Improving the Monitoring of Multiple Myeloma Patients Using Blood Based Markers of Disease

Lauren Campbell

A thesis submitted in partial fulfilment of the requirements of Oxford Brookes University for the degree of Doctor of Philosophy

November 2022

Abstract

Multiple myeloma (MM) is a haematological malignancy characterised by the proliferation of clonal plasma cells in the bone marrow, usually leading to the secretion of a monoclonal immunoglobulin termed the M-protein. Multiple myeloma is a complex disease, and despite great advances in both detection and treatment of the disease, there remains a vast amount of interest in improving patient outcomes. In the era of novel agents and clinical trials, patients with MM are living longer and there is a requirement for frequent monitoring to assess treatment response and signs of relapse. Despite these improvements in treatment, even those patients achieving deep responses are still relapsing due to the presence of very low levels of residual disease, termed minimal residual disease (MRD). Increasingly sensitive techniques are needed to detect this MRD and, as MM is a bone marrow based disease, the majority of techniques focus on bone marrow based assays. However, as bone marrow sampling is invasive, painful and unpleasant for patients it is important to investigate whether blood-based assays could be equally or more informative of disease status.

This project therefore aimed to obtain evidence as to whether blood-based assays can act as a safe and effective marker for disease in the bone marrow of treated multiple myeloma patients. Firstly, a bone marrow flow cytometry assay was validated to act as the gold standard technique. Next, three blood-based assays were investigated: circulating tumour plasma cell measurement, heavy/light chain measurement and mass spectrometry evaluation. This study shows the significance of blood-based assays at different disease stages in both transplant and non-transplant treated patients. These results lead to the suggestion of a testing algorithm focused on the addition of heavy/light chain measurement which can be pre-emptively used to influence decisions regarding bone marrow analysis and treatment, ultimately improving the patient experience.

Acknowledgements

Firstly, I would like to thank my supervisors. Thank you to Ross Sadler for supporting me in applying for and completing this PhD and always believing I could do it. Thank you to Lisa Ayers for your constant support and encouragement. Thank you also to Dave Carter and Ryan Pink for always being there to answer my questions and to Karthik Ramasamy for his support in my development and his collaboration.

Thank you to my colleagues in Immunology. Special thanks to Liz Bateman for your support and reassurance. Also, a special thank you to Remi and Zaynab for the encouragement, coffee and biscuits! Thank you also to all my family and friends outside of Immunology who are always so supportive despite not understanding what I do or why.

Thank you to all the research study staff across the different sites of this study and to Oscar Berlanga from Binding Site for his assistance in the with the mass spectrometry analysis of samples.

Thank you to the National Institute for Health and Care Research (NIHR) for the funding of this study as part of the Clinical Doctorate Research Fellowship. This funding has been invaluable in progressing my career as a clinical academic.

Most importantly, thank you to the participants of this study and their enthusiasm for the aims of this work.

Contents

List of Figures

Chapter 1 – Introduction

Chapter 2 - Methods

Chapter 3 - Validation of a next generation flow cytometry assay for bone marrow and peripheral blood samples

Chapter 4 - Impact of bone marrow minimal residual disease and circulating tumour plasma cell status on survival in multiple myeloma patients

Chapter 5 - Immunoglobulin heavy/light chain analysis in the measurement of monoclonal proteins in multiple myeloma patients

List of Tables

List of Abbreviations

aPC – Abnormal plasma cell

ASCT – Autologous stem cell transplant

ASO-qPCR – Allele specific oligonucleotide quantitative PCR

BM – Bone marrow

CI – Confidence intervals

CLP - Common lymphoid progenitor

CR – Complete response

CSR – Class switch recombination

CT – Computerised tomography

CTPC – Circulating tumour plasma cell

CZE – Capillary zone electrophoresis

EDTA - Ethylenediaminetetraacetic acid

EQA – External quality assessment

FBC – Full blood count

FCS – Foetal calf serum

FDA – United States Food and Drug Administration

FISH – Fluorescence in situ hybridisation

FLC – Free light chain

FLCr – Free light chain ratio

FSc – Forward scatter

GC – Germinal centre

HDT – High dose therapy

HLC – Heavy/light chain

HLCr – Heavy/light chain ratio

HR – Hazard ratios

IFE – Immunofixation

iFLC – Involved free light chain

Ig – Immunoglobulin

iHLC – Involved heavy/light chain

IMWG – International Myeloma Working Group

κ – Kappa free light chain

λ – Lambda free light chain

LLOQ – Lower limit of quantitation

LOB – Limit of blank

LOD – Limit of detection

MFC – Multi-parameter flow cytometry

MM – Multiple myeloma

M-protein – Monoclonal immunoglobulin

MR – Minimal response

MRD – Minimal residual disease

NDMM – Newly diagnosed multiple myeloma

NGF – Next generation flow cytometry

NGS – Next generation sequencing

NHS – National Health Service

NK - Natural Killer

OS – Overall survival

PB – Peripheral blood

PBS – Phosphate buffered saline

PC – Plasma cell

PET – Positron emission tomography

PFS – Progression free survival

PR – Partial response

RRMM – Relapsed/refractory multiple myeloma

RT – Room temperature

sCR – Stringent complete response

SHM – Somatic hypermutation

SPEP – Serum protein electrophoresis

SSc – Side scatter

uFLC – Uninvolved free light chain

uHLC – Uninvolved heavy/light chain

UPEP – Urine protein electrophoresis

VGPR – Very good partial response

Chapter 1 - Introduction

1.1 Multiple Myeloma

1.1.1 Background of multiple myeloma

Haematological malignancies are blood cancers which arise in blood and lymph-forming tissues and are the fourth most common cancer in the UK (Roman *et al.,* 2022). Haematological malignancies are classified according to the cell lineage involved; lymphoid malignancies originate from lymphocytes, including both the T and B cell lineages. Examples of lymphoid malignancies include non-Hodgkin lymphoma, lymphocytic leukaemias and multiple myeloma (MM). MM is part of a group of diseases called plasma cell dyscrasias, which gives an indication of the abnormal malignant plasma cells in the bone marrow (BM) which underlie this disease. MM is a malignancy of older adults, with a median age at diagnosis of 66 years (Kyle *et al.,* 2003). The annual incidence in the UK is close to 6,000, with this expected to continue to increase and double in the next 15 years. Annual UK death rates are around 3,100. (CRUK, 2022, Ríos-Tamayo *et al.,* 2018). Worldwide, there are approximately 154,000 cases and 101,000 deaths per year attributed to MM (Fitzmaurice *et al.,* 2017). Both the incidence and mortality rate of MM increase with age (Tomlinson, 2018). Future increases in the incidence of MM will be due to both population growth and an aging population. Overall survival rates in MM have vastly improved in recent years following advances in drug therapy and optimal use of autologous stem cell transplants (ASCT) (Kumar *et al.,* 2016). The median survival has increased from 3 years to 6 years in the past two decades, with 52.6% of UK MM patients surviving for 5 years or more (Rollig *et al.,* 2015, ONS, 2019). Despite these improvements, MM remains an incurable disease and the majority of patients will relapse and require further treatment (Sonneveld and Broijl, 2016). Relapsed MM is the recurrence of the disease after an initial response to therapy. Careful attention is needed in the monitoring of patients to detect and treat disease relapse before further morbidity occurs. It is necessary to ensure that treatment is neither too early, leading to unnecessary side effects and costs incurred, nor too late and close monitoring of patients through regular laboratory assays helps guide therapeutic decision making (Fernandez de Larrea *et al.,* 2014).

1.1.2 Biology of multiple myeloma

The normal cell which becomes cancerous in MM is a cell of the haematopoietic system called a plasma cell. This is an end-stage B cells which produces functional immunoglobulin to protect against infections (Boyle *et al.,* 2014). B cells have a central role in adaptive immunity and produce a diverse range of immunoglobulins, also known as antibodies, to fight infection. Figure

1.1 shows the development of a plasma cell from a stem cell in the bone marrow through to a mature B cell exiting a germinal centre as a plasma cell or memory cell. The long-lived plasma cells can then return to the bone marrow and can be kept alive for long periods of time by survival signals from the bone marrow microenvironment. It is these long-lived plasma cells which become cancerous in MM.

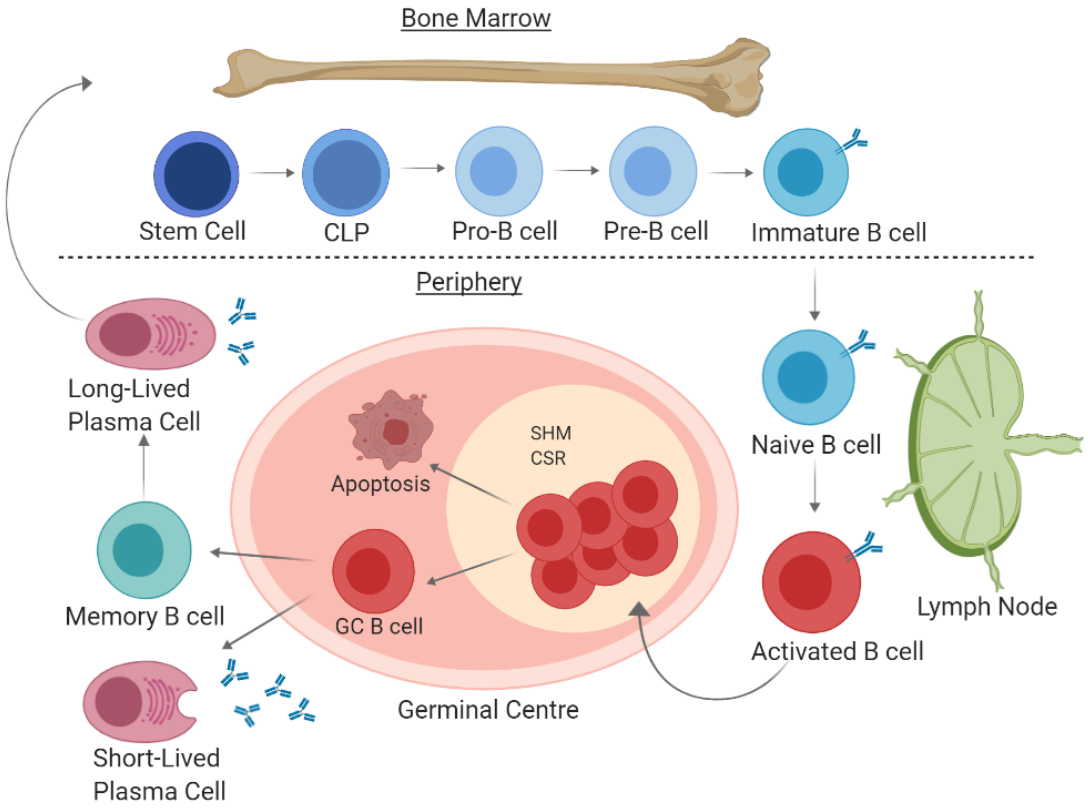


Figure 1.1: B cell development and generation of plasma cells

B cells develop from haematopoietic stem cells in the bone marrow. After maturation in the bone marrow B cells migrate to lymphoid tissues as naïve B cells. Upon activation by a specific antigen these B cells can enter a germinal centre. Within the germinal centre the B cell undergoes clonal expansion, class switch recombination (CSR) and somatic hypermutation (SHM). CSR involves switching of the isotype of the antibody produced by the B cells. SHM increases the affinity of the antibody for the antigen and involves multiple rounds of selection and apoptosis. The germinal centre produces high-affinity antibody-secreting plasma cells and memory B cells. The memory B cells can terminally differentiate into antibody-secreting long lived plasma cells which can reside indefinitely in the bone marrow when given the right survival signals by the bone marrow microenvironment.

CLP: common lymphoid progenitor GC: Germinal Centre

Created with BioRender.com

Normal plasma cells located in the BM produce specific immunoglobulins against a target and once this target has been eliminated the majority of plasma cells will undergo apoptosis, whilst some will remain as long-lived plasma cells in the BM to act as memory cells (Greaves, 2000). MM is the outcome of these plasma cells, also called abnormal plasma cells (aPCs), creating neoplasms, initially in the BM. The development of MM is thought to occur through a number of sequential genetic hits in the germinal centres, that deregulate the behaviour of a normal plasma cell (Da Via *et al.,* 2020). Through large population-based studies it was found that MM was consistently preceded by the premalignant plasma-cell proliferative disorder monoclonal gammopathy of unknown significance (MGUS) (see section 1.1.4) (Landgren *et al.* 2009, Weiss *et al.* 2009). There are shared genetic mutations between MGUS and MM tumours such as immunoglobulin heavy locus (IGH) translocations, aneuploidy, chromosome 13 deletions and dysregulation of a cyclin D gene (Zingone and Kuehl, 2011). These are followed by multiple genetic molecular events associated with the evolution of MGUS to MM, with somatic mutations conferring a selective advantage to subclones of abnormal plasma cells which then become dominant (Hemminki *et al.,* 2021).

In myeloma cells, many of the genetic mutations which give the cells cancerous properties are found in the IGH region on chromosome 14 (Walker *et al.,* 2014). These mutations are primarily chromosomal translocations, but different mutations can occur throughout disease progression. The mutations juxtapose oncogenes into the proximity of IGH enhancers which drive expression of the translocated oncogene (González *et al.,* 2007). Table 1.1 shows some of the main genetic alterations found in MM plasma cells and their association with disease prognosis.

Table 1.1: Table of common genetic alterations found in MM plasma cells

Mutation Type	Chromosomes involved	Risk factor for shortened survival	Frequency in MM patients %	References
Trisomy	3, 5,7, 9, 11, 15 or 17	Standard risk	42-57%	Walker *et al.,* 2010 Kumar *et al.,* 2012 Sonneveld *et al.,* 2016 Rajkumar *et al.,* 2020 Hanamura, 2022
Translocation	t(4;14)	High risk	10-15%	
	t(14;16)	High risk	4-5%	
	t(11;14)	Standard risk	15-20%	
	t(14;20)	High risk	<1-1%	
Deletion	del(13)/del(13q)	High risk	44-50%	
	Del(17p)	High risk	7-10%	
	Del(1p)	High risk	30%	
Gain	1q (+1q)	High risk	35-40%	

The clonal cells of MM initially remain in the BM where they form an important relationship with the BM microenvironment which contributes to the clonal cells long-term survival (Abe, 2011). With the advancement of disease in some patients some of these clonal cells develop the capability to leave the BM and enter the circulation and proliferate in extra-medullary locations such as the spleen, liver and mucosa-associated tissues (Blade *et al.,* 2011). These patients can present as either plasma cell leukaemia or extra-medullary myeloma.

1.1.3 M-protein in multiple myeloma

The cells involved in MM are terminally differentiated clonal plasma cells which all produce the same monoclonal immunoglobulin (Ig), known as the M-protein. In contrasting behaviour to a population of normal plasma cells, which produce a spectrum of immunoglobulins, a population of abnormal MM plasma cells will produce an M-protein of only one heavy and/or light chain type. An individual plasma cell can only produce a single antibody of a single isotype but can produce thousands of copies of this antibody per second (Lam and Bhattacharya, 2018). The combination of this and the clonal expansion of the abnormal plasma cells being far greater than normal plasma cells means the M-protein is produced in large amounts over any of the polyclonal antibodies produced in normal responses (Bolomsky and Young, 2022). The presence of an elevated level of this M-protein is an important serological feature of MM and the M-protein can appear in the serum or urine before any clinical symptoms are recognised. A gammopathy is a disturbance in the body's production of immunoglobulin. Therefore MM and its precursor conditions are also known as monoclonal gammopathies.

A biomarker is a biological parameter which is used as an indicator of normal biological processes, pathogenic processes, or response to therapy (BDWG, 2001). In MM the M-protein acts as a biomarker which is used for diagnosis, prognosis and monitoring. Using techniques described in Table 1.4 the monoclonal M-protein is detected in excess above the usual polyclonal antibody background present in a healthy immune system. This M-protein is different in each patient, but similarities exist as to the presence of certain heavy and/or light chains. During a normal immune response B cells, and therefore plasma cells, can switch which Ig heavy chain constant region will form the immunoglobulin (IgM, IgG, IgA, IgD or IgE) alongside a pair of one of two Ig light chains, either kappa (κ) or lambda (λ) to produce an intact Ig (see Figure 1.2a). A clone of myeloma plasma cells will only produce one type of Ig (M-protein) e.g. only IgGκ. A subset of MM patients will produce an M-protein which is a free light chain (FLC) only e.g. κ light chains not attached to a heavy chain constant region (see Figure 1.2b). This is termed light chain multiple myeloma (LCMM).

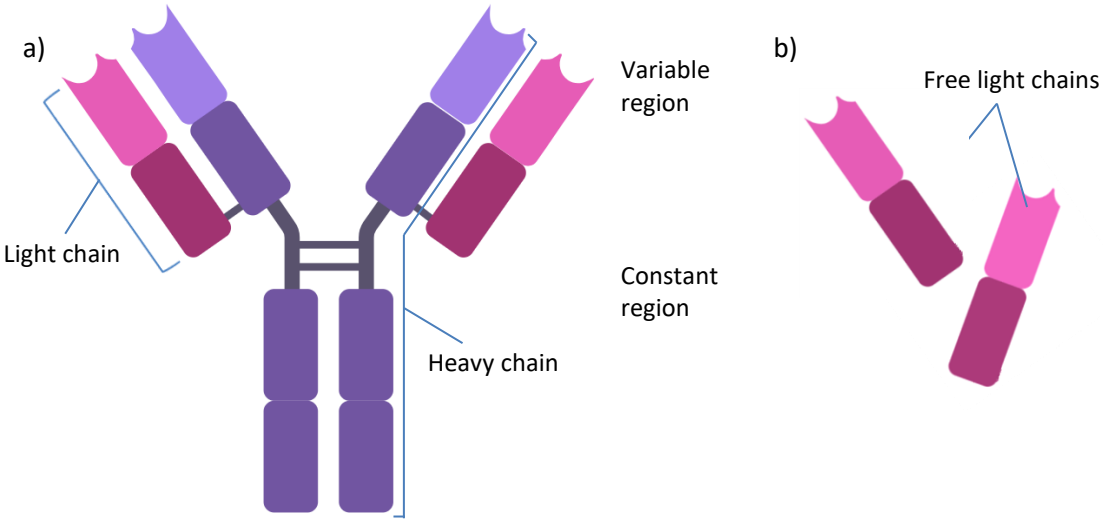


Figure 1.2: a) Intact IgG Immunoglobulin Structure b) Free light chains

1.2a) The M-protein is made up of four units, 2 light chains and 2 heavy chains. Heavy chains can be one of five types and light chains can be one of two types, either kappa or lambda. Shown is an IgG molecule but all Ig isotypes have the same general structural features. Each IgG molecule consists of two identical heavy chains (purple) linked by disulphide bonds, and each heavy chain is linked to a light chain (pink) by a disulphide bond. A given Ig molecule will only have either κ or λ light chains, never both. The constant regions determine the chain isotype whilst the variable regions are involved in antigen binding.

1.2b) Excess FLC are produced by both normal and abnormal plasma cells and are released into the circulation unattached to a heavy chain. In humans κ light chains are produced more often than λ light chains, giving a κ/λ ratio of 2:1 (Janeway *et al.,*2001). When a clone of abnormal plasma cells produces FLC either in excess to intact M-protein or as the M-protein itself (as in LCMM) this ratio will be changed.

FLC can be detected separately from intact Ig and both the concentration of the κ and λ light chains and the κ/λ ratio (FLCr) is used to determine monoclonality (see Table 1.4). FLC are produced by plasma cells (both normal and abnormal) in excess to heavy chains. Under normal circumstances these excess chains are swiftly secreted and then cleared by the kidneys, with FLC having a short serum half-life of 2-6 hours (Hutchison *et al.,* 2008, 2011). Non-secretory and oligo-secretory forms of MM also exist, where either no detectable M-protein or very low levels of M-protein are produced (Corso and Mangiacavalli, 2017, Migkou *et al.,* 2020). MM can be associated with any of the heavy and light chain types but some are far more common than

others (see Table 1.2). Both the clonal proliferation of abnormal plasma cells in the bone marrow and the excess production and circulation of M-protein contribute to the symptoms associated with MM (see Section: 1.1.5).

Table 1.2: Isotype distribution in MM

Isotype	Cases (%)	References
IgG	52-60%	Kyle *et al.,* 2003, Rafae *et al.,* 2018, Migkou *et al.,* 2020
IgA	21-27%	
IgM	<1%	
Light chain only	13-20%	
IgD	2%	
IgE	<1%	
Non-secretory/Oligo secretory	1-12%	

1.1.4 Development of multiple myeloma

The definitions of the different monoclonal gammopathies, which consist of MM and its precursors, comprise of clinical symptoms, evidence of end organ damage and detection of the M-protein (see Table 1.3 for definitions). As MM is a heterogeneous disease, rates of progression through the disease stages vary. It has been recognised that there are a number of stages in the development of MM, including two asymptomatic phases termed monoclonal gammopathy of undetermined significance (MGUS) and smouldering multiple myeloma (SMM) (Kyle and Rajkumar, 2009). It has been found that MM is consistently preceded by these asymptomatic premalignant conditions (Landgren *et al.,* 2009). MGUS is a common pre-malignant disorder which a systematic review found in 3.2% of 50 years or older Caucasians (Wadhera and Rajkumar, 2010). In a longitudinal study of >1000 MGUS patients being followed up for a median of 34.1 years the cumulative risk of progression from MGUS to MM was 10% at 10 years post diagnosis, 18% at 20 years, 28% at 30 years and 36% at 40 years (Kyle *et al.,* 2018). This gives a chance of progression of around 1% per year; therefore the majority of patients with MGUS will not progress to symptomatic disease. SMM is an intermediate stage between MGUS and MM and defines patients who meet the laboratory criteria for MM but who show no signs of end-organ damage (Willrich and Katzmann, 2016). The risk of progression of SMM to MM is higher than for MGUS, at about 10% per year for the first 5 years of diagnosis, 3% for the next 5 years and 1% per year for the next 15 years (Kyle *et al.,* 2007). Individual times to progression are however far more variable than these statistical progression rates (Wu *et al.,* 2018).

Table 1.3: International Myeloma Working Group (IMWG) diagnostic criteria for MM and related disorders

Disorder	Definition
Non-IgM MGUS	All 3 criteria must be met for the disease to be diagnosed: • Serum M-protein (non-IgM type) <3 g/dL • Clonal bone marrow (BM) plasma cells <10% • Absence of end-organ damage, such as hypercalcemia, renal insufficiency, anaemia, and bone lesions (CRAB) that can be attributed to the plasma cell proliferative disorder
Light-Chain MGUS	All 6 criteria must be met for the disease to be diagnosed: • Abnormal FLC ratio (<0.26 or >1.65) • Increased level of the involved light chain • No immunoglobulin heavy chain expression on immunofixation • Absence of end-organ damage, such as CRAB • Clonal bone marrow plasma cells <10% • Urinary monoclonal protein <500 mg/24 hours
MGRS	Presence of nephrotoxic M-protein in kidneys with different degrees of renal function impairment.
SMM	Both criteria must be met for the disease to be diagnosed: • Serum M-protein (IgG or IgA) 3 g/dL, or urinary M-protein 500 mg/24 h, and/or clonal bone marrow plasma cells 10% to 60% • Absence of myeloma-defining events or amyloidosis
MM	Both criteria must be met for the disease to be diagnosed: • Clonal bone marrow plasma cells 10% or biopsy-proven bony or extramedullary plasmacytoma Any 1 or more of the following myeloma-defining events: • Evidence of end-organ damage that can be attributed to the underlying plasma cell proliferative disorder, specifically: • Hypercalcemia: serum calcium >25 mmol/L (>1 mg/dL) higher than the upper limit of normal or >275 mmol/L (>11 mg/dL) • Renal insufficiency: creatinine clearance <40 mL/min or serum creatinine >177 µmol/L (>2 mg/dL) • Anaemia: haemoglobin value >2 g/dL below the lower limit of normal or a haemoglobin value <10 g/dL • Bone lesions: ≥1 osteolytic lesion on skeletal radiography, CT, or PET/CT • Clonal bone marrow plasma cell percentage 60% • Involved/uninvolved serum FLC ratio 100 (involved FLC level must be 100 mg/L) • >1 focal lesion on MRI studies (≥5 mm in size)

Table 1.3 adapted from Rajkumar *et al.,* (2014); Gandolfi *et al.,* (2018); Willrich *et al.,* (2018)
MGRS = Monoclonal gammopathy of renal significance

1.1.5 Clinical manifestations of multiple myeloma

MM is a malignant disease which features end-organ damage and significant patient morbidity (Palumbo and Anderson, 2011). Clinical manifestations of MM are related to tissue and/or organ impairment caused by the abnormal plasma cells themselves or by the excessive M-protein produced by these cells. The CRAB acronym is used to describe these manifestations: elevated serum Calcium, Renal function impairment, Anaemia and/or Bone involvement (IMWG, 2003).

