcin Kowanetz

Mocci, S., Shankar, G., Sandler, A., and Felip, E. (2017) Phase II trial of atezolizumab as first-line or subsequent therapy for patients with programmed death-ligand 1-selected advanced non-small-cell lung cancer (BIRCH). *J Clin Oncol* 35(24):2781.

Phan, L., Jin, Y., Zhang, H., Qiang, W., Shekhtman, E., Shao, D., Revoe, D., Villamarin, R., Ivanchenko, E., Kimura, M., Wang, Z.Y., Hao, L., Sharopova, N., Bihan, M., Sturcke, A., Lee, M., Popova, N., Wu, W., Bastiani, C., Ward, M., Holmes, J.B., Lyoshin, V., Kaur, L., Moyer, E., Feolo, M., and Kattman, B.L. (2020) ALFA: Allele Frequency Aggregator. National Center for Biotechnology Information, U.S. National Library of Medicine. [online] Available at: www.ncbi.nlm.nih.gov/snp/docs/gsr/alfa/ [Accessed 1st April 2022].

Phgfoundation.org, (2011). *Next steps in the sequence: The implications of whole genome sequencing for health in the UK.* [online] Available at: https://www.phgfoundation.org/documents/report_next_steps_Sequence.pdf [Accessed 18th September 2020].

Potapov, V., and Ong, J.L. (2017) Examining Sources of Error in PCR by Single-Molecule Sequencing. Plos One 12(1). https://doi.org/10.1371/journal.pone.0169774

Qiagen.com (2022) *Next Generation Sequencing*. [online] Available at: www.qiagen.com/gb [Accessed 14th July 2022].

Ramalingam, S.S., Hellmann, M.D., Awad, M.M., Borghaei, H., Gainor, J., Brahmer, J., Spigel, D.R., Reck, M., O'Byrne, K.J., Paz-Ares, L., Zerba, K., Li, X., Geese, W.J., Green, G., Lestini, B., Szustakowski, J.D., Chang, H., and Ready, N. (2018) Tumor mutation burden (TMB) as a biomarker for clinical benefit from dual immune checkpoint blockade with nivolumab (nivo) + ipilimumab (ipi) in first-line (1L) non-small cell lung cancer (NSCLC): identification of TMB cutoff from CheckMate 568. *Am Ass Can Res* 78(13).

Ramos-Paradas, J., Hernández-Prieto, S., Lora, D., Sanchez, E., Rosado, A., Caniego-Casas, T., Carrizo, N., Enguita, A.B., Muñoz-Jimenez, M.T., Rodriguez, B., Perez-

Gonzalez, U., Gómez-Sánchez, D., Ferrer, I., Ponce Aix, S., Nuñez Buiza, Á., Garrido, P., Palacios, J., Lopez-Rios, F., Garrido-Martin, E.M., and Paz-Ares, L. (2021) Tumor mutational burden assessment in non-small-cell lung cancer samples: results from the TMB harmonization project comparing three NGS panels. *J Immunother Cancer* 9(5):e001904. doi: 10.1136/jitc-2020-001904.

Ribas A., Hamid O., Daud A., Hodi F.S., Wolchok J.D., Kefford R., Joshua A.M., Patnaik A., Hwu W.J., Weber J.S., Gangadhar, T.C., Hersey, P., Dronca, R., Joseph, R.W., Zarour, H., Chmielowski, B., Lawrence, D.P., Algazi, A., Rizvi, N.A., Hoffner, B., Mateus, C., Gergich, K., Lindia, J.A., Giannotti, M., Li, X.N., Ebbinghaus, S., Kang, S.P., and Robert, C. (2016) Association of Pembrolizumab with tumor response and survival among patients with advanced melanoma. *JAMA* 315:1600–1609.

Rimm, D.L., Han, G., Taube, J.M., Yi, E.S., Bridge, J.A., Flieder, D.B., Homer, R., West, W.W., Wu, H., Roden, A.C., Fujimoto, J., Yu, H., Anders, R., Kowalewski, A., Rivard, C., Rehman, J., Batenchuk, C., Burns, V., Hirsch, F.R., and Wistuba, I.I. (2017) A Prospective, Multi-institutional, Pathologist-Based Assessment of 4 Immunohistochemistry Assays for PD-L1 Expression in Non-Small Cell Lung Cancer. *JAMA Oncol* 3:1051–1058.

Rizvi, N.A., Hellmann, M.D., Snyder, A., Kvistborg, P., Makarov, V., Havel, J.J., Lee, W., Yuan, J., Wong, P., Ho, T.S., Miller, M.L., Rekhtman, N., Moreira, A.L., Ibrahim, F., Bruggeman, C., Gasmi, B., Zappasodi, R., Maeda, Y., Sander, C., Garon, E.B., Merghoub, T., Wolchok, J.D., Schumacher, T.N., Chan, T.A. (2015) Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* 348(6230):124–128.