Bone marrow infiltration of rapidly expanding abnormal plasma cells replaces normal haematopoietic tissue and alters the bone marrow microenvironment leading to anaemia. These changes in the bone marrow microenvironment also cause both an increase in osteoclast activity and a concurrent decrease in osteoblast activity. This leads to lytic bone lesions and hypercalcaemia. The excessive free light chains are the primary cause of renal damage in MM, overwhelming the kidneys filtering systems and forming obstructive casts (Goldschmidt *et al.,* 2000). A review of 1027 sequential patients newly diagnosed with multiple myeloma at the Mayo Clinic, USA between 1985 and 1998 found the symptoms at diagnosis to be anaemia (73%), bone pain (58%), elevated creatinine (48%), fatigue/weakness (32%), hypercalcaemia (28%), and/or weight loss (24%) (Kyle *et al.,* 2003).

1.1.6 Overview of multiple myeloma testing

Due to the complex nature of MM and the M-protein, no single laboratory test is used for diagnosis and monitoring; instead a testing strategy must be defined by the laboratory to effectively detect this disease. Current tests used to identify an M-protein are total immunoglobulin measurements, serum protein electrophoresis (SPEP), urine protein electrophoresis (UPEP), immunofixation electrophoresis (IFE) and serum free light chains (FLC) which have varying levels of sensitivity. These tests give indirect evidence of the presence monoclonal PCs in the BM and are termed conventional methods. These are accompanied by tests to help to detect tissue and organ impairment including full blood count, calcium levels, renal assessment using serum creatinine, imaging for bone lesions and bone marrow films to determine clonality (see Table 1.3).

1.1.7 Conventional methods

Total immunoglobulin measurement of IgG, IgA and IgM concentration in patients' sera is an initial step in M-protein detection and can show a rise in the concentration of a certain immunoglobulin, and potentially a fall in the concentration of those not produced by the clonal

abnormal plasma cells. However, the basic concept of M-protein detection is the differentiation between monoclonal and polyclonal immunoglobulins; the techniques used to determine total immunoglobulins do not distinguish between monoclonal and polyclonal immunoglobulins, therefore further techniques are needed to identify and characterise the M-protein. SPEP with measurement of the M-protein concentration, either by capillary zone electrophoresis (CZE) or densitometry is currently the gold standard in the assessment of patients with MM both at the time of diagnosis and during therapy (Kyle and Rajkumar, 2009). In this technique the M-protein present has a specific electrophoretic mobility which means it is seen as a distinct band against the background of polyclonal antibodies which all have slightly different mobilities. For patients with LCMM FLC are rapidly cleared by the kidneys into the urine and therefore are often not detectable by SPEP. In these cases UPEP can be performed to detect these monoclonal FLCs being excreted. IFE is a qualitative test which determines the isotype of the M-protein. IFE is more sensitive than SPEP but cannot quantify the M-protein (Willrich and Katzmann, 2016).

The FLC assay measures levels of free κ and free λ light chains in the serum and can calculate κ/λ ratios (FLCr) for which normal ranges have been produced. The calculation of ratios is important as this accounts for the variable renal clearance of the different light chains, variations in blood volume and tumour suppression of the non-tumour derived FLC whilst also defining clonality (Bradwell *et al.,* 2003). This test was added into the international guidelines for assessing monoclonal gammopathies in 2009 (Dispenzieri *et al.,* 2009). Although the detection of monoclonal free light chains has historically been performed by UPEP, obtaining urine samples for immunological investigations from patients is challenging; studies from both primary and secondary care have shown that, at best, only 40% of patients have UPEP performed and in some studies, this is as low as 5% (Hill *et al.,* 2006, Robson *et al.,* 2009). The FLC 'Freelite'™ assay released by The Binding Site (Birmingham, UK) quantifies free κ and λ light chains using antisera which is directed against epitopes which are only exposed when the light chain is unbound to a heavy chain as these sites are involved in the strong binding of light chains to heavy chains. Therefore the only light chains quantified are those that are free, even in a solution with a high concentration of intact immunoglobulins (Bradwell *et al.,* 2001). Studies have investigated the role of FLC measurement in a number of different patient populations including non-secretory MM (Drayson *et al.,* 2001), light chain amyloidosis (Dispenzieri *et al.,* 2006), MGUS (Rajkumar *et al.,* 2005), SMM (Dispenzieri *et al.,* 2008), LCMM (Bradwell *et al.,* 2003), newly diagnosed MM (NDMM) (Kyrtsonis *et al.,* 2007) and relapsed/refractory MM (RRMM) (Vij *et al.,* 2015). This assay has been found to be useful in both the diagnosis, prognosis and monitoring of these patient populations. The International Myeloma Working Group (IMWG) produces response

criteria using these conventional methods, categorising patients into different response levels (Durie *et al.,* 2006) (see section 1.2.1). See Table 1.4 for a summary of these conventional methods.

Table 1.4: Detecting the M-protein using conventional methods

Method	Description of method	Pros	Cons
Total Immunoglobulin Measurement	Nephelometric and turbidimetric techniques to quantify total IgG, IgA and IgM. These techniques determine the concentration of immunoglobulins in serum using the scattering or absorbance of light shone through the solution respectively. Increased scattering or absorbance of light is associated with an increased concentration of immunoglobulins in the solution.	-Gives an overall immunoglobulin picture. -Can indicate the presence of an M-protein if large enough. -Can show whether immunoparesis is present which can be prognostic.	- Cannot identify the M-protein at lower levels. - Cannot characterise the M-protein isotype. - Does not distinguish between monoclonal and polyclonal immunoglobulins.
Serum Protein Electrophoresis (SPEP)	Electrophoresis is the separation of molecules using electric currents. Serum molecules are separated from each other based on size and net charge and the proteins in serum separate into 6 different fractions: albumin, α1, α2, β, and γ globulin fractions. Immunoglobulins migrate to the γ region The presence of an M-protein is usually signalled by the occurrence of a homogenous monoclonal band in the γ region, although the M-protein can migrate to the α or β regions. The monoclonal band can then be quantified to give the M-protein concentration.	- Quick and well-established method. -The M-protein can be quantified. -Semi-automated techniques. -Suitable for serial measurements and long-term monitoring.	- M-proteins may be missed if migrating to the α or β regions or cannot be quantified due to the presence of other serum proteins. - Small M-proteins in the γ region can be missed when polyclonal immunoglobulin levels are high - Cannot characterise the isotype of the M-protein. - Free light chain only M-proteins may be missed. - Subjective interpretation of results can lead to inconsistency. - Limit of detection

Method	Description of method	Pros	Cons
Immunofixation Electrophoresis (IFE)	IFE is a combination of electrophoresis and immunoprecipitation. Serum proteins are first separated by electrophoresis and then anti-sera specific to either IgG, IgA, IgM, free kappa or free lambda light chains is added. If the M-protein is of that anti-sera type then an immunoprecipitate will form and can be stained for visualisation. In this way the heavy chain and/or light chain of the M-protein can be determined.	- Characterises the M-protein isotype which can be prognostic. - IFE has improved sensitivity against SPEP (lower limit of detection ~0.1 g/dL versus 0.01 g/dL).	- Not a quantitative technique. More labour intensive than other conventional techniques. - Not suitable for serial measurements
Serum free light chain (FLC)	A turbidimetric or nephelometric immunoassay which measures levels of free κ and free λ light chains in the serum and can calculate κ/λ ratios. The assay uses polyclonal antibodies which are directed against epitopes on the free light chain which are only exposed when the light chain is unbound to a heavy chain.	- Detects free light chain only M-proteins. - Calculation of the κ/λ ratio accounts for the variable renal clearance of the different light chains, variation in blood volume and tumour suppression of the non-tumour light chain. - Automated method.	- Only characterises the free light chains. - Affected by kidney function, infection, inflammation. - Sensitivity for detecting low levels of free light chains has been questioned. - Antigen excess can occur at high free light chain concentrations.

Information adapted from: Wood *et al.,* (2010); Zhu *et al.,* (2021)

1.1.8 Treatment of multiple myeloma

High-dose induction chemotherapy is the standard treatment for all patients, using a combination of novel agents (proteasome inhibitors, immune modulators) and dexamethasone. For younger and fitter patients this is then followed by an ASCT. The patient will first receive 4-8 cycles of induction chemotherapy and then have their stem cells collected after stem cell

mobilisation. The introduction of induction chemotherapy followed by ASCT has improved progression free survival (PFS; time from start of treatment to disease progression) and overall survival (OS; time from start of treatment to death) in a number of randomised trials (Attal *et al.,* 1996, Child *et al.,* 2003, Fermand *et al.,* 2005). After induction and ASCT patients will undergo consolidation and/or maintenance therapy. Consolidation therapy is different to maintenance therapy as it is short term and is used after successful induction or ASCT therapy to deepen and perhaps prolong responses (Al Hamed *et al.,* 2019). Maintenance therapy lasts for more cycles than both induction and consolidation therapy but is less intensive. It is used to delay disease progression either after induction therapy or after an ASCT has been carried out. In elderly patients there must be a finer balance between treatment efficacy and toxicity and ASCT is often not carried out, with the cut off age for performing a transplant in Europe being between 65 and 70 years (Ludwig *et al.,* 2014). This contributes to the disparity in OS seen between young and old MM patients which is unsurprising as age, co-morbidities and patient fitness all play a role in treatment allocation and intensity (Rosko *et al.,* 2017). A longitudinal study which included both transplant and non-transplant patients confirmed the clinical survival advantage still present with transplant patients (Oliva *et al.,* 2016). However, a 2011 study examining survival data found that older patients (>70 years) were indeed beginning to see improved survival outcomes and that survival rates were accelerating over time (Pulte *et al.,* 2011). There are now a large number of therapies for clinicians to choose from and patients are often entered into clinical trials. With the large number of treatments on offer and different treatments used during the course of the disease (diagnosis, disease control, progression, relapse) the need to quickly assess a patient's response is more important than ever.

Although both the PFS and OS of patients has greatly improved with advanced treatments the majority of MM patients will still relapse (Pulte *et al.,* 2011, Kumar *et al.,* 2014, Majithia *et al.,* 2016, Sonneveld and Broijl, 2016, Chen *et al.,* 2017, Goswami *et al.,* 2019). This is primarily due to incomplete eradication of the disease, with residual disease still being present after treatment. When the amount of residual disease is very low and undetectable by standard techniques it is termed minimal residual disease (MRD). In a typical course of MM response to treatment is followed by increasingly aggressive relapses until ultimately the disease does not respond to any treatments and the patient dies (Rustad *et al.,* 2017). MM is an incredibly heterogeneous disease with varying progressive courses, responses to treatment and outcomes. There is a growing need to enhance the differentiation of patients to not only allow for altering of treatment but also to allow informed discussions between clinician and patient and to work towards personalised therapies.

1.2 Minimal Residual Disease

1.2.1 Response assessment in multiple myeloma

In the era of novel agents and clinical trials, patients with MM are living longer and there is a requirement for frequent monitoring to assess treatment response and signs of relapse. Assessing disease burden in response to therapy is a significant part of disease management. It is important to detect response to treatment in order to either reduce therapy, and therefore toxic therapeutic side effects, or to detect when patients are not responding and change the regime to improve the response. There is not only a cost benefit to stopping therapy or changing ineffective therapy but a benefit to the patient's quality of life also, with side effects to treatment including second primary malignancies, infections, gastrointestinal disorders and rashes (Palumbo *et al.,* 2014, McCarthy *et al.,* 2017, Dimopoulos *et al.,* 2019).

The IMWG created uniform response criteria which standardised disease assessment across clinical trials and health services, initially to enable treatment successes and responses to be more easily compared (Durie *et al.,* 2006). These criteria range from poor responses to treatment, termed minimal response (MR), to the highest response levels, complete response (CR) and stringent complete response (sCR) (see Table 1.5).

Table 1.5: IMWG Standard Response Criteria.

Response Status	Response Criteria
Stringent Complete Response (sCR)	CR (below) plus: Normal FLCr AND Absence of abnormal clonal PCs in BM using immunofluorescence or immunohistochemistry
Complete Response (CR)	Negative IFE of the serum and urine AND <5% plasma cells in BM
Very Good Partial Response (VGPR)	Serum or urine M-protein only detectable by IFE (not SPEP/UPEP) OR ≥90% reduction in serum M-protein AND Urine FLC <100mg/day
Partial Response (PR)	≥50% reduction of serum M-protein OR ≥90% reduction of urine FLC or <200mg/day (Serum FLC measurement may be used if M-protein not measurable by SPEP or UPEP)
Minimal Response (MR)	≥25% reduction of serum M-protein OR ≥50% reduction in urine FLC

Adapted from Durie *et al.,* (2006) and Kumar *et al.,* (2016)

With developments in treatment 100% of patients will respond to treatment and up to 80% will achieve CR (Mailankody *et al.,* 2015). Achievement of CR is associated with higher rates of PFS compared with patients who only achieve VGPR or lower (Davies *et al.,* 2001, Harousseau *et al.,* 2010, Gay *et al.,* 2011). In 2008 a prospective analysis of the influence of response status on PFS and OS in patients receiving HDT and ASCT showed that the response status at the post-transplantation time point (100 days post-ASCT) was a surrogate marker for survival, with CR patients having significantly longer PFS (median 61 vs 40 months; $p<0.00001$) than those in VGPR (Lahuerta *et al.,* 2008). The same trend was also true for OS but to a non-significant degree ($p=0.1$). Whilst the achievement of CR has not consistently been shown to improve overall response (Facon *et al.,* 2007, Paiva *et al.,* 2008, Usmani *et al.,* 2018). A meta-analysis of the association between CR and outcomes in transplant-eligible patients concluded that attainment of CR/sCR leads to improved PFS and OS, especially when treatment includes novel agents (van de Velde *et al.,* 2007).

Despite CR/sCR being achieved by an increasing number of patients, the majority of MM patients are still relapsing, suggesting the presence of undetected residual disease in these patients and that the current ways of defining CR are not sensitive enough to evaluate or predict the long term response to therapy (Chanan-Khan and Giralt, 2010). The heterogeneity in the outcomes seen in CR patients will be due in part to the presence of minimal residual disease (MRD) not detected by the serological or cytological methods used to define CR/sCR. In a long-term follow up study it was show that the majority (94%) of patients in CR who achieved long-term survival (10 years PFS) also proved to be MRD negative (Barlogie *et al.,* 2014). A recent meta-analysis showed that MRD negativity was associated with favourable survival outcomes regardless of the response status (CR, VGPR etc.) (Munshi *et al.,* 2020). MRD has therefore become the focus of intense research over the last decade.

1.2.2 Causes and biology of MRD

It is the abnormal plasma cells which evade initial therapy that contribute to MRD and the subsequent relapse of MM patients. Survival signals and drug-resistance, both mediated by the bone marrow environment, and the persistence of MM stem cells are necessary for MM plasma cell survival in the face of treatment and the persistence of residual disease (Brennan and Matsui, 2009, Shain *et al.,* 2009, Meads *et al.,* 2016). The presence of sub-clones of MM cells displaying different drug sensitivities is termed intratumoral heterogeneity (Barlogie *et al.,*2014). As previously mentioned in section 1.1.2, multiple myeloma is consistently preceded by the precursor MGUS, both of which are characterised by monoclonal plasma cells in the bone

marrow (Weiss et al. 2009). This progression of the disease can lead to existence of two clones of plasma cells, one progressing to active myeloma and the other remaining as a secondary MGUS clone which is suppressed or even eliminated by the active clone (Kyle et al. 2003, Campbell et al. 2017). This was found in a small subset, around 1%, of active myeloma patients (Campbell et al. 2017). When investigating the effect of anti-myeloma therapy on these two co-existing clones it was found that this therapy was more effective against the MM clones the MGUS clones (Campbell et al. 2017b). This raises the possibility of MRD in some patients being the result of the continuing existence of an MGUS plasma cell clone, rather than resistance of the active myeloma clone. However, this work did not include flow cytometric or genetic evaluation of the plasma cell clones and further work will be needed in clinical trials to determine the effect of these clones on relapse.

The evaluation of MRD is an area under great scrutiny and development. With the use of more and more effective therapies, detecting these residual cells requires increasingly sensitive techniques. In clinical trials MRD is being used to monitor patients and treatment efficacy; however this has not yet become a standard tool in routine practice (Sonneveld *et al.,* 2017). The current status of MRD detection is as a research test which has very promising clinical utility as a risk-stratification tool for therapy decisions. It is important that MRD is now looked at as a potential surrogate endpoint for PFS and/or OS which can therefore be used to aid in clinical decisions such as ending or switching treatments. The financial advantage of detecting and acting on MRD is apparent: new drugs are not only aimed at eradication of disease but are increasingly aimed at maintenance therapy. Ensuring the right patient is on the correct treatment is paramount given the cost of such drugs, coupled with the survival advantage (Teitelbaum *et al.,* 2013).

Studies have confirmed that MRD status (positive, meaning MRD is present, or negative, meaning MRD is absent) is a stronger predictor for progression and survival than the CR/sCR status (Rawstron *et al.,* 2013, Paiva *et al.,* 2008). With newer, more sensitive tests developed to detect MRD, different disease states past CR/sCR can now be defined and prompted the IMWG to include MRD detection in the most recent IMWG response assessment guidelines (Kumar *et al.,* 2016) (see Table 1.6). These criteria specify exact definitions of MRD negativity assigned by different methods of detection; however, patients who are MRD negative by one technique are not necessarily negative by the others, further complicating the classification of response to treatment.

Table 1.6: IMWG MRD Response Criteria

Technique/Result	Criteria definition*
Sustained MRD negative	MRD negative in the bone marrow (by flow or sequencing and imaging as described below), confirmed a minimum of 1 year apart.
Flow MRD negative	Absence of phenotypically abnormal clonal plasma cells by NGF on bone marrow aspirates using the Euro-Flow SOP (or validated equivalent method) with a minimum sensitivity of 1 in 10^5 or greater nucleated cells.
Sequencing MRD negative	Absence of clonal plasma cells by NGS on bone marrow using the LymphoSIGHT platform (or validated equivalent method) with a minimum sensitivity of 1 in 10^5 or greater nucleated cells.
Imaging plus MRD negative	MRD negative by NGF or sequencing as described above plus disappearance of every area of increased tracer uptake found at baseline or a preceding PET/CT or decrease to less mediastinal blood pool standard uptake volume or decrease to less than that of surrounding normal tissue.

** These criteria require achieving CR on the basis of the standard IMWG response criteria*
Table adapted from Kumar *et al.,* (2016)

1.2.3 Methods of detection of MRD

The past decade has seen a 4-fold increase in the number of publications on MRD in MM (Galtseva *et al.,* 2018). The main techniques which have emerged for detecting MRD in MM are multi-parameter flow cytometry (MFC) (see section 1.3) and the more sensitive next-generation flow cytometry (NGF), allele-specific oligonucleotide quantitative PCR (ASO-qPCR), advanced imaging and next-generation sequencing (NGS), which are reflected in the new IMWG response assessment criteria (Table 1.6). The FDA have recently authorised the first NGS assay to detect MRD in the BM of MM patient (Ching *et al.,* 2020). Spanish and UK groups (San Miguel *et al.,* 2002, Rawstron *et al.,* 2002) were the first to demonstrate the prognostic value of MFC MRD status, both showing that MM patients who were MRD negative at day 100 post-ASCT had longer OS and PFS. The large MFC studies carried out by Rawstron and colleagues brought MRD to the forefront of the MM literature (Rawstron *et al.,* 2013) and it is MFC studies which continue to be at the leading edge of MRD publications (Galtseva *et al.,* 2017). Assessment of response and monitoring of disease by MFC, PCR and NGS is commonly used in other haematological malignancies including chronic myelogenous leukaemia (CML), chronic lymphocytic leukaemia

(CLL), and acute lymphoblastic leukaemia (ALL) (Kantarjian *et al.,* 2008, Brüggemann *et al.,* 2012, Thompson *et al.,* 2018).

Although PCR and NGS techniques have until recently always been more sensitive than MFC they have lower applicability to MM patients due to needing both a baseline sample and patient-specific probes (Martinez-Lopez *et al.,* 2015, Puig *et al.,* 2014, Takamatsu *et al.,* 2017, Carulli *et al.,* 2019). A baseline sample may not always be possible due to sampling errors, pre-analytical errors, analytical errors, and patient wishes. Even when a sample is obtained there is still a possibility (~30%) of not being able to produce successful primers (Martinez-Sanchez *et al.,* 2008, Ladetto *et al.,* 2014). Recent advances in MFC technology have increased the sensitivity of these assays down to the 10^{-6} range and this more sensitive technique is now known as next-generation flow cytometry (NGF). When compared to the conventional 8-colour MFC technique, NGF was shown to have higher sensitivity, giving a 47% MRD positivity rate versus 34% by conventional MFC in the same patient population (p=0.003) (Flores-Montero *et al.,* 2017). Here, the general flow cytometry technique is referred to as MFC and the advanced sensitive technique as NGF. Some large studies evaluated the concordance in MRD testing between MFC, ASO-qPCR and NGS, finding a high level of concordance between methods (Martinez-Lopez *et al. ,*2014, Moreau *et al.,* 2019, Medina *et al.,* 2020).

A limitation to all the mentioned MRD detection techniques is the need for a bone marrow sample. As myeloma lesions can be distributed through the bone marrow, a bone marrow aspirate may not successfully pick up abnormal MM cells when taken from a single area. This must always be taken into account when assessing MRD in the bone marrow, and it is therefore recommended to interpret MRD BM results in conjunction with imaging results, such as positron emission tomography (PET). These imaging techniques can allow for detection of focal lesions or extra-medullary disease which can be missed by bone marrow sampling (Moreau *et al.,* 2017) and have been linked with response duration (Rasche *et al.,* 2018). However, PET imaging is not suitable for all MM lesions, has low inter-observer reproducibility and uses technology not available to every centre (Fulciniti *et al.,* 2015). More recently a study found a strong correlation between MRD results by PET/CT and those by MFC of the bone marrow (Zamagni *et al.,* 2020).

When considering the optimal technique to incorporate into routine service a number of factors must be considered. Sensitivity and specificity of the assay are important considerations but also reliability, cost, standardisation, labour time and effects on patients must be considered. For this study NGF MRD detection has been chosen as the gold standard technique due to the higher

applicability to the greatest number of patients, access to flow cytometry in routine NHS laboratories and the presence of flow cytometry skills already present within the study team.

1.2.4 Evidence for utility of MRD

Systematic reviews and meta-analyses in the area of MRD are complicated by many factors including: the heterogeneous definitions used to define MRD and assign response criteria, different detection methods used and therefore different sensitivities, pre-analytical variations, and various treatment conditions, follow up times, and measurement time points. Some of these variations will arise through this study, however it is important to assess the utility of assays in real world settings seen in the NHS. MRD has become such an intense area of research due to the excess of studies showing its prognostic usefulness, some of which are highlighted in Table 1.7.

A meta-analysis of MRD articles published between 1990 and 2016 found a strong association between depth of response and overall outcome in ASCT patients (Munshi *et al.,* 2017). This association was true even when using data from patients undergoing more primitive therapies and using less sensitive MRD assays. Out of a possible 430 articles, 21 met the criteria for this analysis. Median PFS and OS were 26 and 82 months respectively for MRD positive patients, and 54 and 98 months for MRD negative patients. There was a significant association between MRD negativity and improved PFS (HR 0.41; 95% CI 0.36-0.48; $p<0.0001$) and OS (HR 0.57; 95% CI 0.46-0.71; $p<0.0001$) compared with MRD positivity. Due to the nature of the studies used, these results could not be extrapolated to those patients not undergoing ASCT. However, an updated meta-analysis which included over 12,000 patients was performed in 2020 and further established the role of MRD negativity in survival outcomes for both transplant eligible and transplant ineligible populations, including those relapsed/refractory patients (PFS (HR, 0.33; 95% CI, 0.29-0.37; $p<0.001$) and OS (HR, 0.45; 95% CI, 0.39-0.51; $p<0.001$) (Munshi *et al.,* 2020). These analyses supported the inclusion of MRD status as an endpoint in clinical trials, spanning different response states and treatments, different detection methods and varied clinical response levels, showing its broad applicability. The authors concluded that despite the varied detection methods used and the possibility of publication bias, MRD can supersede the prognostic value of CR across all disease settings regardless of the method used and should be used as an endpoint in clinical trials. Another recent meta-analysis evaluated the use of MRD status as a surrogate for PFS in NDMM patients (Avet-Loiseau *et al.,* 2020). They evaluated the data from 6 randomised clinical trials with a total of 2208 MRD samples. Despite these clinical trials using various treatment methods, including both transplant and non-transplant patients,

and different methodologies and therefore sensitivities of MRD testing in the bone marrow the analysis concluded that MRD could indeed be used as a surrogate for PFS in clinical trials for patients with NDMM.

Even though these meta-analyses (Munshi *et al.,* 2017 and 2020, Avet-Loiseau *et al.,* 2020) proved better PFS and OS in treated MRD negative NDMM patients, another trial showed that, despite better PFS and OS curves than MRD-positive patients, around 25% of MRD-negative patients still experience relapse at 36 months (Attal *et al.,* 2017). In these meta-analyses the majority of studies were at the 10^{-4} sensitivity threshold, though the 2020 analysis included studies at 10^{-5} and even 10^{-6} thresholds. The association of MRD negativity with improved survival outcomes was present at all sensitivity levels and there was an association between further improved PFS and OS and MRD negativity at increased sensitivity levels. This research indicates that even more sensitive detection of MRD will be useful in classifying these patients. Multiple other works not included in these meta-analyses have set out to determine the relationship between MRD status and survival (see Table 1.7).

A longitudinal prospect analysis of 50 MM patients concluded that MRD status predicted both clinical and biochemical relapse by a median of 9 and 4 months respectively. This study only included patients who had achieved at least a VGPR response after therapy which included both transplant and non-transplant treatments (Oliva *et al.,* 2016). This study is important as MRD testing was repeated after consolidation, maintenance and further during long-term follow up. Notably, this study showed that those patients who achieved a larger tumour burden decrease had more favourable survival outcomes, reproducing what has previously been shown by Rawstron *et al.* (2015). This study also confirmed that an MRD positive status is an adverse prognostic factor, even in patients who had attained CR. These studies show that depth of response, as well as magnitude of MRD depletion, has an impact on prognosis which is independent of the method of MRD detection. Rawstron *et al.* (2015) proposed that MRD was most useful not when used as a fixed variable, but instead by determining the log reduction in MRD, using it as a continuous variable. They found that there was a 1-year survival benefit for each 1-log depletion in tumour burden in patients who underwent ASCT. If MRD quantitation has the possibility to be more informative and a more powerful predictor than MRD status, then the accurate quantification of MRD is even more necessary. Recently, Kriegsmann *et al.* (2020) performed a thorough evaluation of both NGS and MFC MRD data on 125 MM patients and concluded that the data strongly support the implementation of MRD status as a primary endpoint and a surrogate outcome for MM clinical trials.

1.2.5 Timing of MRD assessment

Despite the IMWG criteria defining sustained MRD negativity (see Table 1.6) as two instances of confirmed MRD negativity over a year apart, the true number of required monitoring samples to define this has not been fully elucidated. Alongside this, the optimal time points for monitoring have also not been defined and prospective longitudinal studies will be required to clarify these. It has been found that patients who reached an MRD negative status both before and after ASCT had better survival than those who were only MRD negative at one of these time points, suggesting evaluation at both time points is needed for optimal prognostic utility (Schinke *et al.,* 2017). Further to this, a longitudinal study in transplant-eligible patients concluded that regular MRD monitoring is important and that MRD re-appearance always indicated disease progression (Gu *et al.,* 2018). By assessing different MRD progression patterns against PFS/OS this group suggests that it is the sustainment of an MRD negative status (over the 24 months of this study), rather than attainment of MRD negativity at one time point, which is most associated with optimal survival outcomes. They also suggest that the 1-year sustained follow up outlined by the IMWG criteria for "sustained MRD negativity" should be extended to 2 years (Gu *et al.,* 2018).

A small longitudinal study testing MRD by PCR techniques found that some disease progression was not always preceded by an MRD positive status (Ferrero *et al.,*2015). As there was a median of 8 years of follow up in this study it has the advantage of following the majority of patients until clinical relapse. However, the timings of sampling may contribute to the study's conclusion as the last MRD assessment occurred 2 years before disease progression. This highlights the role for a blood-based assay which is able to be performed frequently in MRD detection. Extending this analysis, the group interestingly found that patients who initially obtained MRD negative status and then converted to an MRD positive state had survival outcomes which were intermediary between the other two (MRD negative/MRD positive) groups. The MRD analysis in this study was performed using PCR-based methods; it would be interesting to follow patients up long-term using blood-based assays alongside the accepted MRD techniques. As MM is a rather indolent disease, it is the durability of response which is important and therefore testing at only early time points is not as relevant in this disease as with MRD testing in other haematological malignancies such as ALL (acute lymphoblastic leukaemia).

1.2.6 MRD in transplant ineligible patients

Some, usually older, patients are not eligible for ASCT which contributes to the disparity in OS seen between young and old MM patients; this is unsurprising since age, co-morbidities and patient fitness all play a role in treatment allocation and intensity (Rosko *et al.,*2017). However, a 2011 study examining survival data found that older patients (>70 years) were indeed beginning to see improved survival outcomes and that survival rates were accelerating over time (Pulte *et al.,*2011). The majority of MM patients are elderly and therefore a large proportion of patients are not suitable for an ASCT. It is in these patients, who likely have co-morbidities, where the equilibrium between effective treatment and treatment toxicities must be most carefully balanced. Within this group it is hoped that sensitive response assessment and monitoring can help to evade both under- and over-treatment.

The preponderance of clinical trials and research studies in determining MRD status in MM patients has been to focus on newly diagnosed patients who have undergone ASCT. There is a paucity of data related to transplant ineligible, often elderly, patients due to their exclusion from the majority of randomized trials and the fact that until recently they were infrequently achieving CR. However, a few groups have looked into the utility of MRD detection in transplant ineligible patients. There are 4 main studies where the utility of IMWG response assessment in transplant ineligible patients has been evaluated (Mateos *et al.,*2010, Morgan *et al.,*2010, Paiva *et al.,*2016, Facon *et al.,*2019). Two of these studies were evaluating different treatment regimens for these ineligible patients and found that, as expected, those patients achieving CR had significantly better survival outcomes than those achieving VGPR or below, regardless of treatment (Mateos *et al.,*2010, Morgan *et al.,*2010). Mateos *et al.,* (2010) also showed that immunophenotypic remission (using early MFC MRD techniques) had clear improved survival outcomes compared to those patients in CR only. Paiva *et al.,* (2016) revealed MFC MRD status to be one of the most significant independent prognostic factors in elderly transplant-ineligible patients. This study found that although there was a different in survival of those patients with MRD levels $>10^{-5}$ and $\leq 10^{-5}$, there was no survival difference in patients who were in CR and less than CR when MRD was present, showing the increased utility of MRD status over the original IMWG response criteria. This is supported by a more recent trial which showed patients achieving MRD negativity by NGF experiencing significantly longer PFS than those who remained MRD positive (Facon *et al.,*2019).

Other studies which have included elderly patients or non-intensive treatment arms show differing results in regards to the predictive value of MRD status in survival outcomes. The

PETHEMA/GEM2005 trial (Mateos *et al.* 2014) showed a significant increase in PFS and OS for MRD negative patients versus MRD positive patients, whereas the MRC Myeloma IX trial showed no significant difference in survival between the two MRD statuses (Rawstron *et al.,*2013). Therefore, although recent publications suggest MRD assessment can be used as a marker to evaluate the efficacy of different treatment strategies in transplant eligible patients, there has not been such in depth research or agreement with transplant-ineligible patients. The role of MRD is still an unknown phenomenon in this patient group, but the use of a blood-based test for MRD over a BM is undoubtedly beneficial in this cohort. With highly effective therapies now being available outside of transplant, it is important to try to calculate response driven treatment plans for elderly patients.

1.2.7 Conclusion on MRD

Studies have shown that MRD has clinical impact regardless of the treatment, patient risk factors, detection methods and sensitivity of technique. Currently MRD assessment is still considered a research tool and is not used for treatment decisions outside of clinical trials (Martin and Huff, 2019). The studies highlighted here make a compelling case for the use of MRD in MM response assessment but also highlight the current heterogeneity in MRD assessment including method of detection, timings of assessment and long-term utility. There is a need to evaluate MRD outside of clinical trials in an adaptable but sensitive way. The current reliance on bone marrow based MRD analysis makes this test unsuitable for regular monitoring of patients and alternative blood-based detection methods would have great value in this field.

Table 1.7: Important papers on MRD and their main findings.

Study	Technique (Sensitivity)	Patient Population	Key Messages
Paiva *et al.,*2008	MFC (10^{-4})	295 ASCT patients	- PFS longer in MRD negative patients at day 100 than MRD positive patients at day 100 post-ASCT. - MRD is the most important independent prognostic factor for PFS (HR=3.64, p=0.002) and OS (HR=2.02, p=0.02).
Rawstron *et al.,*2013	MFC (10^{-4})	>350 ASCT patients >200 non-ASCT patients	- MRD status significant predictor of PFS both post-induction and post-ASCT for transplant patients.

			- This prediction was not affected by adverse cytogenetics. - MRD status not predictive of PFS post-induction for non-transplant patients.
Rawstron *et al.,*2015	MFC (10^{-4})	397 ASCT patients	- The level of MRD is highly informative, such that an approximate 1-year OS benefit is demonstrable for each log of tumour depletion.
Paiva *et al.,*2017	NGF (10^{-6})	458 ASCT patients	- MRD status had a greater effect on PFS and OS than cytogenetic risk. - Cellular and imaging MRD techniques are both needed in order to detect extramedullary disease.
Lahuerta *et al.,*2017	MFC (10^{-4} – 10^{-5})	482 ASCT patients 127 non-ASCT patients	- MRD negative patients showed a marked increase in PFS/OS compared to MRD positive in both groups. - Survival benefit previously attributed to CR found to be due to MRD negativity
Flores-Montero *et al.,*2017	NGF (10^{-6})	332 patients	- Enhanced sensitivity gives MRD status more prognostic utility.
Perrot *et al.,*2018	NGS (10^{-6})	~200 ASCT patients	- Different levels of MRD-NGS cut off could predict different PFS and OS times/outcomes.
Carulli *et al.,*2019	MFC (10^{-5})	46 patients	- 8-colour dried reagent tube is useful for the routine monitoring of treated MM patients. - MRD evaluation is most useful in those achieving CR or sCR.
Zamagni *et al.,*2019	MFC (10^{-5}) and PET/CT	182 patients	- Strong correlation between MRD results by PET/CT and MFC in the bone marrow. - PET/CT could be complementary to BM MRD assessment to detected extramedullary disease.

1.3 Multi-parameter flow cytometry (MFC) for MRD evaluation

1.3.1 Current state of MFC in MRD evaluation

In the healthcare service it is important to determine the most effective analytical methods, not only in terms of sensitivity and specificity, but also applicability, repeatability and costs – in time, skill and reagents. Being a national service, it is also important to have a method which can be standardised across the country. The IMWG consensus criteria do not specify which method is best to use and instead recommend using the methods which are locally available (Kumar *et al.,* 2016). From the current methods used to determine MRD, MFC, and potentially the more sensitive NGF, is the best suited to implement and use as the gold standard method.

There are a number of reasons why MFC has been the most widely studied technology for use in MRD detection, which apply both in the NHS and worldwide. Flow cytometers are readily available in the majority of diagnostic laboratories, panels for MRD detection can be standardised and software can be produced to homogenise analysis (Flores-Montero *et al.,* 2017). In the NHS, access to flow cytometers is more widespread than access to NGS techniques and MFC is currently routinely used in the diagnosis of MM and other haematological malignancies and in MRD detection in malignancies such as acute leukaemias (Chatterjee *et al.,* 2016). MFC involves the use of fluorochrome dye-labeled monoclonal antibodies (mAb) directed against certain cell surface or intracellular proteins. The flow cytometer uses light scattering parameters and the differing fluorescent intensities of the different fluorochromes to identify cells. By looking for different cell markers using specific patterns of fluorochromes, cells can be identified and discerned from each other. Both normal and abnormal cells are identified by specific combinations of surface and/or intracellular antigens. MFC can analyse a large number of events from one sample and provide information on a cell-by-cell basis. A larger number of cells are required for MRD detection than bone marrow immunophenotyping at diagnosis due to the abnormal plasma cells being present in very low numbers. The more events which are acquired, the more sensitive the assay is and as previously stated, increased sensitivity has a better predictive value for PFS (Paiva *et al.,* 2016).

1.3.2 Disadvantages of MFC

A survey of MRD assessment practices with responses from 11 major testing institutions in the US highlighted the heterogeneity in MFC MRD testing for MM patients. Variances were especially seen in total number of cells analysed, definitions of the presence of MRD and the panel of antibodies used to define abnormal clones of plasma cells (Flanders *et al.,* 2013). However, consensus guidelines have since been published and heterogeneity has started to

decrease (Salem *et al.,* 2016). Despite the EuroFlow recommendations, the number of tubes to be used, the panel of antibodies, and the number of events to be acquired are all variables which have not been fully agreed internationally. Even with the use of standardised assays and panels, each laboratory will need to fully validate these for their own use which requires time, money and resources which routine NHS laboratories do not often have. The centralisation of analysis could help to solve standardisation problems; however bone marrow samples need to be processed within 48 hours maximum, and ideally within 24 hours of being taken. One study found the absolute number of plasma cells declined by up to 40% in samples stored for 48 hours versus those tested immediately after collection (Royston *et al.,* 2016). Therefore transport to central hubs may not be possible during this time frame.

The majority of laboratories running haematological flow cytometry assays will not have experience with complex panels and the machines and software are not equipped to handle MRD protocols. To achieve the minimum required 10^{-5} sensitivity for an MRD assay 5 million cells need to be analysed per patient (Kumar *et al.,* 2016). This is a lengthy procedure and could therefore prove to be a significant disruption to the workflow in a routine laboratory where samples need to be analysed within a specified time frame. An important disadvantage of MFC methodologies is the knowledge and expertise required to analyse and interpret the results. Automated identification and reference databases have been created to solve this issue (Flores-Montero *et al.,* 2017); however, these are difficult to implement across the NHS due to the high cost of the software and large data storage requirements versus the relatively low sample input. The detection of rare events is complicated by the need to identify very small numbers of specific cells against a background of many other cells and cell types. In the health service this is further confounded by the need to perform these tests on limited sample volumes.

The most critical aspect affecting the outcome of MRD testing by MFC is sample quality. Bone marrow samples are considered precious samples which cannot be easily retaken. There must be a large enough sample of bone marrow taken to give the appropriate number of cells for maximum sensitivity. The sample must also not be overly haemodiluted ('contaminated' with peripheral blood during the collection procedure) as this could lead to false negative MRD results. Internal quality checks must be performed alongside the MRD evaluation in order for the results to be of clinical value, including a quality check of the whole sample cellularity by identifying cell populations such as B cell precursors, eosinophils, monocytes and erythroblasts. Sampling errors not only lead to haemodilution but also to unusable clotted samples being

received. In a pilot study performed prior to this study, 12% of bone marrow samples received were not usable (Campbell *et al.,* 2019).

Bone marrow involvement in MM is patchy and multifocal, still, bone marrow aspirates are often taken from one anatomical site, the pelvic bone, as bone marrow is easily accessible at this site. Therefore, with each sample taken, there is a chance that the bone marrow being taken is normal whilst large amounts of abnormal bone marrow exists in the patient. MM can also exist as extra-medullary disease which will not be detected in a bone marrow sample. The issues related to bone marrow sampling means that although MRD positive results are very informative, MRD negative results may always have a chance of being false negatives. Due to this possibility of false negatives, it is currently much safer to make clinical decisions based on an MRD positive results rather than an MRD negative.

1.3.3 Bone marrow MFC conclusion

As MM is a bone marrow based disease, the majority of MRD studies have focused on bone marrow based assays. However, to effectively monitor patients this would require sequential invasive procedures at frequent or semi-frequent intervals. The procedure of bone marrow aspiration is painful and inconvenient for patients, has associated health risks and is therefore not suitable for monitoring purposes. Repeated sampling has the potential to have a detrimental effect on patients both physically and psychologically. In a prospective study of over 200 haematological malignancy patients, 70% reported pain during the procedure, with 32% reporting severe pain (Linden *et al.,* 2009). From our own experiences with local MM patient groups the bone marrow aspiration procedure is an unpleasant and often feared experience and patients are keen to take part in research which might minimise the frequency of this procedure for current and future patients. In a survey of MM clinicians, 25% stated that the discomfort of the bone marrow aspiration procedure was a concern about using MRD status to guide decision making (Derman *et al.,* 2022). This same survey also showed that lack of access to MRD tests and high costs were common reasons for MRD not being used clinically. As bone marrow samples are invasive, it is important to investigate whether blood-based techniques will be more or equally informative at the same disease time points as the gold standard BM MRD evaluation.

1.4 Blood-based assays for MRD evaluation

1.4.1 The need for a blood-based MRD method

Despite the great advances in MRD analysis, the advantage of frequent sensitive MRD analysis would be greatly outweighed by the significant detrimental effect of the quality of life for these

patients. Sensitive, non-invasive, blood-based assays which can detect the presence of MRD are needed to optimize the management of MM patients. Peripheral blood tests are appealing due to the ease and rapidity of sample collection and testing, the opportunity for more frequent sampling and longitudinal monitoring of disease burden, ease of standardisation across laboratories, and reduced costs at all stages of the testing procedure. A test which can be implemented in the majority of local laboratories is also preferable due to the cell loss which occurs during transport delay of bone marrow samples (Royston *et al.,* 2016). The use of blood-based assays can be seen as particularly useful in elderly patients who may find bone marrow aspirations particularly traumatic but who would benefit greatly from close monitoring in order to balance treatment efficacy with treatment-related toxicity. It is also hoped that blood-based assays can detect loss of response in treated patients before the signs of clinical relapse. For the aforementioned reasons, the IMWG recommends the development of a blood-based MRD monitoring tool as the ultimate goal in MRD testing (Kumar *et al.,* 2016).

Assessing MRD in blood is an intimidating concept, due to the need to detect very low levels of M-protein or abnormal plasma cells against a polyclonal background. In principle, the sensitive detection of M-protein is the most favoured peripheral blood MRD detection method, this is because the M-protein is already routinely measured, and a sensitive test can be easily added to the patient pathway and implemented into laboratories. The detection of the M-protein has advantages in this extremely heterogeneous disease. A single patient may have multiple clones of malignant plasma cells present at one time (Melchor *et al.,* 2014); these clones may display different surface markers and therefore recognition of these clones by flow cytometry can prove complicated and inaccurate. The M-protein is representative of the product of all clones and could therefore be a more reliable test. M-protein detection is also not affected by the presence of patchy bone marrow involvement or extramedullary disease. Traditional techniques to detect the M-protein, such as SPEP and UPEP, which used to be sufficient to define response to treatment in 90% of cases (Chee *et al.,* 2009), are becoming incapable of monitoring patients achieving deeper and deeper responses due to their limited sensitivity (Willrich *et al.,* 2016).

1.4.2 Heavy/light chain (HLC) assay

The Hevylite™ assay (The Binding Site, UK) evaluates the heavy/light chain (HLC) pairing of the M-protein. For patients with intact immunoglobulin MM this provides the ability to assess the involved (e.g. IgGλ in a patient with an IgGλ M-protein) and uninvolved (e.g. IgGκ in an IgGλ M-protein patient) immunoglobulin separately for the first time, providing more specific information than the currently used SPEP (Koulieris *et al.,* 2012). The manufacturers HLC normal

reference ranges are used (Katzmann *et al.,* 2013) and the acronyms for the different HLC measurements which can be performed are in Table 1.8 and Figure 1.3 using an IgGλ M-protein patient as an example. Currently an abnormal involved HLC (iHLC) and/or abnormal HLC ratio (HLCr) is considered an abnormal HLC result. An abnormal iHLC is considered to be above the normal range, an abnormal uninvolved HLC (uHLC) is considered to be below the normal range.

Table 1.8: Definitions of HLC measurements and calculations

Term	Definition	Comment
iHLC	Involved HLC	The HLC isotype that is produced by the tumour e.g. IgGλ in an IgGλ producing tumour.
uHLC	Uninvolved HLC	The HLC isotype with the same heavy chain but alternate light chain to that produced by the tumour e.g. IgGκ in an IgGλ producing tumour.
HLCr	HLC Ratio	The ratio of the concentration of HLC κ to λ for a particular immunoglobulin isotype e.g. IgGκ/IgGλ. This ratio helps to indicate monoclonality.
dHLC	Difference in HLC	The difference between the iHLC and the uHLC for a particular immunoglobulin isotype e.g. IgGλ-IgGκ in an IgGλ producing tumour.
sHLC	Summated HLC	The total sum of iHLC + uHLC e.g. IgGλ+IgGκ to give a total HLC isotype result comparable to total immunoglobulin.
HLC matched pair suppression	When uHLC concentration is below the normal range	The HLCr must also be abnormal for HLC pair suppression to be concluded.
I/U HLCr	Involved/uninvolved HLC ratio	Quantitative marker with the involved HLC as the numerator and uninvolved HLC of the same isotype as the denominator. Allows grouping and comparison of e.g. IgGκ and IgGλ isotype patients together. No normal ranges exist for I/U HLCr.

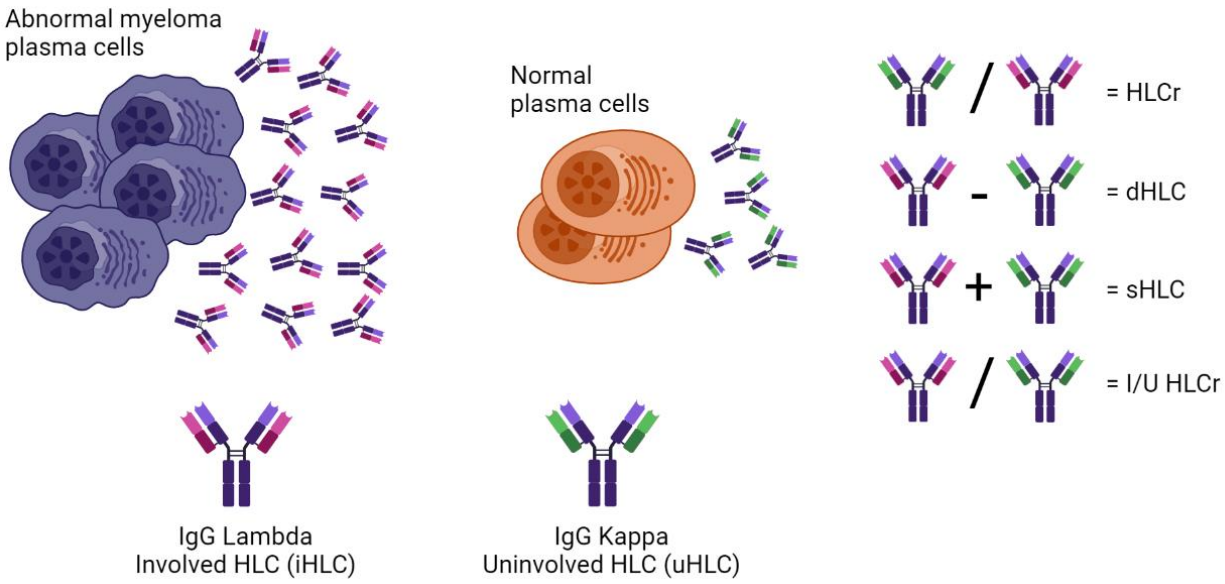


Figure 1.3: Definitions of HLC measurements in an IgG lambda M-protein patient
Highlighting the use of the definitions from Table 1.8 in an IgG Lambda M-protein patient.

In a similar way to the FLC assay, the HLC assay allows definition of clonality and evaluation of individual tumour clones which the conventional SPEP/IFE techniques do not. The assessment of the uninvolved immunoglobulin allows assessment of immunosuppression in the bone marrow as a decrease in the uninvolved immunoglobulin will likely be due to the suppressive effect of the malignant clone on other normal plasma cells. The limit of detection of the HLC assay is similar to or better than the sensitive IFE assay (Rios-Tamayo *et al.,* 2021).