Rizvi, H., Sanchez-Vega, F., La, K., Chatila, W., Jonsson, P., Halpenny, D., Plodkowski, A., Long, N., Sauter, J.L., Rekhtman, N., Hollmann, T., Schalper, K.A., Gainor, J.F., Shen, R., Ni, A., Arbour, K.C., Merghoub, T., Wolchok, J., Snyder, A., Chaft, J.E., Kris, M.G., Rudin, C.M., Socci, N.D., Berger, M.F., Taylor, B.S., Zehir, A., Solit, D.B., Arcila, M.E., Ladanyi, M., Riely, G.J., Schultz, N., and Hellmann, M.D. (2018) Molecular Determinants of Response to Anti–Programmed Cell Death (PD)-1 and Anti–Programmed Death-Ligand 1 (PD-L1) Blockade in Patients With Non–Small-Cell Lung Cancer Profiled With Targeted Next-Generation Sequencing. *J Clin Oncol* 36(7):633–641.

Robert, C., Long, G.V., Brady, B., Dutriaux, C., Maio, M., Mortier, L., Hassel, J.C., Rutkowski, P., McNeil, C., Kalinka-Warzocha, E., Savage, K.J., Hernberg, M.M., Lebbé, C., Charles, J., Mihalcioiu, C., Chiarion-Sileni, V., Mauch, C., Cognetti, F., Arance, A., Schmidt, H., Schadendorf, D., Gogas, H., Lundgren-Eriksson, L., Horak, C., Sharkey, B., Waxman, I.M., Atkinson, V., and Ascierto, P.A. (2015) Nivolumab in previously untreated melanoma without BRAF mutation. *N Engl J Med* 372:320–330. https://www.nejm.org/doi/10.1056/NEJMoa1412082?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub=www.ncbi.nlm.nih.gov

Roberts, H. (2020). *Tumour Mutational Burden: A novel biomarker for immunotherapy stratification in lung cancer patients.* BSc. Manchester Metropolitan University.

Samorodnitsky, E., Jewell, B.M., Hagopian, R., Miya, J., Wing, M.R., Lyon, E., Damodaran, S., Bhatt, D., Reeser, J.W., Datta, J., and Roychowdhury, S. (2015).

Evaluation of Hybridization Capture Versus Amplicon-Based Methods for Whole-Exome Sequencing. *Hum Mutat* 36(9):903-14.

Seiwert, T.Y., Haddad, R., Bauml, J. Weiss, J., Pfister, D.G., Gupta, S., Mehra, R., Gluck, I., Kang, H., Worden, F., Eder, J.P., Tahara, M., Burtness, B., Liu, S.V., Webber, A., Huang, L., Mogg, R., Cristescu, R., Cheng, J., and Chow, L.Q. (2018) Biomarkers predictive of response to pembrolizumab in head and neck cancer (HNSCC). *Am Ass Can Res* 78 (13).

Sha, D., Jin, Z., Budczies, J., Kluck, K., Stenzinger, A., and Sinicrope, F.A. (2020) Tumor Mutational Burden as a Predictive Biomarker in Solid Tumors. *Cancer Discov* 10(12):1808-1825.

Sherry, S.T., Ward, M., and Sirotkin, K. (1999) dbSNP—Database for Single Nucleotide Polymorphisms and Other Classes of Minor Genetic Variation. *Genome Res* 9:677–679.

Spigel, D.R., Schrock, A.B., Fabrizio, D., Frampton, G.M., Sun, J., He, J., Gowen, K., Johnson, M.L., Bauer, T.M., Kalemkerian, G.P., Raez, L.E., Ou, S.H.I., Ross, J.S., Stephens, P.J., Miller, V.A., and Ali, S.M. (2016) Total mutation burden (TMB) in lung cancer (LC) and relationship with response to PD-1/PD-L1 targeted therapies. *J Clin Oncol 34(suppl).*

Stenzinger, A., Endris, V., Budczies, J., Merkelbach-Bruse, S., Kazdal, D., Dietmaier, W., Pfarr, N., Siebolts, U., Hummel, M., Herold, S., Andreas, J., Zoche, M., Tögel, L., Rempel, E., Maas, J., Merino, D., Stewart, M., Zaoui, K., Schlesner, M., Glimm, H., Fröhling, S., Allen, J., Horst, D., Baretton, G., Wickenhauser, C., Tiemann, M., Evert, M., Moch, H., Kirchner, T., Büttner, R., Schirmacher, P., Jung, A., Haller, F., Weichert, W., and Dietel, M. (2020) Harmonization and Standardization of Panel-Based Tumor Mutational Burden Measurement: Real-World Results and Recommendations of the Quality in Pathology Study. *J of Thoracic Onc* 15(7): 1177-1189.

Strickler, J.H., Hanks, B.A., and Khasraw, M. (2020) Tumor Mutational Burden as a Predictor of Immunotherapy Response: Is More Always Better? *Clin Cancer Res* (2021) 27(5):1236–1241.

Sul, J., Blumenthal, G.M., Jiang, X., He, K., Keegan, P., and Pazdur, R. (2016) FDA approval summary: pembrolizumab for the treatment of patients with metastatic non-small cell lung cancer whose tumors express programmed death-ligand 1. *Oncologist* 21:643–650.