1.4.3 Current HLC usage in multiple myeloma

Although the HLC assay was first introduced in 2009 it is yet to become extensively used in either the diagnosis or monitoring of MM (Bradwell *et al.,* 2009). In a 2017 survey sent to 71 MM testing centres from around the world 6/40 (15%) respondents used HLC to assess treated MM patients (Holstein *et al.,* 2018). None of these respondents were from UK based sites and therefore this survey does not represent the current state of HLC measurement in the NHS. Although the HLC assay is still not part of the routine diagnostic work up of MM patients it is being used as a monitoring tool by some and has been included in a few clinical trials (Greil *et al.,* 2017).

The HLC assay has initially proved most useful in M-protein detection and quantification of those M-proteins which migrate into the β-region during SPEP (Boyle *et al.,* 2014). It is difficult to quantify β-region migrating proteins as other proteins (e.g. C3, transferrin) also migrate to this

region and may mask the protein and/or affect its quantification. Because of this, HLC utility has been most thoroughly investigated in IgA MM patients, who have the highest frequency of β-migrating M-proteins.

1.4.4 HLC in diagnosis

A large study investigating the prognostic and clinical utility of FLC and HLC in MM found that highly abnormal HLC ratios (<0.29/>73) at diagnosis were associated with shorter PFS (Lopez-Anglada *et al.,* 2018). In addition to this, iHLC levels of >5g/L (for IgA or IgG patients) were associated with both shorter PFS and OS. This study suggested that the HLCr has prognostic utility at diagnosis, whereas the iHLC level is prognostic after treatment. It is acknowledged that at diagnosis immunoparesis of the uninvolved isotypes (e.g. IgA and IgM in an IgG M-protein patient) is independently associated with a poorer outcome (Kastritis *et al.,* 2014), although this has not always been shown to be true for all patient groups (Ludwig *et al.,* 2016). Boyle *et al.,* (2014) found that 80% of IgA M-protein MM patients in their study (126/157) presented with isotype paired suppression at diagnosis (e.g. suppression of IgAλ in an IgAκ M-protein producing patient) but this was not found to be prognostic for PFS or OS.

1.4.5 HLC in monitoring and MRD

The M-protein is important to detect and quantify not only for diagnosis but also for monitoring for the duration of the disease. During and after treatment, quantifying the M-protein may get harder as the plasma cell clone responds to treatment. Abnormal HLC results have been shown months before SPEP detection in ASCT patients who have gone on to relapse and the prognostic utility of HLC analysis has been demonstrated where abnormal results have been associated with reduced PFS (Ludwig *et al.,* 2013). Harutyunyan *et al.* (2016) determined the HLC levels of involved and uninvolved immunoglobulins and their relationship to outcomes in 189 MM patients. They showed that patients with normal uHLC levels showed a much longer PFS than those with below normal uHLC levels (p=0.0019). This was also the case when comparing those with normal iHLC levels and above normal iHLC levels (p=0.041). In assessing the clinical relevance of the HLC assay in treated MM patients Batinic and colleagues produced Kaplan-Meier curves stratifying HLCr values against overall survival (OS) (Batinic *et al.,* 2015). Patients with abnormal HLCr values at any clinical stage had significantly shorter OS than those who did not. This study found HLCr to be abnormal in the presence of a normal free light chain ratio (FLCr) in patients who were in CR, VGPR and SD response states, showing that an algorithm using both tests is the most sensitive. The prognostic significance of the HLC assay was further tested using Cox multivariate regression analysis and it was found that both an abnormal HLCr and a

high beta-2-microglobulin concentration were independent predictor of OS. However, when only those patients achieving VGPR or CR were included in the analysis the significance of an abnormal HLCr was not as strong. Ludwig *et al.* (2013) reported that an abnormal HLCr could be detected in approximately 25% of patients with CR. This group further showed that HLC-pair suppression at the best response time point was correlated with poorer survival in both newly diagnosed and relapsed/refractory MM (Ludwig *et al.,* 2016). These results suggest the promising prognostic utility of the HLC assay in MM patients.

Our group has previously carried out a pilot study comparing the sensitivity of FLCr/HLCr normalisation together with MFC MRD negativity in predicting survival outcomes (Campbell *et al.,* 2018). From 23 patients post-ASCT normalisation of FLCr/HLCr results showed good diagnostic performance compared to MFC, if MFC is considered the gold standard. However, a significant prediction for PFS was not found using combined FLCr/HLCr normalisation alone (p=0.27). This study demonstrated the potential of using FLC/HLC as surrogate markers for BM MRD negativity. Previous studies examining the correlation of FLC alone with MRD found that a normal FLC did not always associate with MRD negativity (Matinez-Lopez *et al.,* 2015). However, we demonstrated that combined FLC and HLC normalisation at day 100 post-ASCT broadly correlated with a better outcome in terms of PFS (Campbell *et al.* 2018).

Kraj *et al.* (2012) compared the use of IFE, HLC and FLC in the assessment of remission in ASCT MM patients. They found IFE to be more sensitive than the HLC and FLC assays in the detection of residual disease. However, they also found that patients classed as CR could be further split depending on HLCr and FLCr values, with those whose ratios were normalised having longer PFS than those who had abnormal ratios. Beaumont-Epinette *et al.* (2015) also assessed the correlation between HLC and IFE for the detection of residual disease after treatment and in the early detection of relapse. Confirming earlier studies, this group also found the HLC assay to be less sensitive than IFE in detecting residual disease. In a subset of patients however (3/15), the HLCr values became abnormal prior to SPEP and IFE and the patients subsequently relapsed. Despite this, the group concluded that the HLC assay provided no additional information to current techniques for monitoring intact Ig MM patients.

Miyazaki and Suzuki (2018) aimed to evaluate the sensitivity of the HLC assay, particularly in patients with deep responses equivalent to CR/sCR. This study compared the utility and sensitivity of the HLC assay to the currently used conventional methods and found addition of HLC testing significantly increased the sensitivity of blood-based disease detection (p=0.02 for

HLCr, p=0.004 for HLC-matched pair suppression). This study did not do any follow up on these patients so the relevance between the HLC assay results and patient survival could not be determined. A large study of >400 NDMM patients compared HLC response assignment (normal/abnormal) to SPEP (Michallet *et al.,* 2018). Patients ranged in IMWG response assignment from PR to CR, with interestingly the VGPR group having poor concordance between methods (45% agreement) compared to PR (79%) and CR (92%). This study advocated for the additional value of HLC analysis in both response assessment and prognosis. Importantly, this paper compared HLC results with MFC analysis although the sensitivity of this MFC analysis was 10^{-4}, which does not meet the current standard for MRD analysis. However, at this sensitivity the concordance between the two methods was high and this paper concluded that HLC analysis indicated not only the presence of residual tumour cells but also the normalisation of the immune system by evaluating the polyclonal production of matched uninvolved immunoglobulin (uHLC). For this study, however, comparisons between techniques occurred at only one time point during the patient's treatment. A single centre prospective study evaluated 25 MM patients in CR with a median follow up of 52 months. HLCr and FLCr were analysed for all patients either post-induction (for non-ASCT) or post-ASCT (D'Auria *et al.,* 2017). This study found FLCr normalisation in CR patients to be significantly associated with better PFS (p=0.049) – thus confirming the need for a further response level past CR as outlined in the 2016 IMWG response criteria (Kumar *et al.,* 2016). However, HLCr was not found to have an impact on PFS. This small study concluded that the use of HLC in monitoring, and especially MRD monitoring, is doubtful, however only a small number of CR patients were evaluated and larger studies are needed to draw absolute conclusions. Table 1.9 contains a summary of other important studies looking into the use of HLC in MRD detection.

1.4.6 HLC conclusions

The IMWG state that bone marrow based molecular MRD techniques are 3 orders of magnitude more sensitive than serum M-protein detection by IFE and FLC (Kumar *et al.,*2016). However, no comparison has been made for the newer HLC assay. Both the recent IMWG response and MRD assessment criteria and the NICE MM guidelines recommend further investigation into the potential role for HLC in deep response evaluation, particularly in those patients achieving CR post-treatment (Kumar *et al.,* 2016, NICE 2016). A normalisation of the HLC result may reflect the reconstitution of healthy non-malignant plasma cells in the BM. Unlike MFC MRD detection, the HLC assay could indicate the functional result of MRD negativity, in addition to quantifying residual disease. More data are needed, particularly amongst patients achieving CR, to allow conclusions to be drawn on the utility of the HLC assay and its relationship to the FLC assay which

is already in the IMWG criteria (Durie *et al.,* 2006). HLC analysis in MM can therefore greatly improve patient management for several reasons including acting as an early indicator of relapse, being a prognostic marker, showing response to treatment and refining of response criteria. The ease of use, standardisation, interpretation, availability of technology and simplicity of obtaining the sample from the patient makes this a key area for research within the NHS.

Table 1.9: A summary of studies investigating HLC measurement in MM

Reference	Patient Group	Key Points
Ludwig *et al.,*2013	156 NDMM patients	- Abnormal HLCr at diagnosis and post-treatment is associated with shorter OS. - The post-treatment association was not present in VGPR patients.
Batinic *et al.,*2015	90 treated MM patients	- Abnormal HLCr is associated with significantly shorter OS. - Paired isotype suppression of the uHLC also associated with significantly shorter OS.
Ludwig *et al.,*2016	203 MM patients 156 NDMM 47 relapsed/refractory	- HLC-matched pair suppression is significantly and independently associated with reduced survival.
Garcia de Veas Silva *et al.,*2016	85 patients 30 ASCT	- Severe HLC-matched pair suppression is associated with shorter PFS and OS. - Abnormal HLCr post-transplant is associated with shorter PFS.
D'Souza *et al.,*2017	497 MM patients 331 ASCT 166 Allo-SCT	- HLCr measured pre-ASCT is prognostic for PFS and OS in MM patients.
Suehara *et al.,*2017	120 MM patients	- Abnormal HLC associated with significantly shorter OS in IgA patients reaching at least VGPR.
D'Auria *et al.,*2017	25 MM patients 11 ASCT 14 non-ASCT	- HLC assessment had no significant impact on PFS or OS.
Michallet *et al.,*2018	509 NDMM patients	- HLC monitoring adds prognostic significance to M-protein evaluation in treated MM patients.
Lopez-Anglada *et al.,*2018	>800 ASCT and non-ASCT MM patients	- Highly abnormal HLCr was prognostic at diagnosis. iHLC levels were prognostic after treatment.

1.4.7 Mass spectrometry (MS)

Mass spectrometry (MS) based platforms are currently used in clinical laboratories for a number of protein biomarkers (Jannetto *et al.,* 2016). Recently, MS has been proposed as a novel sensitive screening tool for monoclonal gammopathies, being able to replace both SPEP and IFE

in the detection, isotyping and quantitation of the M-protein. MS methods are capable of detecting very low levels of M-protein in the blood, with detection limits approximately 100 times lower than that of IFE (Barnidge *et al.,* 2014).

Two MS techniques have emerged as potential methods for M-protein analysis: the clonotypic peptide approach and the intact immunoglobulin light chain (intact Ig LC) approach. The intact Ig LC approach is a simpler method and has been adapted to matrix-assisted laser desorption/ionisation-time-of-flight (MALDI-TOF) MS. This development introduced a high throughput, fast and sensitive method which has been clinically validated at the Mayo Clinic and has currently been found to be comparable to IFE by the Mayo Clinic laboratory (Kohlhagen *et al.,* 2020). A pre-commercial intact Ig LC assay, quantitative immunoprecipitation mass spectrometry (QIP-MS), is based on the principle that each M-protein has a unique amino sequence and therefore a unique mass (The Binding Site, UK). This assay can identify, type and, recently, quantify both intact and light chain only M-proteins. This technique originally used denatured intact immunoglobulin light chains to produce two polyclonal mass spectra, one for free kappa and one for free lambda. These mass spectra were normally distributed and any monoclonal light chain could be detected above this background (Barnidge *et al.,* 2014). This technique will be referred to in Chapter 6 as FKFL MS. This technique has been refined further to include an immune-enrichment step which allows identification of the intact immunoglobulin heavy chain also and is referred to as GAMKL MS (Mills *et al.,* 2016).

1.4.8 Mass spectrometry in multiple myeloma

Previous studies have shown that MALDI-TOF-MS demonstrates good analytical performance (Milani *et al.,*2017, Thoren *et al.,* 2018, Kohlhagen *et al.,*2020). In a preliminary small study in collaboration with The Binding Site, our group tested the ability of QIP-MS to detect low level M-proteins in 11 monoclonal gammopathy screening samples. In this small study, QIP-MS positively identified the primary M-protein seen on IFE in all 11 patient samples, including against a polyclonal hypergammaglobulinaemia background (Campbell *et al.,* 2021). In addition to this, extra M-proteins of differing isotypes were detected in 5 patient samples. This suggests that QIP-MS has detection levels below that of IFE. This increased sensitivity compared to the current standard tests is supported by data from other studies (Barnidge *et al.,* 2014, Mills *et al.,* 2016). In Campbell *et al.* (2021) we concluded that QIP-MS has clinical utility as a first-line screening tool for monoclonal gammopathy investigation, identifying monoclonality in patients with higher sensitivity and resolution compared to the current standard methods.

1.4.9 Mass spectrometry in MRD

The most recent IMWG report recommends the further collection of data for the use of mass spectrometry for MRD detection in peripheral blood and its use in guiding the timings for bone marrow tests (Murray *et al.,* 2021). It is hoped that blood-based assays can detect loss of response in treated patients before the signs of clinical relapse. Although mass spectrometry has been found to be more sensitive than IFE its ability to detect minimal levels of M-protein against a background of polyclonal immunoglobulins is still in question, potentially affecting its limit of detection (LOD) (Bergen *et al.,* 2016, Mills *et al.,* 2017, Thoren *et al.,* 2018).

Eveillard *et al.* (2020) compared the performance of MALDTI-TOF-MS of peripheral blood with sensitive NGF of the bone marrow for the detection of MRD. They concluded that MS adds value to the bone marrow testing and is more sensitive for than SPEP/IFE for detecting relapse in treated MM patients. As well as sensitivity, specificity of M-protein detection can also be improved by MS as it has the ability to distinguish between the M-protein and therapeutic monoclonal antibodies more easily than current techniques (Moore *et al.,* 2019). A small study of the QIP-MS system confirmed the higher sensitivity of QIP-MS when compared to SPEP/IFE (Puig *et al.,* 2019). For those patients in CR, QIP-MS showed good correlation with bone marrow MFC MRD at post-induction and post-ASCT time points and overall a moderate correlation with MFC MRD across all time points analysed (Puig *et al.,* 2019).

1.4.10 Mass spectrometry conclusion

MS is currently being investigated for its clinical utility at different points in the monoclonal gammopathy pathway, from screening to MRD detection (Murray *et al.,* 2019, Eveillard *et al.,* 2020). MS needs to be implemented into larger cohorts and the overall concordance with current techniques, both blood and bone marrow based, needs to be asserted at different stages of the monoclonal gammopathy pathway but especially in MRD scenarios. The ability of MS to screen potential monoclonal gammopathy patients, track M-proteins in current patients, detect M-proteins in MRD settings and identify M-proteins in those treated with monoclonal antibody therapy must be explored. Further MS studies are summarised in Table 1.10.

Table 1.10: A summary of studies investigating MALDI-TOF MS measurement of M-protein in MM

Reference	Patient Group	Key Points
Milani *et al.,*2017	257 MM patients	- MS sensitivity is comparable to IFE/FLC across a spectrum on monoclonal gammopathy patients.
Moore *et al.,*2019	31 daratumumab treated patients	- MS offered greater specificity for distinguishing daratumumab from M-proteins compared to IFE.
Eveillard *et al.,*2020	48 NDMM 23 treated MM	- MS adds value to BM MFC MRD measurements. - MS is more sensitive than current peripheral blood methods for detecting relapse.
Kohlhagen *et al.,*2020	182 MM	- Successful implementation of the MS system to replace IFE.
Derman *et al.,*2021	76 treated VGPR or better	- MS may be as sensitive as BM-based MRD evaluation.
Eveillard *et al.,*2021	23 NDMM treated with daratumumab	- MS more sensitive than standard panel of conventional methods. - MS can distinguish M-proteins from daratumumab.

1.4.11 Circulating tumour plasma cells (CTPC)

In the search for a non-invasive test to replace bone marrow sampling another area of interest is the detection of circulating plasma cells (CTPCs) in peripheral blood. CTPCs are abnormal MM plasma cells which have originated in the bone marrow and escaped into peripheral blood. The reasons these cells are released from the bone marrow have not been fully elucidated and could represent both medullary and extramedullary disease (Kis *et al.,* 2017, Mishima *et al.,* 2017). CTPCs have been detected in all stages of MM development and detection of these CTPC has been shown to be a poor prognostic factor in MGUS, SMM and NDMM and RRMM (Peceliunas *et al.,* 2012, Gonsalves *et al.,* 2014a, Terpos *et al.,*2019). The characteristics of these cells strongly resemble those in the bone marrow, although downregulation of some cell surface markers has been shown (Paiva *et al.,* 2013, Terpos *et al.,* 2019).

CTPCs can represent an important alternative way to detect the presence of disease, not only due to using a non-invasive technique, but also as they are quiescent and therefore their presence may be overlooked by detecting M-protein alone (Kumar *et al.,* 2004, Paiva *et al.,* 2013). The two main forms of detection being investigated are MFC/NGF (Paiva *et al.,* 2013) and nucleic acid-based (Mishima *et al.,* 2017). Due to the technology available and the implementation of a plasma cell panel for bone marrow plasma cell detection by MFC/NGF as part of this study, this technique will also be explored in CTPC detection. Up until recently the

relatively low sensitivity of MFC has limited detection of plasma cells to the bone marrow, as far lower numbers of CTPCs are found in the periphery than in the bone marrow. It had previously been thought that the frequency of CTPCs is 1 in 1million nucleated blood cells (Foulk *et al.,* 2018). However, several studies have been able to show a higher presence of CTPCs, with post-induction patients in one study having a median of 58 CTPCS per 150,000 events (Chakraborty *et al.,* 2016). With the introduction of the sensitive NGF method, the exploration of CTPCs has now become a viable area of study.

The Mayo Clinic group has found CTPCs to have prognostic utility in both NDMM and RRMM (Gonsalves *et al.,* 2014a, 2014b). A study evaluating the prognostic utility of CTPCs when measured by MFC found the presence of CTPCs pre-transplant to have high prognostic utility in ASCT patients; with patients having deeper responses if no CTPCs were detected (Chakraborty *et al.,* 2016). However, CTPCs were only detected in 19% of patients so the applicability of this assay may be low. A more recent study using NGF technology detected CTPCs in 26% of treated MM patients, showing a higher detection rate using this more sensitive method and the collection of a larger number of cells (Sanoja-Flores *et al.,* 2019). However, this study did find that there were a significant number of patients who were BM MRD and IFE positive with no detectable CTPCs in paired samples. These findings indicate that CTPC is a less sensitive marker for MRD than BM MRD detection. This study concluded that although CTPCs may not be effective surrogate markers for BM MRD, the presence of CTPCs is an important prognostic marker for PFS and impending disease progression.

1.4.12 CTPC conclusion

There are a number of questions which can be asked of CTPCs including: does the presence of these cells in the periphery pre- or post-ASCT have prognostic utility for outcomes? Do the quantity and phenotype of these cells impact on outcomes? Can these cells be detected even in those patients with only minimal residual disease present? See Table 1.11 for a summary of the key studies into CTPCs in MM.

Table 1.11: A summary of studies investigating CTPC measurement in MM

Reference	Method	Patient Group	Key Points
Gonsalves *et al.,*2014	Six colour MFC, 150,000 events	647 previously treated MM patients	- Presence of CTPCs predicted worse median survival in treated patients. - Presence of >100 CTPCs in relapsing patients predicted worse survival.
Chakraborty *et al.,*2016	6 colour MFC, 150,000 events	840 pre-ASCT patients	- Presence of CTPCs at diagnosis and pre-ASCT associated with inferior PFS and OS (both p<0.001) - Presence of CTPCs an independent predictor of mortality
Chakraborty *et al.,*2017	6 colour MFC, 150,000 events collected	247 NDMM patients	- Monitoring CTPCs before induction therapy and pre-ASCT is predictive of survival in NDMM.
Sanoja-Flores *et al.,*2019	8 colour NGF, 10 million events	137 NDMM patients	- Presence of CTPCs is a powerful prognostic marker for PFS, regardless of treatment or treatment phase.
Terpos *et al.,*2019	8 colour NGF, 6-10million events	182 NDMM patients	- CTPCs detected in 86.8% of MM diagnostic samples - Trend of inferior survival in patients with high numbers of CTPCs over 16 months - Not significant but short follow up

1.4.13 Blood-based MRD detection summary

There remain significant unanswered questions with regards to the use of peripheral blood markers in MRD detection. Similarly, to BM MRD detection, sensitivity limits and positivity thresholds need to be defined, and the optimal timing of MRD measurement must be elucidated. However, with the ease of sample collection and rapidity of peripheral blood tests there is more flexibility in this assessment. Previous studies have indicated a potential for these assays but a lack in sensitivity compared to the bone marrow MRD tests is a common theme.

These tests will be investigated separately, but analysis must include the possibility of creating an algorithm of peripheral blood testing to try to delay or avoid taking a bone marrow sample from patients. In the advent of new and improved technologies and with the recent recognised importance of decentralised care for the immunocompromised, peripheral blood testing must be looked at as a viable alternative to bone marrow testing.

1.5 Aim and Objectives

The intended aim of this research is to obtain evidence which determines whether blood-based assays can act as a safe and effective marker for disease in the bone marrow of treated multiple myeloma patients. In order to do this, 5 objectives have been selected:

1. Validate and standardise a next-generation flow cytometric (NGF) assay for the measurement of low levels of bone marrow and circulating plasma cells.
2. Evaluate whether the NGF assay had prognostic utility when assessing MRD status in the bone marrow.
3. Evaluate whether peripheral blood NGF CTPC analysis and serum based HLC and MS analysis can be used to determine MRD status when using bone marrow NGF as the gold standard.
4. Assess the utility of blood-based assays during the follow up of treated patients in comparison with current conventional techniques.
5. Evaluate the data in order to determine whether the blood-based assays could be implemented into current routine practice in order to improve patient care.

Chapter 2 - Methods

This chapter describes the methods and materials used in this work. Additional method information is detailed in the relevant chapters.

2.1 Research setting

This was a longitudinal observational cohort study which recruited participants within England. Participant recruitment occurred at the following sites: Oxford University Hospitals NHS Foundation Trust (OUH NHS FT), Royal Berkshire NHS FT, Stockport NHS FT, Western Sussex NHS FT, Surrey and Sussex Healthcare NHS Trust, Frimley Health NHS FT, Buckingham Healthcare NHS Trust and the Clatterbridge Cancer Centre NHS FT. All sample processing, analysis and storage was carried out at the Oxford Immunology laboratory. Participating sites entered onto the study at different times but total recruitment time was from April 2019 – June 2021. Due to COVID-19 restrictions, recruitment for the study was paused between March 2020 – September 2020 and January 2021 – May 2021.

2.2 Ethical Approval

Ethical approval was granted by REC and HRA (REC reference: 19/NE/0025, IRAS ID: 252000) in January 2019 and was approved by the Sponsor, Oxford University Hospitals NHS FT. The ethical approval covers all recruitment sites under local arrangements. The study was conducted in accordance with the principles of the Declaration of Helsinki and Good Clinical Practice. The study was also carried out in line with the Data Protection Act 2018. Informed consent was obtained from each participant prior to entering the study.

2.3 Participant recruitment

Potential participants were identified and approached in haematology clinics by their treating clinician, clinical nurse specialist, clinical care team or myeloma service coordinator. The member of staff provided the potential participant with a participant information leaflet and gave a verbal explanation of the study. Participants who expressed an interest were given the opportunity to receive more information on the study. Informed consent was then obtained prior to initiating any study related activity. For the majority of participants written informed consent was obtained, however due to restrictions during COVID-19 some participants gave verbal consent over the telephone under a special amendment.

There were no exclusion criteria based on demographics, ethnicity or other personal criteria. Participants were excluded if they were unwilling or unable to give informed consent. Inclusion criteria required the patients to have a diagnosis of MM, to be attending haematology clinics at one of the recruitment sites and to be able and willing to give informed consent.

Participants were recruited into 2 different groups depending on the stage of their disease and/or their treatment plan. Group 1 are diagnosed MM patients due to undergo ASCT and Group 2 are diagnosed MM not due to undergo ASCT. Participants in the study received their usual care and were only asked to give extra blood and bone marrow samples when these were being taken as part of their normal appointments. Exact numbers of participants, baseline samples and follow up samples are stated in each results chapter. See Figure 2.1 for study flow chart.

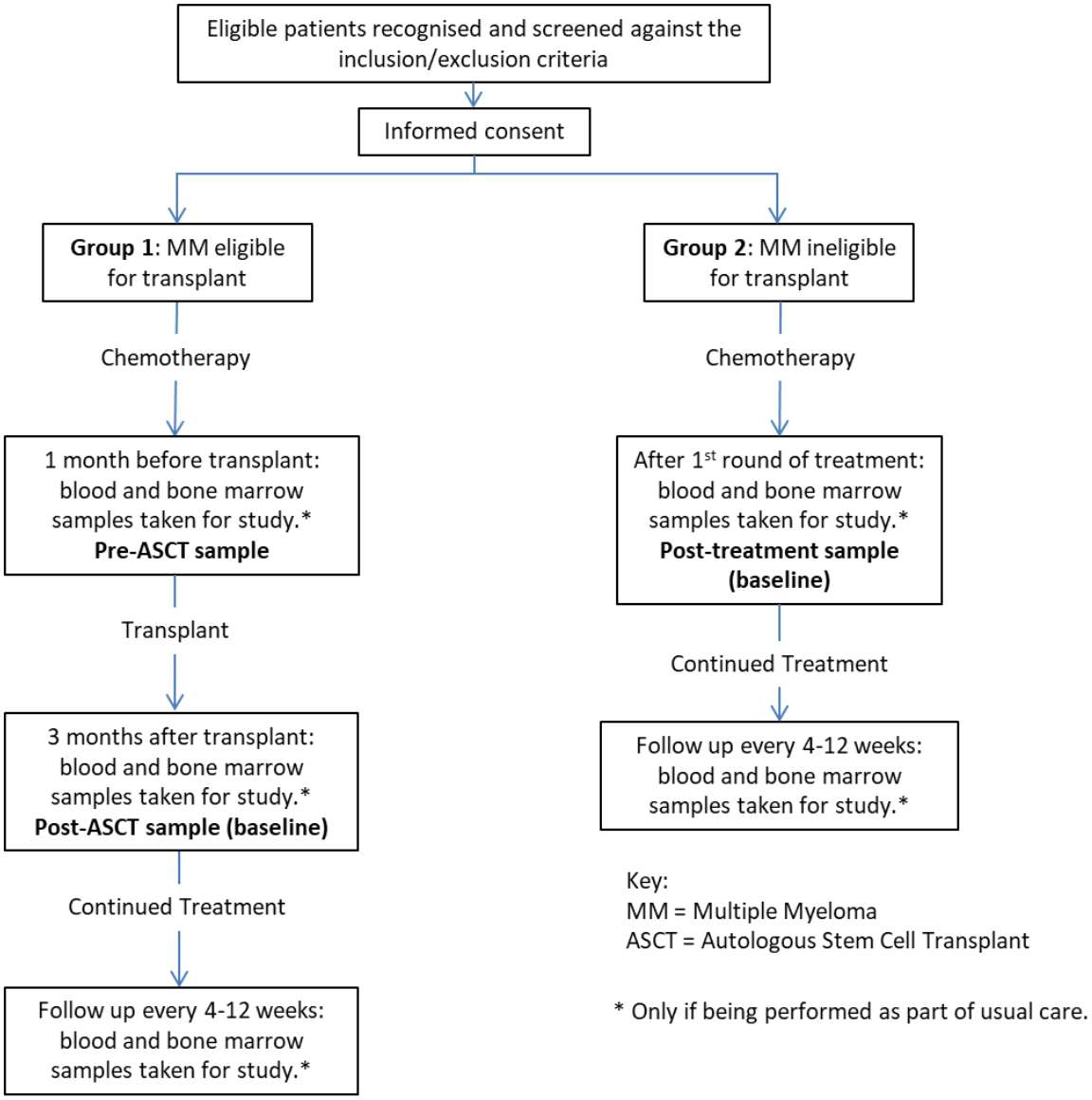


Figure 2.1: Flow diagram of the study measurement schedule

2.4 Study samples

Study samples were obtained based on clinician preference and no specific protocol was followed. The serum samples for this study were received, centrifuged, aliquoted and stored at -80^{o}C within 72 hours of collection prior to testing at a later date. Peripheral blood EDTA and bone marrow EDTA samples were stored at room temperature and tested within 48 hours and 24 hours of collection respectively.

2.5 Participant therapy

Participants were recruited both in and out of other clinical trials and participation in clinical treatment trials did not affect their status in this study. Patients received therapy according to

provider preference with the aim of obtaining a maximal response according to the IMWG response criteria. Participants received high-dose therapy based on: bortezomib + thalidomide + dexamethasone (VTD), bortezomib + lenalidomide + dexamethasone (VRD), carfilzomib + lenalidomide + dexamethasone (KRD), daratumumab + bortezomib + dexamethasone (DVd) and bortezomib + cyclophosphamide + dexamethasone (VCD) regimes.

2.6 Sample analysis

2.6.1 Flow cytometry analysis

Please see Chapter 3 section 3.2 for details of peripheral blood and bone marrow flow cytometry techniques.

2.6.2 Conventional serum methods

The tests performed on each serum sample routinely included total immunoglobulins (G, A and M), SPEP, IFE to confirm negative SPEP and FLC analysis. Please see Appendix 1 for standard operating procedures for these techniques. Total immunoglobulins, SPEP and FLC analysis were performed at the local participating centre for each participant. IFE analysis to confirm negative SPEP was performed by the study team at Oxford. Total immunoglobulins were measured using turbidimetry (Abbott, USA) with nationally accepted normal ranges used, SPEP and IFE were performed using standard laboratory procedures (Helena Biosciences, UK), FLC analysis was performed on the turbidimetric Optilite analyser using locally produced ranges (The Binding Site, UK) (Campbell *et al.,* 2020). In addition to M-protein testing, other biochemical parameters including serum creatinine, full blood count and serum calcium levels were retrieved from OUH NHS FT electronic medical records and data supplied by external sites via case report forms.

2.6.3 Heavy/light chain (HLC) analysis

HLC analysis was performed on the turbidimetric Optilite analyser (The Binding Site, UK) according to the manufacturer's instructions by the study team at Oxford (see Appendix 1). The manufacturers HLC normal reference ranges were used (Katzmann *et al.,* 2013) and the acronyms for the different HLC measurements are in section 1.4.2. An abnormal iHLC is considered to be above the normal range, an abnormal uHLC is considered to be below the normal range and an abnormal HLCr could be above or below the normal range.

2.6.4 Mass spectrometry (MS) analysis

MS analysis took place at The Binding Site laboratory (The Binding Site, UK) using a pre-commercial QIP-MS system based on MALDI-TOF-MS and was carried out by The Binding Site MS team. QIP-MS is based on the principle that each M-protein has a unique amino sequence and therefore a unique mass. For this technique polyclonal antibodies (anti-IgG, -IgA, -IgM, -total κ and -total λ light chain, -free κ and -free λ light chain) covalently attached to paramagnetic microparticles were incubated with serum, washed and treated to elute and reduce immunoglobulins. The molecular mass of each immunoglobulin in each patient was measured and clonality was established from the generated mass spectra. The mass spectra were inspected for peaks above a polyclonal background in the light chain mass-to-charge (*m/z*) range and the M-protein detected and isotyped if present. Although this method can now be utilised for quantitation of M-proteins, in the present study MS is only being used for detection and isotyping as the quantitation technology was not available at the time of analysis. Both FKFL MS and the enrichment based GAMKL MS methods were used for this analysis and the methods will initially be analysed separately in order to provide a full evaluation of this technique. Mass spectra was acquired over an *m/z* range of 10,000 to 30,000 and analysis was performed by two independent readers at The Binding Site, with the samples only unblinded after full analysis. MS was initially considered positive when the mass spectra from the GAMKL MS method produced evidence of a monoclonal protein as this is the most likely candidate to be taken forward for commercial use. The FKFL MS method results were then also considered when investigating discordant samples.

2.6.5 Disease assessment

Participants were stratified into response groups according to the IMWG uniform response criteria (Kumar *et al.,* 2016) using complete response (CR), very good partial response (VGPR), partial response (PR) and minimal response (MR). Participants in VGPR, PR, or MR were grouped together as “non-CR”. Disease progression or relapse was determined according to the IMWG consensus statement (Rajkumar *et al.,* 2011). Each participant had a baseline sample which was taken after induction therapy or post-ASCT. Baseline in this study is defined as the initial sample taken post-therapy (see Figure 2.1). Periodic serial samples were then taken at each clinic visit (median every 2 months) during routine follow up from baseline. Response was assessed at baseline and each follow up point. Best response samples are identified as the maximal response at any time post-ASCT or chemotherapy as defined using IMWG criteria (Kumar *et al.,* 2016).

2.6.6 Survival assessment

For post-ASCT/therapy analysis, progression free survival (PFS) was defined as time from ASCT/end of induction therapy to disease progression or death due to any cause. For best response analysis PFS was calculated as time from achievement of maximal response (MR or above) to first observation of relapse or death. Participants without observed events by the end of the follow up period were censored. Overall survival (OS) was not measured due to the short follow up time.

2.6.7 Statistical analysis

Data were analysed using GraphPad Prism 9 (California, USA). Differences in categorical variables were assessed with the Fisher's exact test, differences in continuous variables were evaluated with parametric (t-test) and non-parametric (Mann-Whitney) tests. Bland-Altman and Passing-Bablok regression analysis were used for method comparison. Categorical agreement between methods was assessed using Cohen's kappa statistic. For sensitivity and specificity analyses the assumption was made that bone marrow NGF MRD or serum IFE were the gold standard methods. Kaplan-Meier survival curves were used to plot and compare PFS curves between groups of participants. Survival curves were compared using the two-sided log-rank test to indicate significance and log-rank hazard ratios (HR) with confidence intervals (CI) were calculated. All analyses were two-sided and statistical significance assumed at $p \leq 0.05$.

Chapter 3 - Validation of a next generation flow cytometry assay for bone marrow and peripheral blood samples

3.1 Introduction

Overall survival rates in MM have vastly improved in recent years following advances in drug therapy and optimal use of ASCT (Kumar *et al.,* 2016). Despite these improvements, MM remains an incurable disease and the majority of patients will relapse and require further treatment (Sonneveld *et al.,* 2016). Relapse is due to the continued presence of a small number of abnormal plasma cells after treatment termed minimal residual disease (MRD). The evaluation of MRD is an area under great scrutiny and development. With the use of more and more effective therapies, detecting these residual cells in the bone marrow requires increasingly sensitive techniques. It is MFC/NGF studies which are at the forefront of MRD publications (Galtseva *et al.,* 2017).

3.1.1 Samples for MRD plasma cell detection

MRD has traditionally been measured in bone marrow as this is the primary site of the residual abnormal plasma cells in MM patients. MRD detection in bone marrow has consistently been shown to be significantly associated with both progression free and overall survival (Munshi *et al.,* 2017). When performing a bone marrow aspiration a first pull aspirate is recommended for optimal MRD detection; this is to avoid haemodilution of the sample with peripheral blood, which can lead to false negative results when detecting very small numbers of cells and has been found to be the most common technical error in BM MRD assessment (Joshi *et al.,* 2008). As MRD detection is still a research test for MM patients, current practice uses the first pull for smear preparations, histology and/or routine immunophenotyping; therefore it was expected to be receiving second or third pull samples for this study and will attempt to determine haemodilution through identification of non-plasma cell bone marrow populations in order to give accurate results.

Due to the invasiveness of bone marrow sampling, alternative blood-based methods have been explored for sensitive residual disease detection. One such method is the detection of abnormal circulating tumour plasma cells (CTPCs) in peripheral blood (PB). With improved flow cytometry technology there is potential to sensitively detect the presence of CTPCs and correlate this to the presence of MRD in the bone marrow and with survival outcomes.

3.1.2 MFC for MRD plasma cell detection

There are a number of reasons why MFC has been the most widely studied technology for use in MRD detection which apply both in the NHS and worldwide. Flow cytometers are readily available in the majority of diagnostic laboratories, panels for MRD detection can be standardised and software can be produced to homogenise analysis (Flores-Montero *et al.,* 2017). Early studies which used MFC had a lower sensitivity (10^{-4}) than current protocols and instrumentation allow (10^{-5} – 10^{-6}). These technological developments have led to increases in the number of fluorescently labelled antibodies which can be detected at once, increases in speed of acquisition and increases in the total number of cells which can be acquired, allowing for greater sensitivity.

A standardised MFC MRD procedure has been recommended by the EuroFlow group (Van Dongen *et al.,* 2012, Stetler-Stevenson *et al.,* 2016). This method consists of two tubes and 8 fluorescently labelled antibodies per sample, with the collection of 5 million cells total. Despite being expertly standardised this method does have its drawbacks when considering implementation into a national health service (see section 1.3.1). In an attempt to create a more accessible method, the Memorial Sloan Kettering (MSK) group developed a comparable single-tube 10 colour method (Roshal *et al.,* 2017).

The decision was therefore made to use a single tube MFC panel for this study. This allows for a reduction in both antibody costs and time spent on assay preparation and analysis, important considerations for introduction into routine clinical use. A single tube containing an initial 8 fluorochromes, with the option of 2 drop-ins (see Table 3.1), is available from Beckman Coulter, the company which also supplies the CytoFlex flow cytometer which allows for acquisition of >10million events from a single tube and contained within a single file. The markers selected for use allow normal and abnormal plasma cells to be distinguished, with antigen expression remaining stable from diagnosis to MRD stages of disease (Arana *et al.,* 2018). CTPCs can be detected using the same antigen panel as bone marrow plasma cells, although some differences in expression of CD38, CD138, CD56 and CD81 have been noted (Paiva *et al.,* 2013, Terpos *et al.,* 2019). It is important to validate an assay when introducing it into the laboratory to establish performance specifications to meet the intended use of the assay. The bone marrow MRD MFC assay will also be used as the gold standard method against which the blood based MRD detection assays will be compared.

Table 3.1: Cell surface markers used in the flow cytometry panel

Antigen	Normal plasma cell immunophenotype expression	Aberrant plasma cell immunophenotype expression	Percentage of cases with abnormal expression	References
CD19	+	-	95%	Rawstron *et al.,* 2008
CD27	++	- or dim+	40-68%	Guikema *et al.,* 2003, Rawstron *et al.,* 2008
CD38 (PC marker)	++	dim+	92%	Rawstron *et al.,* 2008
CD45	+	-	69-80%	Pellat-Deceunynck and Bataille 2004, Soh *et al.,* 2017
CD48 (PC marker-optional drop-in)	+	+	80-100%	Hosen *et al.,* 2012, Muccio *et al.,* 2016
CD56	-	++	60-76%	Sahara *et al.,* 2002, Lin *et al.,* 2004
CD81	+	- or dim+	45%	Paiva *et al.,* 2012
CD117 (optional drop-in)	-	+	30-55%	Kraj *et al.,* 2004, Gupta *et al.,* 2018
CD138 (PC marker)	+	+	80-100%	Rawstron *et al.,* 2008
CD200	weak	+ / ++	65-86%	Moreaux *et al.,* 2006, Alapat *et al.,* 2012.

3.1.3 Aims

1. To validate a flow cytometric assay for the measurement of low levels of bone marrow and circulating tumour plasma cells.

2. To standardise MRD analysis across different patients and sample types.

3.2 Methods

3.2.1 Subjects and samples

This was a laboratory based study performed in 2019. As healthy controls have low numbers of CTPCs and bone marrow samples cannot ethically be obtained from this population, enrolled participant samples were also used for assay validation. Up to 20mls of healthy control peripheral blood (PB) samples were collected in EDTA tubes from 5 healthy donors to validate the analytical specificity of the panel. Healthy controls were eligible if they had no history of a plasma cell dyscrasia and were samples were obtained from volunteers amongst the immunology laboratory staff in Oxford. 2 bone marrow samples (2ml each in EDTA) were obtained from myelodysplastic syndrome (MDS) patients, a haematological malignancy where abnormal plasma cells should not be present in the bone marrow. Bone marrow and PB samples were obtained from 8 patients with confirmed MM diagnoses. Samples were collected in EDTA tubes and run within 24 (bone marrow) or 48 (PB) hours of collection.

3.2.2 Sample preparation

Bone marrow: Samples were prepared according to a pre-defined protocol (Campbell *et al.,* 2019, Bayly *et al.,* 2020, Soh *et al.,* 2020) in combination with the kit recommendations (Beckman Coulter, Inc.). Starting sample volumes ranged from 3-5ml. Briefly, samples were mixed thoroughly and placed in a conical 50ml centrifuge tube. Bulk erythrocyte lysis was performed on whole samples using an ammonium chloride lysing reagent (8.30g NH_4Cl, 2ml 0.5M EDTA solution, 1L distilled water, pH 7.15) at a volume ratio of 1:9 (Stetler-Stevenson *et al.,* 2016). Lysis was performed at room temperature with the samples agitated on a roller for 15 minutes followed by centrifugation at 1400 RPM for 7 minutes. Samples were washed twice in PBS/5% FCS (phosphate buffered saline with 5% foetal calf serum (Fisher Scientific, UK)), centrifuging at 1400 RPM for 7 minutes between each wash. Cells were re-suspended in 150µl PBS/5%FCS. Viability and number of cells were assessed by light microscopical analysis of trypan blue staining using a haemocytometer, and re-suspension volume adjusted to give ~$5x10^6$ - $10x10^6$ nucleated cells per 100ul using the calculation below.

$$\text{No. of cells per ml in original mixture} = \left(\frac{\text{No. of cells counted}}{\text{No. of complete 4x4 grids counted}} \right) \times 10^4 \times \text{Sample dilution}$$

100ul of sample was added to a DuraClone RE PC Tube (Beckman Coulter, Inc.). This is a pre-mixed, dry reagent cocktail containing fluorescent antibodies to stain 3-$10x10^6$ leucocytes in

100ul. The fluorescent labels in the reagent cocktail and their cell surface targets are: CD81 - fluorescein isothiocyanate (FITC) (Clone JS64), CD27 – R-phycoerythrin (PE) (Clone 1A4CD27), CD19 – PE-cyanine 5.5 (PC5.5) (Clone J3-119), CD200 – PE-cyanine 7 (PC7) (Clone OX-104), CD138 - allophycocianin (APC) (Clone B-A38), CD56 – APC-Alexa Fluor 750 (APC-A750) (Clone N901), CD38 Pacific Blue (Clone LS198-4-3), CD45 Krome Orange (Clone J.33). The sample and antibodies were incubated for 15 minutes at room temperature in the dark and then washed with PBS/5%FCS and the pellet re-suspended in 500ul PBS/5%FCS.

Peripheral blood: Preparation of PB samples followed the same method as bone marrow samples but starting volumes were higher with a range of 4-20ml. Varying sample volumes are adjusted for by cell counting prior to the addition of ~$5x10^6$-$10x10^6$ nucleated cells per 100ul to the DuraClone RE PC tubes.

All samples were run immediately after processing and therefore a fixative agent was not required. Total preparation time was 70-100 minutes.

3.2.3 Sample acquisition

Samples were acquired on a 3 laser (488 nm, 638 nm, 405 nm) CytoFlex flow cytometer (Beckman Coulter, Inc.). Daily start up and shut down cleaning was performed according to manufacturer's instructions. Compensation settings were applied using compensation beads with defined fluorescence provided with the DuraClone RE PC tubes and manually adjusted using samples which contained the cells of interest. Voltages were adjusted to allow visualisation of all cells of interest. These settings were saved for future experiments and the settings re-checked with every DuraClone RE PC tube lot number change. Quality control (QC) beads (Beckman Coulter, Inc.) were run upon each use of the CytoFlex for verification of the flow cytometer's optical alignment and fluidics system. Rare populations within samples will be randomly distributed and therefore Poisson statistics apply, where precision will increase as more events are acquired. The total number of events acquired per tube was set to allow a minimum of 5 million and maximum of 10 million cells to be acquired, following best practice guidelines (Mailankody *et al.,* 2015). The acquisition rate was set at approximately 3000 events/second in order to allow all rare cell events to be captured. Each tube took approximately 30 minutes to run and the tube was agitated at regular points during the run to avoid the cells settling.

3.2.4 Flow cytometry plot analysis

Due to the antibodies being freeze dried in one tube, isotype controls and fluorescence minus one (FMO) staining could not be used for these antibodies. Instead, using density plots and gating on cells which are negative for the markers of interest, positive expression gates could be set for each marker. Separate templates were created to be used for either bone marrow or PB samples. These templates should require minimal adjustment between patients.

3.2.5 Assay validation

3.2.5.1 Limit of blank (LOB)

This is the highest event result which is likely to be observed for a blank sample and assesses the contamination rate in the gating strategy. 5 healthy donor samples were run in order to detect plasma cells with an abnormal phenotype. A conservative LOB is calculated from the mean and SD from the blank results:

LOB = Mean blank + 3(SD blank)

3.2.5.2 Limit of detection (LOD)

The LOD is closely related to the LOB and is the lowest level of cells which can be reliably detected above the level of the blank. The LOD is calculated as:

LOD = LOB + 3(SD blank)

3.2.5.3 Lower limit of quantitation (LLOQ)

It needs to be ensured that the abnormal events detected are part of a discrete population. The desired sensitivity for this assay is 0.001% (Galtseva *et al.,* 2018). Therefore it needs to be ensured that the assay can detect at least 50 abnormal cells out of 5 million with high precision. The LLOQ was evaluated using dilution and spiking experiments of a healthy control PB sample with either a bone marrow or PB MM sample containing a known amount of abnormal plasma cells. The number of abnormal events in 100ul of the sample was determined and spiking experiments carried out using cells from a healthy control. 1:2, 1:4, 1:10, 1:32 and 1:100 dilutions were performed to obtain set numbers of events. Dilutions were each run 3 times and the %CV calculated. The LLOQ will be accepted if the following criteria are met: all 3 replicates are >LOD, an appropriate titration effect is evident, and the %CV is ≤30%.

3.2.5.4 Accuracy

There is no reference standard with which to compare closeness of agreement of this assay. However, there is a pilot External Quality Assessment (EQA) scheme in place for MRD detection in MM samples. 2 samples are supplied and all participating centres must identify and quantify the level of disease in these samples and report the immunophenotype of the abnormal plasma cells detected. Samples are manufactured using a plasma cell myeloma sample, stabilised whole blood and stem cell harvest material. Comparison of results with consensus will be used to measure accuracy.

3.2.5.5 Precision

Intra-assay precision is established by running samples in triplicate and calculating the percentage coefficient of variance (%CV). These were performed for different disease levels using the same samples as the LLOQ experiments as recommended by consensus guidelines (Wood *et al.,* 2013). Due to the unstable nature of plasma cells outside of the body inter-assay precision is not able to be performed. A desirable %CV target is <10%, however for rare event analysis such as MRD an assay precision of ≤30% is acceptable near the limit of detection (LOD) (Lee *et al.,* 2006).

3.2.5.6 Sample quality

Bone marrow sample quality can be assessed to ensure it is a representative sample and excessive haemodilution is not present. This is performed by detecting other cell populations expected in normal, diagnostic MM and post-treatment MM bone marrow samples (see Table 3.2). These populations are detected using cell surface markers contained within the MRD MFC panel. If such populations cannot be detected then the results of the MRD analysis should be called into question and any results, especially MRD negativity, reported with extreme caution.

Table 3.2: Cell types to be used for quality assessment of bone marrow samples

Cell Type	Markers
Eosinophils	SSC^{hi}, $CD45^{++}$, $CD81^{+}$
Natural Killer (NK) cells	$CD45^{+}$, CD56+
B cell precursors	$CD19^{+}$, $CD38^{++}$, $CD45^{dim}$, $CD81^{++}$, $CD27^{+/-}$
Memory T/B cells	$CD45^{+}$, $CD38^{-/dim}$, $CD27^{+}$

3.2.5.7 Patients on anti-CD38 therapy

Therapeutic anti-CD38 antibodies, under the name daratumumab, are becoming more frequently used in MM patients and these could reduce the expression of CD38 on plasma cells, therefore interfering with the $CD138^+/CD38^+$ gating strategy recommended to initially identify plasma cells (Flores-Montero *et al.,* 2017). By using the drop-in antibody CD48 plasma cells can be gated using $CD138^+/CD48^+$ (Hosen *et al.,* 2012). Gating using CD48 was tested on MM patients not undergoing anti-CD38 therapy in order to compare plasma cell numbers (normal and abnormal) between the two staining and gating methods to ensure the $CD48^+$ method is identifying the same cells.