Sun, J.X., Frampton, G., Wang, K., Ross, J.S., Miller, V.A., Stephens, P.J., Lipson, D., and Yelensky, R. (2018) Abstract 1893: A computational method for somatic versus germline variant status determination from targeted next-generation sequencing of clinical cancer specimens without a matched normal control. *Cancer Research* 74:19(suppl).

Targeted Oncology, (2019). *Bristol-Myers Squibb reports fourth quarter and full year financial results [news release].* [online] Available at: https://www.targetedonc.com/view/bms-withdraws-nivolumabipilimumab-application-in-tmbhigh-nsclc [Accessed 18th September 2020].

Taube, J.M., Klein, A., Brahmer, J.R., Xu, H., Pan, X., Kim, J.H., Chen, L., Pardoll, D.M., Topalian, S.L., and Anders, R.A. (2014) Association of PD-1, PD-1 Ligands, and

Other Features of the Tumor Immune Microenvironment with Response to Anti–PD-1 Therapy. *Clinical Cancer Research*. 20(19):5064–5074.

The 1000 Genomes Project Consortium (2015) A global reference for human genetic variation. *Nature* 526:68-74.

Topalian, S.L., Hodi, F.S., Brahmer, J.R., Gettinger, S.N., Smith, D.C., McDermott, D.F., Powderly, J.D., Carvajal, R.D., Sosman, J.A., Atkins, M.B., Leming, P.D., Spigel, D.R., Antonia, S.J., Horn, L., Drake, C.G., Pardoll, D.M., Chen, L., Sharfman, W.H., Anders, R.A., Taube, J.M., McMiller, T.L., Xu, H., Korman, A.J., Jure-Kunkel, M., Agrawal, S., McDonald, D., Kollia, G.D., Gupta, A., Wigginton, J.M., and Sznol, M. (2012) Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N Engl J Med* 366:2443-2454.

Ukneqas.org.uk, (2022). *UK NEQAS.* [online] Available at: https://ukneqas.org.uk, 2022 [Accessed 31st October 2022].

Underhill, H.R. (2021) Leveraging the Fragment Length of Circulating Tumour DNA to Improve Molecular Profiling of Solid Tumour Malignancies with Next-Generation Sequencing: A Pathway to Advanced Non-invasive Diagnostics in Precision Oncology? *Mol Diagn Ther* 25(4):389-408.

Vega, D.M., Yee, L.M., McShane, L.M., Williams, P.M., Chen, L., Vilimas, T., Fabrizio, D., Funari, V., Newberg, J., Bruce, L.K., Chen, S.J., Baden, J., Carl Barrett, J., Beer, P., Butler, M., Cheng, J.H., Conroy, J., Cyanam, D., Eyring, K., Garcia, E., Green, G., Gregersen, V.R., Hellmann, M.D., Keefer, L.A., Lasiter, L., Lazar, A.J., Li, M.C., MacConaill, L.E., Meier, K., Mellert, H., Pabla, S., Pallavajjalla, A., Pestano, G., Salgado, R., Samara, R., Sokol, E.S., Stafford, P., Budczies, J., Stenzinger, A., Tom, W., Valkenburg, K.C., Wang, X.Z., Weigman, V., Xie, M., Xie, Q., Zehir, A., Zhao, C., Zhao, Y., Stewart, M.D., and Allen, J. (2021) TMB Consortium. Aligning tumor mutational burden (TMB) quantification across diagnostic platforms: phase II of the Friends of Cancer Research TMB Harmonization Project. *Ann Oncol* 32(12):1626-1636.

Zaretsky, J.M., Garcia-Diaz, A., Shin, D.S., Escuin-Ordinas, H., Hugo, W., Hu-Lieskovan, S., Torrejon, D.Y., Abril-Rodriguez, G., Sandoval, S., Barthly, L., Saco, J., Homet Moreno, B., Mezzadra, R., Chmielowski, B., Ruchalski, K., Shintaku, A.P., Sanchez, P.J., Puig-Saus, C., Cherry, G., Seja, E., Kong, X., Pang, J., Berent-Maoz, B., Comin-Anduix, B., Graeber, T.G., Tumeh, P.C., Schumacher, T.N.M., Lo, R.S., and Ribas, A. (2016) Mutations Associated with Acquired Resistance to PD-1 Blockade in Melanoma. *N Engl J Med.* 375(9):819–829.

Zeng, Z., and Bromberg, Y. (2019) Predicting Functional Effects of Synonymous Variants: A Systematic Review and Perspectives. *Frontiers in Genetics* 10(914):1-15.

Zhang, K., Fei, S., Lu, J., YangQifan Yang

, Q., Yang, J., Zeng, Y., Liu, Y., Xiong, G., Jiao, X., Tao, Y., Yuan, Z., Gao, W., Wei, Z., Liu, D., XuTongfu Xu

, T., CaiShangli Cai

, S., Li, F., Cao, R., and Liu, L. (2018) The comparison of tumor mutational burden (TMB) in patients of early and late stage lung adenocarcinoma in China. *Journal of Clinical Oncology* 36(15_suppl):8520-8520.