3.2.6 Statistics

All statistics were performed using GraphPad Prism 9 software. The linearity of the spiked samples was assessed by linear regression. The %CV for intra-assay precision and LLOQ samples was calculated using the mean and SD values.

3.3 Results

3.3.1 Sample acquisition

Bone marrow and blood samples were obtained from 8 MM patients enrolled in the study. Samples were processed in order to obtain a minimum of $5x10^6$ leucocytes for labelling and acquisition. The majority of bone marrow samples were able to produce close to $5x10^6$ cells whilst CTC samples had a lower mean cell acquisition of $4.4x10^6$ (see Table 3.3). This lower cell recovery is due to the lower number of cells in PB vs bone marrow and low volumes of samples received from some patients. Samples were run until either >5million total cells were collected or until the tube ran dry.

Table 3.3: Total events collected from MM bone marrow and PB samples

Patient	Total bone marrow cellular events	Total PB cellular events
1	4,966,993	4,638,368
2	5,478,865	3,235,883
3	5,043,597	4,823,095
4	5,149,561	4,592,931
5	4,317,452	3,995,515
6	5,189,843	5,443,265
7	4,849,260	4,963,385
8	4,967,967	4,138,368
Mean	4,995,442	4,478,851
SD	333,844	677,848
%CV	7%	15%

3.3.2 MFC analysis

The gating strategy for both normal and abnormal plasma cell detection was developed from the 5 healthy control PB samples, 2 MDS bone marrow samples and reviewed using MM patient samples (see Figure 3.1). The sample in Figure 3.1 is from an MM patient with a relatively large amount of disease in the bone marrow to allow visualisation of cell populations.

Initially, doublets are excluded by plotting FSC width vs FSC area (Figure 3.1: Panel A). They are excluded to ensure only single cells are counted and no false positive or negative signal results occur. A forward scatter (FSc) vs side scatter (SSc) plot is then applied to the singlets to capture all the cells and exclude dead cells and debris which may be present e.g. unlysed or fragmented red blood cells, platelet aggregates (Figure 3.1: Panel B). It is important to exclude these as a false increase in the number of total cells studies could affect the final MRD result, however a large gate is set in order to capture as many rare cell events as possible. These cells are then plotted on a CD45 vs SSc plot to help discriminate neutrophils, lymphocytes, monocytes and red cell debris (Figure 3.1: Panel C). A generous CD45 lymphocyte gate is applied on the CD45 vs SSc plot in order to capture both CD45+ and CD45- plasma cells. From this a plasma cell gate is set on a CD38 vs CD138 plot, using $CD38^{+}$ $CD138^{+}$ as the plasma cell markers (Figure 3.1: Panel D (PC1+PC2)). This gate needed to be slightly adjusted for some patients due to dimmer CD138 expression. All subsequent marker plots are used to distinguish phenotypically normal from abnormal plasma cells using those cells captured in the plasma cell gate (see Table 3.1). Per the

EuroFlow recommendations, confirmation of the clonal nature of plasma cells must be evidenced by the expression of 2 or more aberrant markers and the cells must be negative for at least one marker from the panel (Van Dongen *et al.,* 2012).

As the gating for the plasma cells differed slightly between BM and PB samples, two templates have been created for the different sample types. This will help to standardise the data analysis and avoid an overreliance on manual gating which will lead to operator-specific variations. An abnormal population of plasma cells will be reported as both the number of events, alongside the total “all cell” events and the %MRD as a % of all cells. Absolute counts cannot be reported as the absolute count of the original sample is unknown.

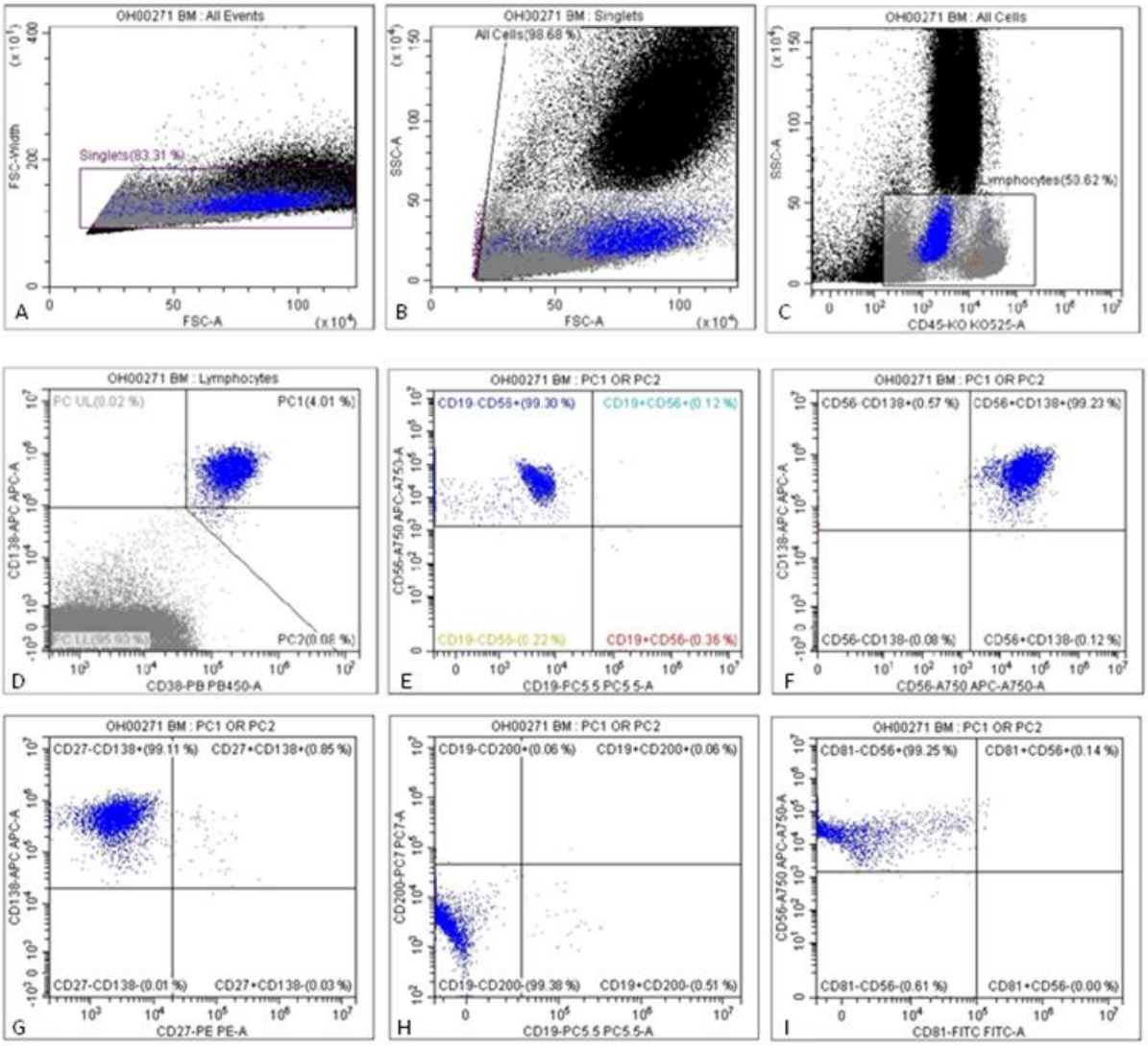


Figure 3.1: MRD bone marrow gating strategy template illustrating an example of bone marrow from a treated MM patient.

This Figure shows an example plot of a pre-ASCT bone marrow from a MM patient (patient OH00271). The bivariate dot plots create a sequential gating strategy for the identification of low levels of abnormal plasma cells using an 8-colour single tube flow cytometric assay. A rectangular plot is used to include the single cell population and exclude doublets (panel A). Panels B – D are used to resolve plasma cells from all cells collected. From the plasma cell gates, PC1 and PC2 on panel D, the phenotypes of the plasma cells are defined using panels E-I and the immunophenotype information in Table 1. Abnormal plasma cells (blue population: $CD19^{-}CD56^{+}CD27^{-}CD200^{-}CD81^{-/dim}CD45^{dim}$) and a small population of normal plasma cells (red population: $CD19^{+}CD56^{-}CD27^{+}$) can be seen in this panel.

The CTPC template (not shown here) follows the same format but with a more generous plasma cell gate in panel D to capture $CD38^{dim}/CD138^{dim}$ plasma cells.

3.3.3 Assay validation

3.3.3.1 Limit of blank (LOB)

5 healthy control PB samples (1-5) and 2 MDS patient bone marrow samples (6-7) were used to calculate the LOB (Table 3.4). Fewer events were collected from the MDS bone marrow samples due to the low starting volume. The LOB can be calculated by number of abnormal events detected or by the percentage of abnormal cells in total events. Due to the range in events collected % abnormal plasma cells was used to determine the LOB.

Table 3.4: Results from acquisition of limit of blank samples

Sample No.	Total Cellular Events	Abnormal plasma cell events	% Abnormal plasma cell of total cellular events
1	3,765,152	2	0.00005%
2	5,199,480	4	0.00008%
3	5,945,967	3	0.00005%
4	4,182,011	3	0.00007%
5	5,702,931	4	0.00007%
6	1,131,532	1	0.00009%
7	968,637	1	0.0001%
Mean	**3,842,244**	**2.6**	**0.00007%**
SD	2,059,808	1.3	1.89×10^{-7}%

The LOB is calculated as

Mean blank + 3(SD blank) = $0.00007 + 3(1.89 \times 10^{-7}) = 0.00013\%$

If 5 million cells are collected this gives a LOB of 7 events.

3.3.3.2 Limit of detection (LOD)

Using the values from the LOB experiments, the LOD is calculated as:

LOB + 3(SD blank) = $0.00013 + 3(1.89 \times 10^{-7}) = 0.0002\%$

If 5 million cells are collected this gives a LOD of 10 events.

3.3.3.3 Lower limit of quantitation (LLOQ)

Bone marrow - The results of spiking abnormal bone marrow plasma cells into healthy control PB are shown in Table 3.5. 5 million total cell events were acquired for each sample and each spiking dilution was run in triplicate. Observed versus expected abnormal PC events were examined and analysed by linear regression. Correlation between the observed and the expected abnormal PC levels was very close to 1.0 (r^2 = 0.987), showing that at all spiking levels, results were closely correlated with the concentration of cells added (Figure 3.2). At a mean of 19 events the %CV was 18.98% which is within the accepted %CV limit for rare event analysis with an LOD below 0.01% (Lee *et al.,* 2006). In order to be conservative and avoid any artifacts, 25 cells was chosen as the LLOQ, giving a sensitivity of 0.0005% if 5 million total cell events are acquired and 0.0003% if 10 million total cell events are acquired. Therefore, if less than 25 abnormal cells are detected per sample it can classed as MRD negative (pending sample quality checks).

Table 3.5: Bone marrow spiking experiment results for evaluation of LLOQ

Spiking Dilution	Expected Abnormal PC Events (%MRD)	Observed Abnormal PC Events	Observed Abnormal PC Events Mean (%MRD)	SD	%CV
Neat	N/A	1913	1879 (0.04%)	33.51	1.78%
		1846			
		1878			
1:4	470 (0.01%)	415	455 (0.01%)	39.04	8.58%
		493			
		457			
1:10	188 (0.004%)	141	160 (0.003%)	16.65	10.43%
		165			
		173			
1:32	59 (0.001%)	44	48 (0.001%)	4.51	9.33%
		53			
		48			
1:100	19 (0.0004%)	15	19 (0.0004%)	3.61	18.98%
		20			
		22			

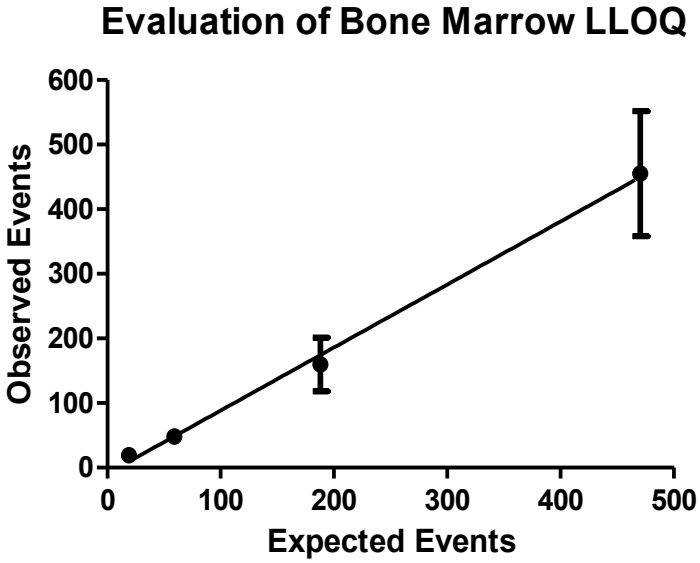


Figure 3.2: Linearity and LLOQ assessment of spiked bone marrow samples. Samples were run in triplicate and plotted from 1:4 to 1:100 dilutions with 95% CI error bars shown. Linear regression delivered a slope of 0.9754 and a correlation coefficient of $r^2 = 0.987$ ($p<0.0001$).

Peripheral blood - The results of spiking abnormal CTPCs into healthy control PB are shown in Table 3.6. 5 million total cell events were acquired for each sample and each spiking dilution was run in triplicate. Due to the lower starting concentration of abnormal CTPCs, fewer dilutions than for the bone marrow sample were needed to reach the desired sensitivity. Observed versus expected abnormal PC events were examined and analysed by linear regression. Correlation between the observed and the expected abnormal PC events was very close to 1.0 ($r^2 = 0.9479$) (Figure 3.3). At a mean of 28 events the %CV was 10.78% which is within the accepted %CV limit for rare event analysis with an LOD below 0.01% (Lee *et al.,* 2006). At a mean of 6 events the %CV was 32.87% alongside under-recovery of CTPCs at 0.0002% and therefore this is not an acceptable LLOQ levels. Due to the low %CV at 28 events (0.0006%) the LLOQ for the bone marrow plasma cells will be replicated and set at 25 events in 5million cells (0.0005%) for peripheral blood. Therefore, if less than 25 abnormal cells are detected per sample it can be classed as CTPC negative.

Table 3.6: Peripheral blood spiking experiment for evaluation of LLOQ

Spiking Dilution	Expected Abnormal PC Events (%MRD)	Observed Abnormal PC Events	Observed Abnormal PC Events Mean (%MRD)	SD	%CV
Neat	N/A	55	51 (0.001%)	3.51	6.84%
		51			
		48			
1:2	26 (0.0005%)	25	28 (0.0006%)	3.06	10.78%
		31			
		29			
1:4	13 (0.0002%)	7	6 (0.0001%)	2.08	32.87%
		8			
		4			

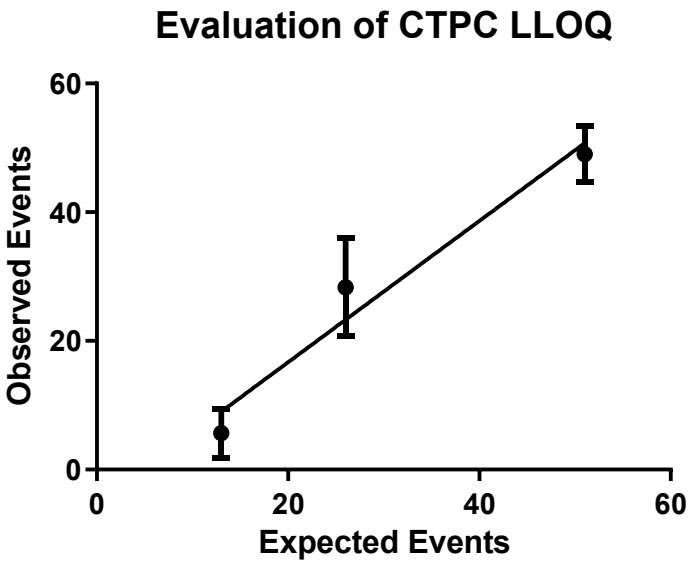


Figure 3.3: Linearity and LLOQ assessment of spiked whole peripheral blood samples.

Samples were run in triplicate and plotted from neat to 1:14 dilutions with 95% CI error bars shown. Linear regression delivered a slope of 1.098 and a correlation coefficient of $r^2 = 0.9479$ ($p<0.0001$).

3.3.3.4 Accuracy

Due to no standard reference material being available a pilot EQA sample has been introduced for MRD detection in plasma cell myeloma. 2 stabilised specimens to mimic bone marrow samples from a known MM patient were supplied, taken at different post-treatment time points. There were 46 participants in the trial. The reported included the total number of cells analysed, percentage MRD population and staining intensity for each marker. A satisfactory performance was attained for all trial samples with z scores within the defined limits and % MRD results close to the all participants mean (Table 3.7). Accuracy will continue to be monitored under this EQA scheme 4 times per year.

Table 3.7: Results and Performance in MM MRD EQA scheme

Percentage MRD Population	Our Result (%)	Robust Mean (%)	Robust SD (%)	z score (Limits: -3.50 to +3.50)
Sample 1	0.70%	0.90%	0.56%	-1.48
Sample 2	0.30%	0.48%	0.36%	-1.25
Sample 3	0.02%	0.07%	0.03%	-1.67
Sample 4	0.001%	<0.005%	N/A	N/A
Sample 5	0.10%	0.09%	0.04%	0.25
Sample 6	0.04%	0.05%	0.02%	-0.25
Sample 7	0.48%	0.44%	0.15%	0.27
Sample 8	0.02%	0.02%	0.01%	0.00

3.3.3.5 Precision

Spiked health control samples were run in triplicate as part of the LLOQ experiments. For each of the % MRD levels down to 0.0004% the %CV remained below 20%. This is an acceptable %CV for rare event analysis (Lee *et al.,* 2006).

3.3.3.6 Sample quality

As the quality of bone marrow aspirate samples can vary greatly, and will have an impact on the legitimacy of the MRD result, it is important to ensure that the flow cytometry panel could identify normal cell populations alongside the plasma cells. Experienced flow cytometrists can examine the CD45 vs SSc plot to conduct an initial check that monocytes, erythrocytes and granulocytes are present. Further to this, cell lineages which could be identified through simple gating strategies were chosen to allow quality checks to be performed quickly on every bone marrow sample (Figure 3.4). If these cell populations are not present in sufficient numbers (as

determined by total cell numbers acquired, experience with the assay and reported normal ranges) the quality of the bone marrow sample should be called into question (Flores-Montero *et al.,* 2017).

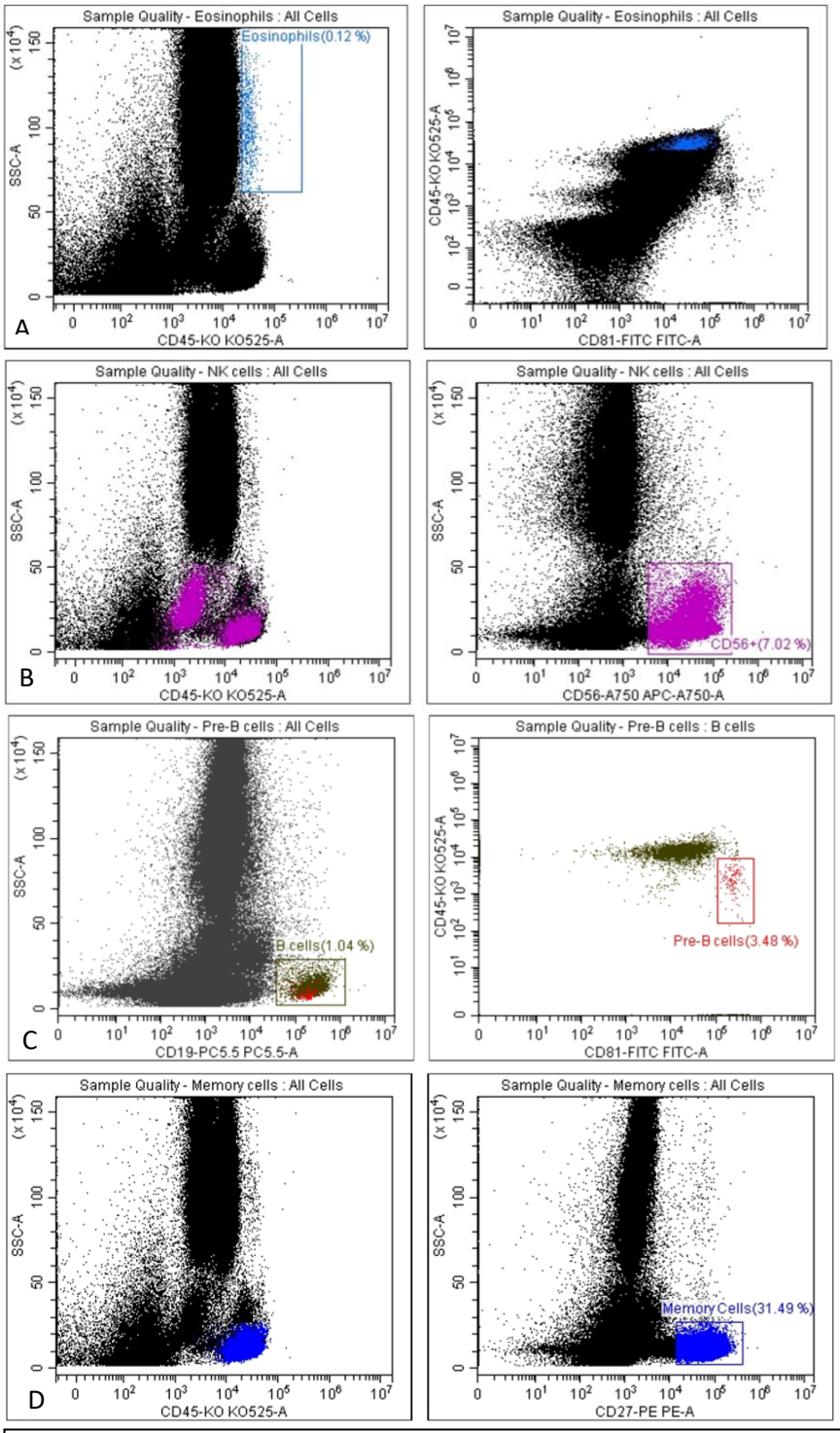


Figure 3.4: Illustrating examples of post-acquisition checks for bone marrow sample quality.

The control populations are as follows and are illustrated using bone marrow from an MM patient: **Panel A)** Eosinophils - SSC^{hi}, $CD45^{++}$, $CD81^{+}$. **Panel B)** NK cells - $CD45^{+}$, CD56+ (note the $CD45^{dim}CD56^{+}$ which are abnormal plasma cells). **Panel C)** Precursor B cells - $CD19^{+}$, $CD45^{dim}$, $CD38^{++}$, $CD81^{++}$, $CD27^{+/-}$ **Panel D)** Memory T/B cells - $CD45^{+}$, $CD38^{-/dim}$, $CD27^{+}$.

3.3.3.7 Patients on anti-CD38 therapy

The aim of this experiment was to compare the accuracy of the assessment of MRD in patients using the original $CD38^+/CD138^+$ plasma cell gating versus $CD48^+/CD138^+$ gating. The CD48 antibody is a drop-in to a pre-made freeze dried antibody tube and is used at a volume of 5µl. 3 PB samples from MM patients not on anti-CD38 therapy were stained using both the original panel and the original panel plus CD48-electron coupled dye (ECD) (clone: J4-57) antibody. Different compensation settings are needed to include the ECD fluorochrome. Compensation was performed using manufacturer supplied compensation beads as discussed in section 3.2.3. Compensation settings either including or excluding ECD can be easily switched between depending on the needs of the sample. An example of the $CD38^+/CD138^+$ plasma cell gating versus $CD48^+/CD138^+$ gating for one patient is shown in Figure 3.5.

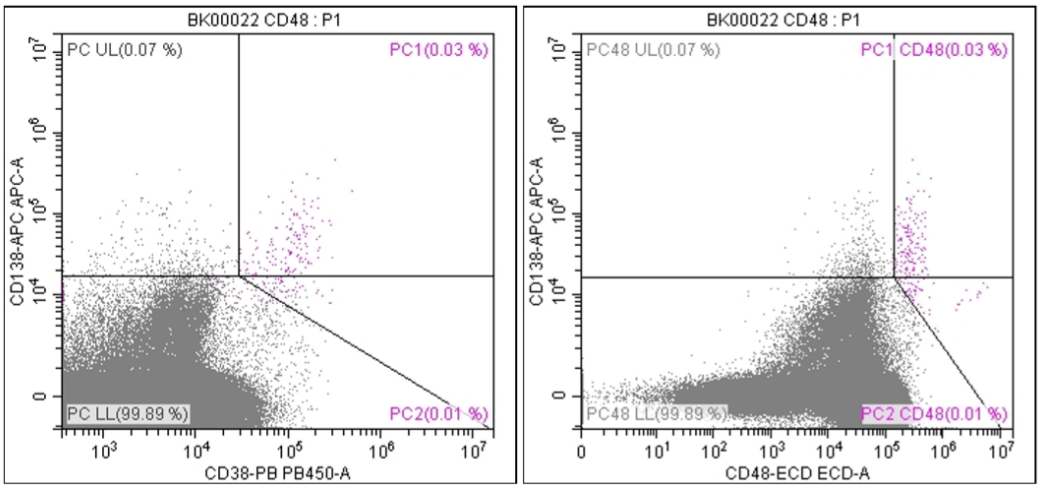


Figure 3.5: Example of CD38+/CD138+ vs CD48+/CD138+ gating of plasma cells on the same MM patient PB sample.

Cells captured by both methods are highlighted in pink. The plasma cell % of total cells acquired was the same for both gating methods, with the two gating methods capturing the majority of the same cells.

When evaluating the two gating methods the number of gated plasma cell events and the % plasma cells of total cells acquired were compared (Table 3.8). For all 3 samples the %CV between the two methods for plasma cell events was <10%. At the lower limit of around 25 events the %CV was 9% which gives confidence that these gating methods are equivalent and the use of CD48 gating is a suitable alternative for those patients not expressing CD38 on their plasma cells.

Table 3.8: Comparing CD38+/CD138+ vs CD48+/CD138+ gating of plasma cells.

	CD38+/CD138+ Gating		CD48+/CD138+ Gating				
Sample	Plasma Cell Events	% Plasma Cells of Total Cells	Plasma Cell Events	% Plasma Cells of Total Cells	Mean Plasma Cell Events	SD Plasma Cell Events	%CV Plasma Cell Events
1	215	0.04%	205	0.04%	210	7.07	3%
2	21	0.0007%	24	0.0008%	22.5	2.12	9%
3	43	0.01%	44	0.01%	43.5	0.71	2%

3.4 Discussion

The experiments described in this chapter aimed to introduce and validate a flow cytometric assay to primarily test bone marrow plasma cells but also plasma cells found in PB (CTPCs). In order to validate the assay sensitivity, reproducibility and utility across patients and sample types needed to be determined. Validation showed that this assay can be used in both bone marrow and PB samples. The LOB, LOD and LLOQ were comparable to other flow cytometry MRD assays and the validated sensitivity of this assay means that it can be used as a next-generation flow (NGF) MRD technique.

3.4.1 Flow cytometry analysis

The NGF MRD assay is based on consensus recommendations (Arroz *et al.,* 2015) and uses a pre-made 8-antibody panel. Instrument set up and compensation is performed automatically by the CytoFlex, although compensation settings did require some minor manual adjustments for optimisation. Recent studies have shown the usefulness of using dried reagents for both bone marrow and peripheral blood assays and advantages include long stability, room temperature storage, less pipetting of small amounts of antibody and therefore a reduction in accidental operational mistakes (Van der Velden *et al.,* 2017). The single tube assay also decreases the labour input and time required, making this single tube flow cytometry assay more accessible to routine laboratories.

The panel of antibodies chosen for MRD analysis has been previously optimised by Beckman Coulter (UK) based on EuroFlow recommendations and studies examining other useful markers in abnormal plasma cell discrimination (Van Dongen *et al.,* 2012, Stetler-Stevenson *et al.,* 2016). The use of this 8-colour panel allows for abnormal PCs to be distinguished against a background

of normal plasma cells without the knowledge of a diagnostic sample. The large number of immunophenotypic markers used means that any changes in the phenotype of the abnormal cells over time will not affect the ability of the assay to distinguish the abnormal and normal plasma cells from each other (Wood *et al.,* 2013). The panel of antibodies chosen does not include kappa/lambda identification which would help determine the clonality of the plasma cell population. However, two recent studies have shown that kappa/lambda antibodies would only be useful in a small percentage of patients and does not provide additional information to the panel validated here (Dold *et al.,* 2020, Riebl *et al.,* 2021). By verifying that this panel is in line with other national MFC/NGF MRD assays through EQA performance then the gating strategy can be validated. Of note, the NGF assay used in this study was the only centre to acquire >2million events and therefore had the lowest LLOQ of all participating centres.

The same panel of antibodies has been chosen for detecting plasma cells in both bone marrow and PB of MM patients. Paiva *et al.* (2013) noted that CD38, CD138 and CD81 were downregulated on CTPCs compared with bone marrow plasma cells. Due to these observations, a different analysis template has been created for CTPC analysis so that minimal adjustments of each template are needed by the operator, improving standardisation of analysis. In the future, multiple users can be assessed using these templates to evaluate inter-operator analysis.

3.4.2 Assay validation

Due to having no reference standard and the ethical limitations to obtaining healthy control bone marrow, the lower limits of blank, detection and quantitation were calculated using healthy PB samples, disease control bone marrow samples and samples from MM patients with abnormal plasma cells present. One of the biggest challenges in validating a flow cytometry assay is the establishment of a LLOQ. It is important to ensure no false positive or negative results are given out which might affect patient treatment. Spiking and dilution experiments determined expected events and when compared to observed events the %CVs down to a sensitivity of 0.0005% were all below 20%. When identifying rare population, a %CV of ≤30% is acceptable where the LOD is 0.01% or lower (Lee *et al.,* 2006). However, this chosen sensitivity threshold can only be useful and instructive when it is reached by a significant number of patients and further work will determine the percentage of patients reaching this degree of MRD negativity. In another study using NGS, the number of patients reaching MRD negativity at 10^{-6} was 80 out of 131 (61%) and these patients had better PFS at 3 years than those with MRD levels $>10^{-6}$ (83% vs 53%), suggesting this is an appropriate sensitivity cut off to aim for (Avet-Loiseau *et al.,* 2015). The sensitivity of this assay will decrease when fewer events are acquired and the

desired 5 million events to be acquired from each sample will not be reached 100% of the time (Table 3.3). The LLOQ has previously been calculated as (50/total cells) x 100 (Arroz *et al.,* 2016). This would give a LLOQ of 0.001% for 5 million cells collected. However, with the increased sensitivity of this assay it can reliably detect a discrete population of 25 cells and is in line with a previously standardised NGF MRD assay (Flores-Montero *et al.,* 2017).

The testing of the presence of MRD is a time critical evaluation, requiring both the use of fresh samples and for the assay to be run immediately after antibody labelling. Therefore, proof of inter-assay precision is not possible to provide for this assay. However, the small number of samples ran in triplicate for this assay when labelled at the same time produced precise results with %CV values below 20%, even at low concentrations of the cells of interest. The time critical aspect of this assay makes it more difficult to implement into a routine diagnostic laboratory and cut off times for sample collection and arrival to the laboratory would need to be implemented.

With the inclusion of a high number of cell surface markers, this assay can provide an overall assessment of the quality of the bone marrow samples being tested for the presence of MRD. Through the absence of, or significant decrease in, non-plasma cell populations the haemodilution of bone marrow samples with PB can be uncovered. It is essential to consider sample quality when producing MRD reports and the result of quality checks plus the LLOQ of the assay should be stated on each report as recommended by consensus guidelines (Arroz *et al.,* 2016).

Due to the emergence of anti-CD38 therapy as an effective treatment for MM patients the identification of plasma cells using CD38 positivity will be unusable in a certain cohort of patients. The cell surface marker CD48 has been identified as an alternative plasma cell marker (Hosen *et al.,* 2012). Use of the additional drop-in antibodies did not compromise the performance of the assay. It was important to determine not only that this marker could detect plasma cells but also that it did not affect the sample labelling by the other antibodies. Results between the two labelling methods were very comparable and it can be concluded that this antibody drop in can be used to accurately detect plasma cells in those patients undergoing anti-CD38 therapy.

3.4.3 Limitations

The validation of a bone marrow MFC/NGF assay is best performed using negative controls of the same sample type. However, it is both difficult and unethical to obtain healthy control bone marrow due to the invasive nature of sample collection. Therefore, healthy control PB was used a substitute. For the spiking experiments washed cells from bone marrow and PB were used together. Despite this mixing of sample types the results obtained were in line with previous MFC/NGF validations and therefore the assay parameters can be considered valid (Arroz *et al.,* 2016, Flores-Montero *et al.,* 2017). Small sample numbers have been used to carry out this validation due to the patient bone marrow samples being relatively low volume, unrepeatable and the high reagent costs for MFC. The use of 5 or more samples for calculation of LOB and LOD is in line with consensus recommendations (Wood *et al.,* 2013). The performance of this assay will continue to be assessed through the EQA scheme and assessment of the number of patients reaching the lower sensitivity limits of the assay. The results of these assays will also be continuously clinically validated during the long-term follow up of patients. When assessing sample quality of bone marrow samples using this assay currently only the presence or absence of cell populations can be utilised. A normal bone marrow range for these populations and therefore cut offs for haemodiluted samples cannot be produced without access to a number of healthy control bone marrow samples. However, a comparable MFC assay with similar sensitivity has produced normal ranges for bone marrow cell populations and these cut-offs, alongside assay expertise, can be used together to assess sample quality (Flores-Montero *et al.,* 2017).

This NGF assay will have the highest sensitivity when detecting ≥5 million events, with 88% of the MM patient bone marrow samples from this validation acquiring 5 million +/-5% total events; although these acquisition numbers were reached in only 38% of PB samples. From this participant recruitment centres were contacted and asked for the maximum volume of PB sample stated in the protocol (20ml) to be sent. During the course of this study the treatment details of the patient i.e. the use of anti-CD38 daratumumab, may not be available when first testing and therefore the drop-in CD48 antibody might not be added. This will be remedied by asking for current and planned treatment details upon the screening of the patient prior to the first sample collection. All samples which are found to be negative for CD38+CD138+ plasma cells will be followed up to check the anti-CD38 status. If found to be on anti-CD38 therapy these negative results cannot be accepted. The relatively low number of patients on anti-CD38 versus the cost of using an additional antibody for every patient means that the primary panel used will

be the initial 8-colour antibody panel and the CD48 drop-in will only be used once anti-CD38 therapy has been confirmed.

3.4.4 Conclusion

The presented NGF assay has been validated and standardised for the assessment of bone marrow and PB plasma cells using the CytoFlex flow cytometer. This assay has been shown to have a high sensitivity in line with other MRD assays. Reliable assessment is vital in the detection of MRD and the validity of this assay must be continuously assessed during use. The validation of this assay has allowed a standardised protocol and simple template to be created for MRD analysis in both bone marrow and PB. The assay can now be implemented into routine practice and the bone marrow MRD NGF technique will be used as the gold-standard MRD detection method.

Chapter 4 - Impact of bone marrow minimal residual disease and circulating tumour plasma cell status on survival in multiple myeloma patients

4.1 Introduction

Minimal residual disease (MRD) refers to the MM plasma cells which persist after treatment and contribute to relapse of the disease in certain patients. Widely used in clinical trials, MRD evaluation is still not a standard tool used in routine practice (Sonneveld *et al.,* 2017). The current status of MRD detection is use as a research test which has very promising clinical utility as a risk-stratification tool for therapy decisions. It is important that MRD is now looked at as a potential surrogate endpoint for PFS and/or OS which can be used to aid in clinical decisions such as ending or switching treatments. This will benefit patients by potentially limiting medical toxicity and improving quality of life, and will benefit healthcare systems by reducing the financial burden of intensive therapy. It is essential for the clinical management of patients to determine biomarkers which are capable of detecting MRD and predicting disease progression and relapse.

4.1.1 Bone marrow MRD

The assessment of MRD in the bone marrow has shown to be a useful prognostic tool in ASCT-eligible patients both prior to and post-ASCT (Rawstron *et al.,* 2013, Munshi *et al.,* 2017, Schinke *et al.,* 2017). Bone marrow MRD status has also been shown to be an independent prognostic factor in elderly transplant-ineligible patients (Paiva *et al.,* 2016). MRD positivity in the bone marrow is included in the IMWG response criteria with a minimum sensitivity of 10^{-5} to determine MRD status. This requires a validated next generation flow cytometry (NGF) approach with a minimum requirement of 5 million cells analysed. A flow cytometric MRD assay for sensitive bone marrow plasma cell detection which is in line with suggested guidelines has been introduced in Chapter 3 (Flores-Montero *et al.,* 2017). This assay has been technically validated for use and now it's clinical use and how this compares to published studies will be assessed.

4.1.2 Circulating tumour plasma cells (CTPCs)

CTPCs are abnormal MM plasma cells which have originated in the bone marrow and escaped into peripheral blood. The bone marrow NGF approach was applied to peripheral blood with the aim of detecting CTPCs at a high sensitivity (Chapter 3). This assay can be used for the detection of CTPCs in peripheral blood in line with the EuroFlow Consortium guidelines (Sanoja-Flores *et al.,* 2019). CTPCs has gained interest recently for its use in minimally invasive disease monitoring

and it also has the advantage of avoiding potential haemodilution and the clotting seen in bone marrow samples. If CTPC analysis shows comparative sensitivity and prognostic utility to bone marrow MRD detection it could be a viable blood-based alternative.

4.1.3 Aims

1. To determine the prognostic utility of MRD status as determined by the flow cytometric assay in the bone marrow of participants undergoing ASCT.
2. To determine the prognostic utility of CTPC status in both ASCT and non-ASCT patients and how this compared to bone marrow MRD evaluation.
3. To evaluate clinical and laboratory characteristics and their correlation with MRD and CTPC status.

4.2 Methods

4.2.1 Participants

Participants were recruited into 2 different groups depending on the stage of their disease and/or their treatment plan. Group 1 are ASCT eligible participants; Group 2 are non-ASCT eligible chemotherapy only participants. This is a prospective study and the investigation schedule reflects the routine evaluation of MM at the respective centres. Bone marrow aspirates and peripheral blood samples were taken during standard patient visits to clinic. See Methods Chapter 2 for further details.

4.2.2 Study measurement schedule

For Group 1 participants bone marrow MRD analysis was planned to occur at around 1 month prior to ASCT (after induction therapy) and around 3 months post ASCT. CTPC analysis was performed at the same time points and then regularly during participant follow up. For Group 2 participants bone marrow MRD analysis was planned to occur after induction chemotherapy (post-treatment sample) and prior to the commencement of maintenance therapy (see Figure 4.1).

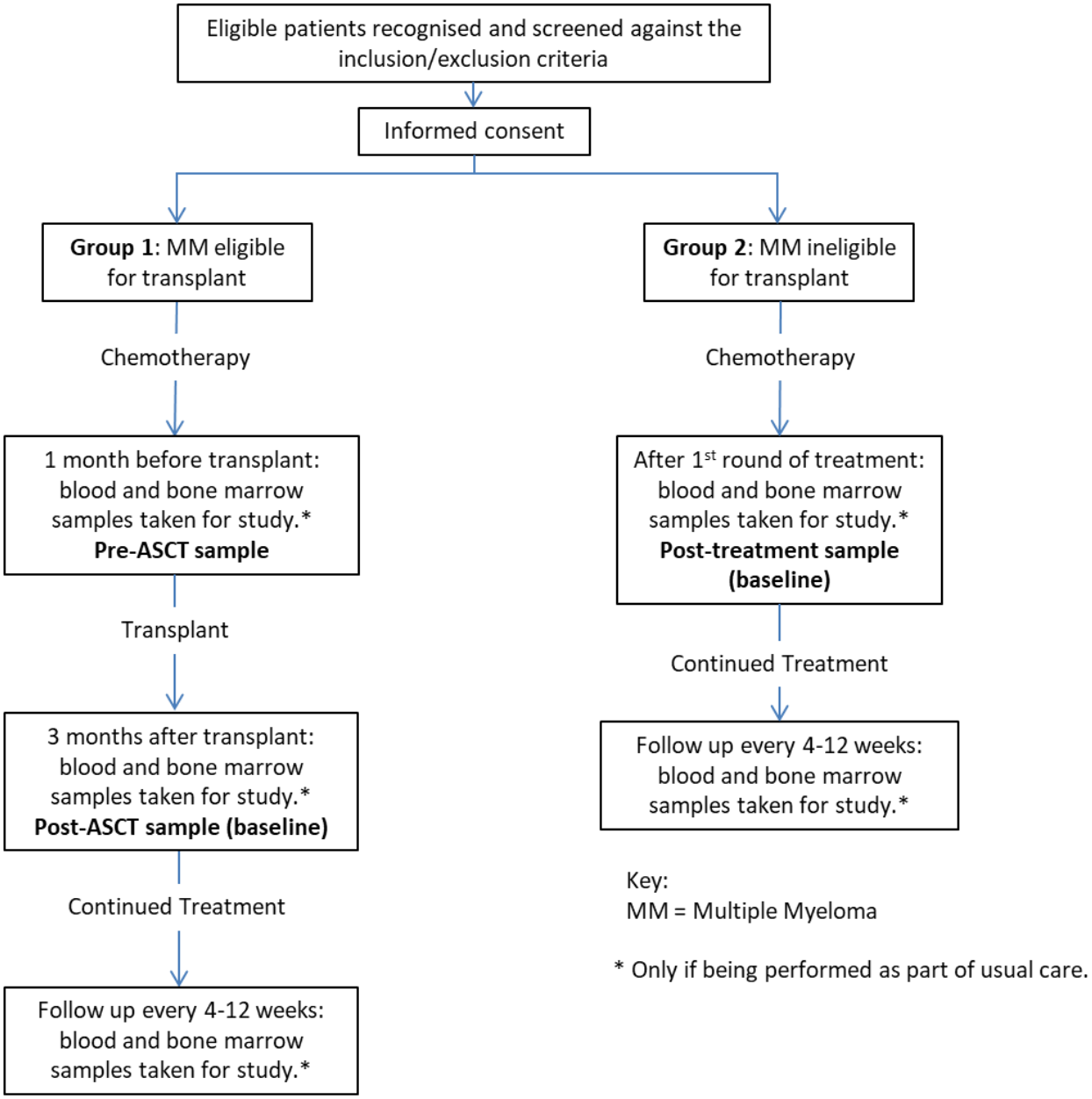


Figure 4.1: Flow diagram of the study measurement schedule.

4.2.3 MRD and CTPC analysis using NGF

Bone marrow and peripheral whole blood samples were analysed using a validated NGF technique (see Chapter 3 section 3.3.2). If less than 25 abnormal plasma cells are detected per 5 million total cell events the samples are classed as MRD negative, if more than 25 abnormal plasma cells are detected then the sample is classes as MRD positive and reported as % abnormal cells as a % of total events. Samples where the minimum of 5 million events cannot be acquired cannot be used to determine the presence or absence of MRD.

4.3 Results

4.3.1 Samples

A total of 72 bone marrow samples were received during the course of this study. 22 samples were unusable due to the following reasons: delayed arrival in the laboratory (>48 hours), incorrect timing of sample e.g. a diagnostic sample, clotted sample, haemodiluted sample, or sample volume too low to produce the required number of flow cytometry events (5 million). Of the remaining viable bone marrow samples, 28 samples were pre-ASCT and 16 samples were post-ASCT from a total of 30 Group 1 ASCT participants. 6 post-therapy samples were received from 6 Group 2 non-ASCT participants. 82% of viable bone marrow samples received had matching peripheral blood samples. For Group 1 participants the median time from the pre-ASCT sample to the date of transplant was 37 days. The median time from the date of transplant to the post-ASCT sample was 105 days. For Group 2 participants the median time from the end of induction chemotherapy to bone marrow analysis was 89 days. Regular follow up blood samples after therapy, and during either maintenance therapy or after treatment was stopped, were received from 113 participants across Group 1 and Group 2, with a median follow up time of 13 months (range 2-26 months).

4.3.2 Group 1 bone marrow MRD and CTPC analysis

4.3.2.1 Pre-ASCT bone marrow MRD

A total of 38 pre-ASCT bone marrow samples were received, with 28 producing a result due to 10 being unusable. All of the 28 pre-ASCT bone marrow samples were MRD positive by NGF. The level of residual disease in the pre-ASCT bone marrow samples varied across 5 logs with a median % abnormal plasma cells of total cells of 0.026% and a mean of 0.21% (range 0.0003% - 2.2%). As all pre-ASCT samples were MRD positive, survival curve comparison could not be completed for MRD status. Instead a survival curve was created based on grouping participants by the % abnormal plasma cells of total cells (Figure 4.2). There was no significant difference between the survival in those participants with higher amounts of bone marrow disease levels versus those with lower amounts (p=0.32). When analysing the relationship between pre-MRD levels and the post-MRD status in the same participant, no significant relationship was seen (p=0.22) although those in the bottom 50% of MRD levels did skew towards having a negative post-ASCT MRD status (Table 4.1).

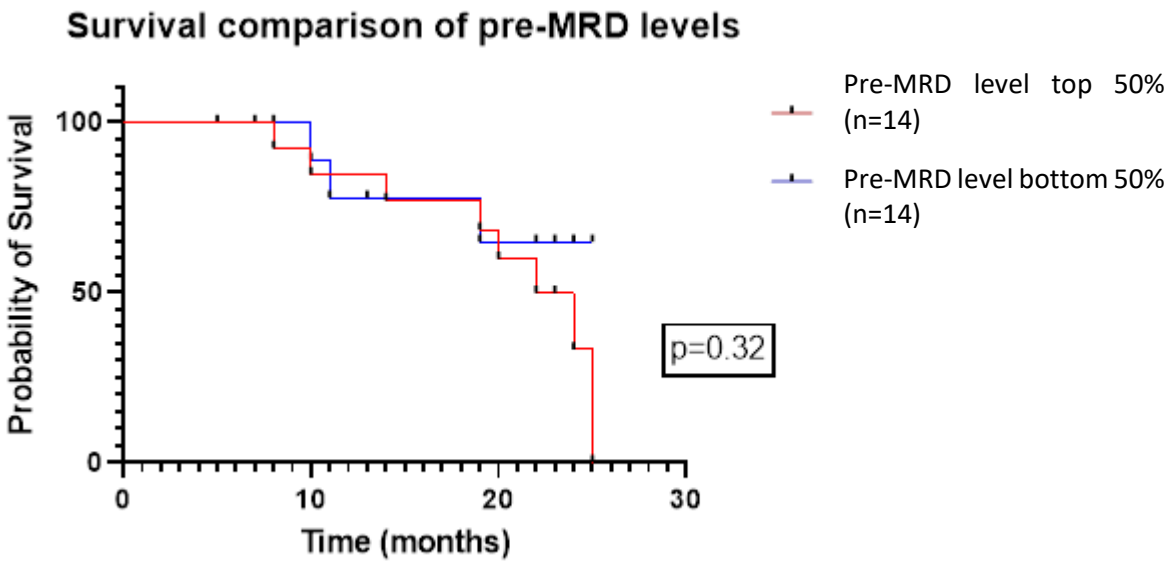


Figure 4.2: PFS curve for MRD level pre-ASCT. Positive MRD results were grouped according to % abnormal plasma cells of the total cells acquired, with the top 50% of levels (mean 0.94%, median 0.22%) vs the bottom 50% of levels (mean 0.006%, median 0.002%). (HR, 1.9; 95% CI, 0.55-6.2; p=0.32). Median survival is 22 months vs undefined.

Table 4.1: Comparison of bone marrow MRD levels at pre-ASCT and bone marrow MRD status at post-ASCT

	Pre-MRD level top 50%	**Pre-MRD level bottom 50%**
Post-ASCT Positive	3	1
Post-ASCT Negative	2	6

Fisher's exact test p=0.22

4.3.2.2 Post-ASCT bone marrow MRD

A total of 25 post-ASCT bone marrow samples were received, with 16 producing a result and 9 samples being of poor quality. Of the 16 samples, 7 were MRD positive (44%) and 9 were MRD negative (56%). Of the MRD negative samples 4 of the participants were classified as being in CR, and 3 of the MRD positive samples were classified as CR at the time of sampling using IMWG criteria based on immunofixation and bone marrow immunohistochemistry status (Kumar *et al.,* 2016). The level of residual disease in the post-ASCT bone marrow samples varied across 3 logs with a median % abnormal plasma cells of total cells of 0.0% and a mean of 0.005% (range 0.0% - 0.038%). From this small sample set survival curves were created for PFS (Figure 4.3). The presence of MRD was associated with an inferior survival outcome, although this association was not significant (p=0.051). Median survival for MRD positive participants was 24 months

versus undefined in MRD negative participants and none of the MRD negative participants relapsed during the course of the study.

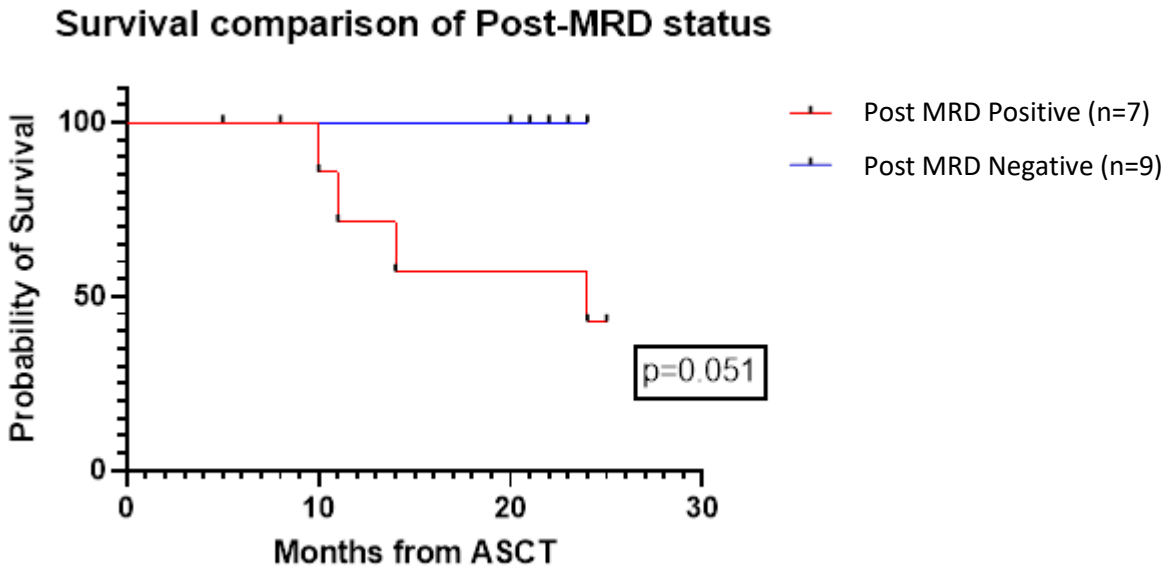


Figure 4.3: PFS curve for MRD status at day 100 post-ASCT (HR, 7.5; 95% CI, 1.0-59; p=0.051). Median survival is 24 months vs undefined.

Post-ASCT MRD status was then corelated with different clinical variables associated with MM response and outcomes including cytogenetics and IMWG response criteria at the same time of sampling (Table 4.2). No significant association was found between MRD status and any of the listed variables.

Table 4.2: Clinical and laboratory associations with bone marrow post-ASCT MRD status.

Clinical Parameter	MRD Positive (n=7)	MRD Negative (n=9)	P value
Age, years	66 (40-85)	60 (52-67)	0.2
Female Sex (%)	39%	16%	0.4
Haemoglobin (g/L)	122 (90-149)	126 (105-146)	0.5
Serum creatinine (µmol/L)	93 (40-669)	68 (48-85)	0.6
Serum calcium (mmol/L)	2.4 (2.1-2.8)	2.3 (2.2-2.4)	0.09
IgG M-protein (%)	67%	83%	0.6
Involved Ig (g/L)	12 (0.13-48)	9 (5-15)	0.5
I/U FLC ratio	52 (1-701)	3 (0.7-10)	0.4
CTPC Positive	3	0	0.06
CTPC (%)	0.01 (0-0.02)	0	0.5
CR IMWG response (n)	3	4	0.7
Cytogenetics (High risk)	4	1	0.5

Values for age, haemoglobin, creatinine, calcium, involved Ig, I/U FLC ratio and CTPC% are shown as median values and ranges.

4.3.2.3 Post-ASCT IMWG response status

In order to see if the detection of bone marrow MRD provided additional prognostic information to currently used IMWG response criteria (Table 1.5) post-ASCT survival curves were produced using IMWG criteria i.e. those in CR vs non-CR at this time point. When looking at MRD positive participants only there are not enough participants to accurately determine the effect of IMWG status on MRD positive survival (MRD positive/CR vs MRD positive/non-CR, p=0.70). When grouping all post-ASCT participants by IMWG status only, regardless of MRD status, a significant difference in survival is not seen in CR vs non-CR participants (Figure 4.4, p=0.51), underscoring the need for more sensitive measures of response. As the IMWG status reflects conventional techniques such as SPEP and IFE, with all CR patients being IFE negative, this highlights the need for more sensitive blood and serum assays to help determine disease status and be used for prognosis.

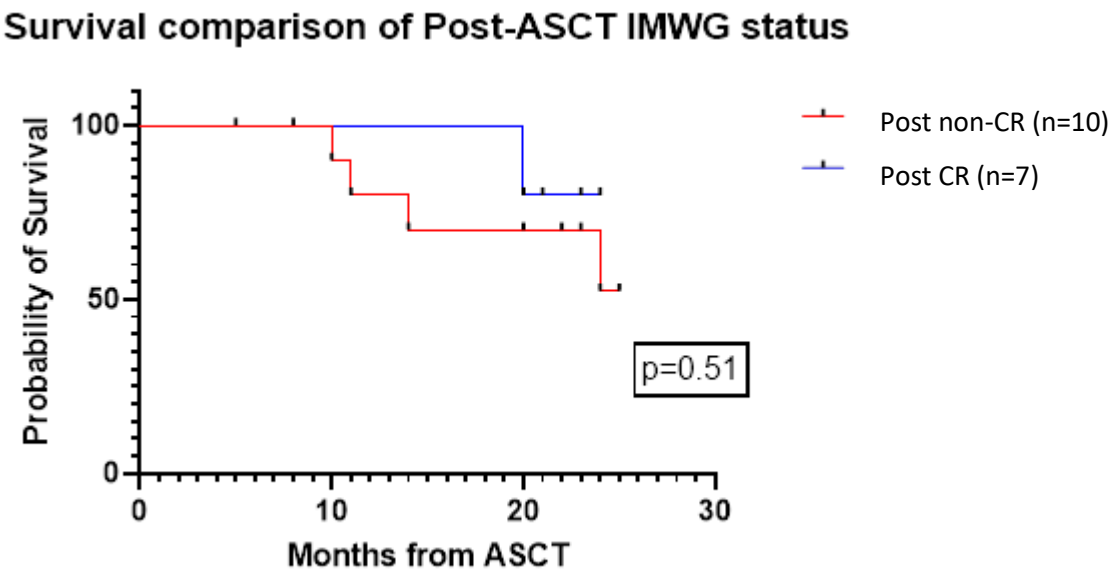


Figure 4.4: PFS curve for IMWG response status at day 100 post-ASCT (HR, 2.1; CI, 0.29-12.3; p=0.51). Median survival is undefined for both groups.

4.3.2.4 Pre-ASCT CTPC status

A total of 20 pre-ASCT peripheral whole blood samples were received. 11 samples were CTPC positive (55%) and 9 samples were CTPC negative (45%), meaning there are 9 samples which have a positive bone marrow MRD result but a negative CTPC result. The level of residual disease in the pre-ASCT CTPC samples varied across 4 logs with a median % abnormal plasma cells of total cells of 0.001% and a mean of 0.002% (range 0.0% - 0.13%).

When survival was assessed for pre-ASCT CTPC negative vs pre-ASCT CTPC positive no significant difference was found between the survival of the two CTPC statuses, though the median survival was 22 months vs undefined respectively and none of the CTPC negative patients relapsed during the course of the study (p=0.072) (See Figure 4.5).

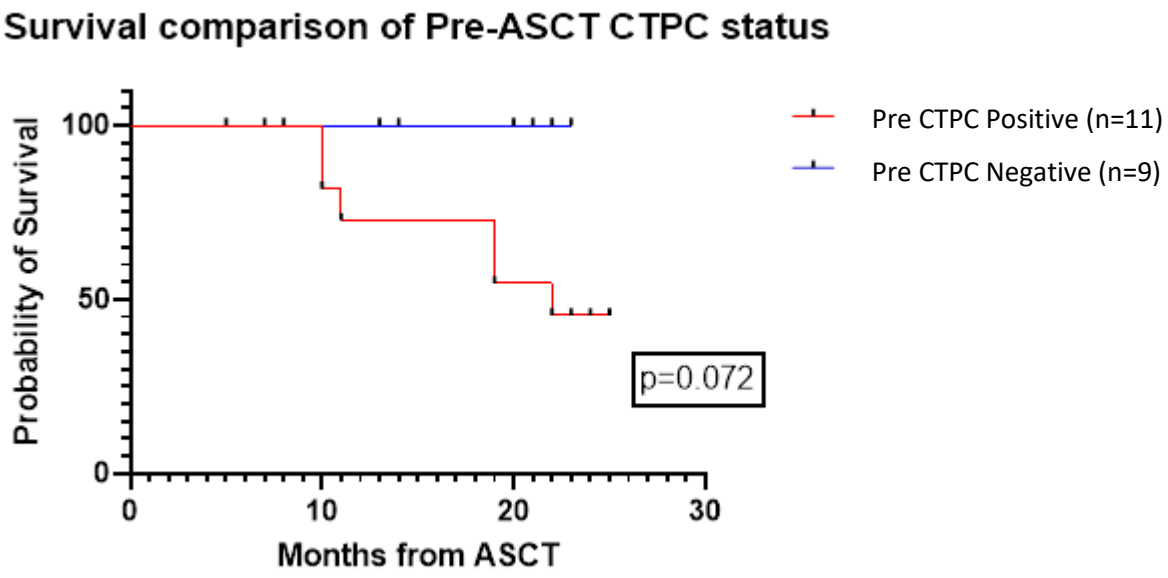


Figure 4.5: PFS curve for CTPC status pre-ASCT (HR, 4.9; 95% CI, 0.9-29; p=0.072). Median survival is 22 months vs undefined.

4.3.2.5 Post-ASCT CTPC status

A total of 16 post-ASCT peripheral whole blood samples were received. 4 samples were CTPC positive (25%) and 12 samples were CTPC negative (75%). All positive CTPC samples with a matching bone marrow sample were also MRD positive. 3 bone marrow MRD positive samples had a matching CTPC negative sample. All MRD negative bone marrow samples were also CTPC negative (Table 4.3). The level of residual disease in the post-ASCT CTPC samples varied across 3 logs with a mean % abnormal plasma cells of total cells of 0.002% (range 0.0% - 0.02%).

Table 4.3: Comparison of bone marrow MRD status and CTPC status post-ASCT

	MRD Positive	**MRD Negative**
CTPC Positive	4	0
CTPC Negative	3	9
		81% concordance

When analysing PFS for the CTPC positive participants versus CTPC negative, no significant difference was found (p=0.68), with median survival being undefined for both groups (Figure 4.6). All the CTPC negative participants who experienced progression had an accompanying bone marrow MRD positive result.

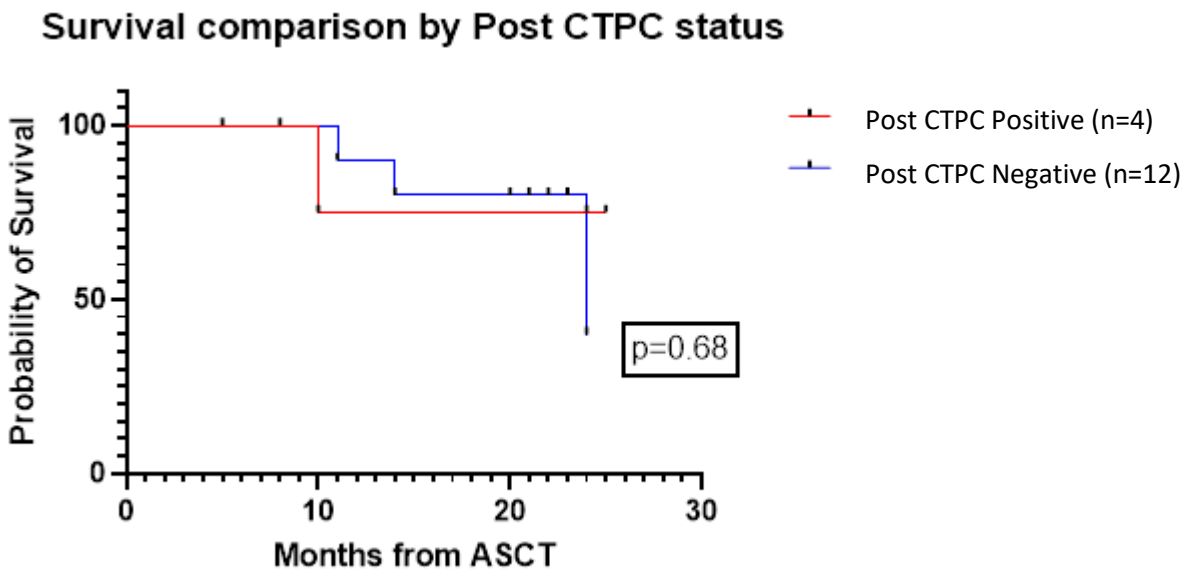


Figure 4.6: PFS survival of post-ASCT CTPC positive vs CTPC negative participants (HR, 0.63; 95% CI, 0.07-5.62; p=0.68). Median survival is undefined for both groups.

To determine whether CTPC status can add prognostic value to an MRD result the group were split according to MRD and CTPC status, looking at those where both were positive, both were negative or a positive MRD result had an accompanying negative CTPC result (Figure 4.7). Interestingly, despite these small numbers there was a significant survival difference between the three groups (p=0.02) and it was those who were MRD positive with CTPC negative who experienced the highest rate of progression. However, when directly comparing the MRD positive with CTPC positive group and the MRD positive with CTPC negative group there was no significant difference in survival (p=0.27). This suggests that it is the MRD status of the bone marrow, and not the CTPC status which has a significant effect on survival. Looking further into the participants in the MRD positive with CTPC negative results who relapsed they consisted of two IgG kappa M-protein and one IgA lambda M-protein participants with a range of % abnormal plasma cells in the bone marrow from 0.008 – 0.04%. There were no defining qualities of these participants which appeared to make them more likely to relapse than MRD positive plus CTPC positive participants.

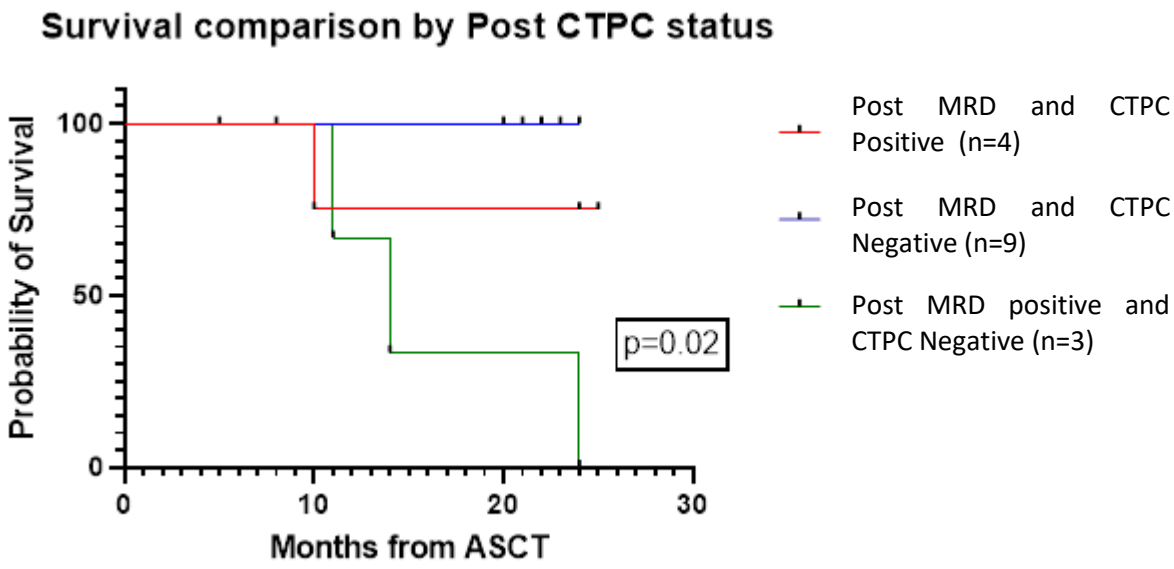


Figure 4.7: PFS survival of post-ASCT MRD+CTPC positive vs MRD+CTPC negative vs MRD positive+CTPC negative (p=0.02)

4.3.3 Group 2 bone marrow MRD and CTPC analysis

Only 6 usable bone marrow samples were received from Group 2 participants post-therapy. 4 of these samples were MRD positive and 2 were MRD negative. Of the 4 MRD positive participants, 2 have relapsed at 10 and 14 months post-therapy. Neither of the MRD negative participants have relapsed. There are not enough data within this group to perform any statistical analysis. All Group 2 post-therapy CTPC samples were negative.

4.3.4 CTPC analysis of groups 1 and 2 at best response

Finally, CTPC status at the time of best response was analysed as determined by IMWG criteria. The best response sample is that where the patient reached their maximal response as defined by IMWG criteria at any point during follow up. Survival analysis was then performed from this time point onwards. CTPC analysis was performed on 113 different participants and assessed their PFS (Figure 4.8). CTPC status at best response had a significant effect on PFS (p=<0.0001). From a prognostic point of view, presence or absence of CTPC appears to be a prognostic marker for progression when looking at the best response time point.

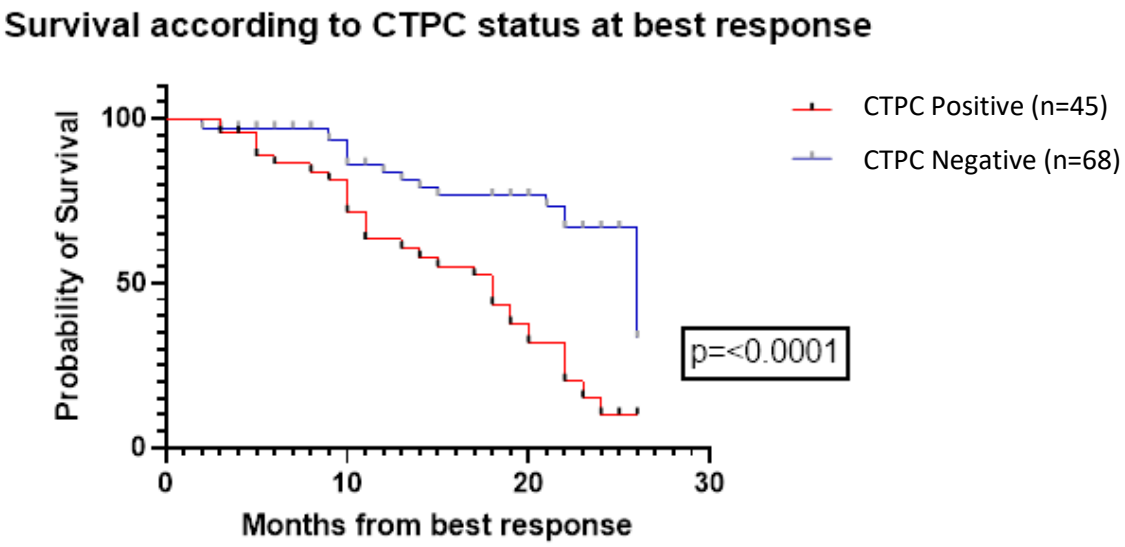


Figure 4.8: PFS of participants post-treatment when CTPC is assessed at time of best response (HR, 3.4; 95% CI, 1.9-6.1; p=<0.0001). Median survival is 18 months vs 26 months.

Table 4.4 gives the clinical and laboratory characteristics of these participants at the time of best response, sub-grouped by the presence or absence of CTPCs. Several parameters were significantly different between the two groups of participants. The presence of CTPCs at best response in the cohort is associated with a lower haemoglobin level and higher serum calcium and involved immunoglobulin levels and more abnormal free light chain ratios. Correlating with these results, as expected, the presence or absence of CTPCs also shows a significant relationship with the IMWG definition of complete response (CR).

Table 4.4: Clinical and laboratory associations with CTPC status at best response.

Clinical Parameter	CTPC Positive (n=45)	CTPC Negative (n=68)	P value
Age, years	68 (40-91)	66 (41-88)	*0.04
Female Sex (%)	31%	35%	0.6
Haemoglobin (g/L)	124 (71-151)	126 (83-166)	*0.03
Serum creatinine (µmol/L)	75 (40-369)	74 (48-192)	0.7
Serum calcium (mmol/L)	2.4 (2.0-2.9)	2.3 (1.9-2.6)	*0.05
IgG M-protein (%)	69%	77%	0.4
Involved Ig (g/L)	13 (0.13-60)	7 (0.14-25)	***0.001
K/L FLC ratio	2.7 (0.00-700)	1.8 (0.02-122)	**0.004
I/U FLC ratio	7 (0.05-1530)	1.9 (0.4-122)	**0.005
CR IMWG response (%)	7%	46%	***<0.0001

Values for haemoglobin, creatinine, calcium, involved Ig, K/L FLC ratio and I/U FLC ratio are shown as median values and ranges.

Due to the relationship seen between CTPC status and IMWG response, and in order to determine whether the presence or absence of CTPCs at best response added value over the currently used IMWG criteria, PFS survival was evaluated in this best response group using IMWG response status. When disregarding the CTPC status and comparing those in CR vs non-CR there was a significant PFS difference between the two groups (p=0.001) and this difference remained when comparing CR and VGPR participants only (p=0.004). The addition of CTPC status to a CR response status gives a significant survival difference between CR participants with CTPC present and CR participants without CTPC (p=0.046, median survival is 19 months vs 25) (Figure 4.9). It should be noted however that out of the 34 participants who were in CR at time of best response, only 3 of these were CTPC positive, and 2 of these relapsed.

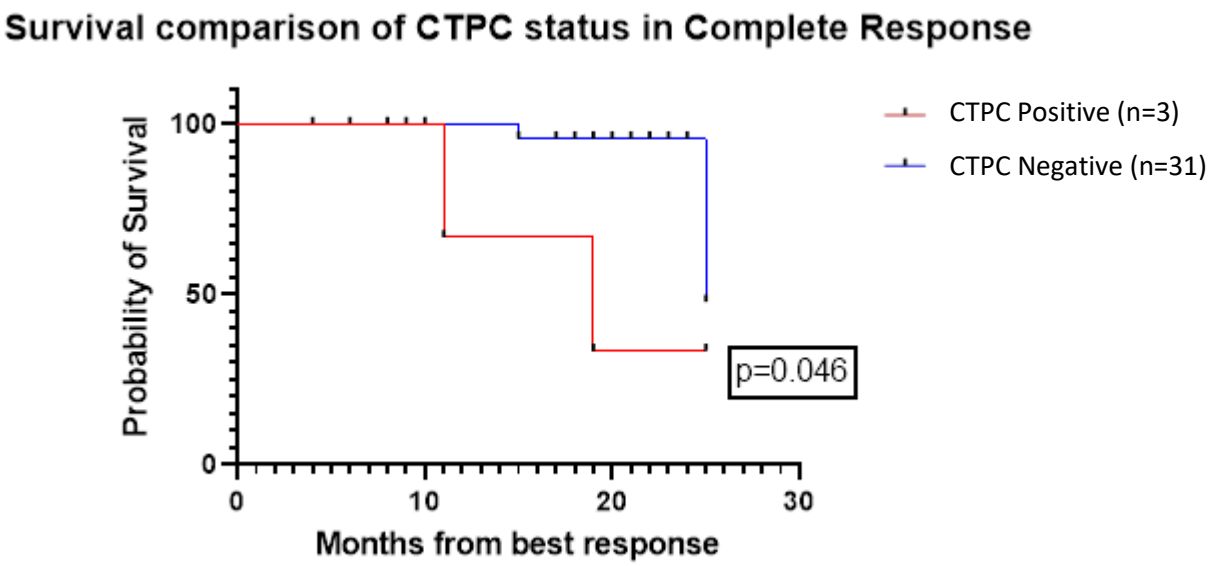


Figure 4.9: PFS of participants when CTPC is assessed in participants in complete response at time of best response (HR, 5.4; 95% CI, 0.1-9.4; p=0.046). Median survival is 19 months vs 25 months.

4.4 Discussion

The evaluation of bone marrow MRD has emerged as an important tool in response assessment and prognosis for treated MM patients. As outlined previously, MFC/NGF is the most appropriate bone marrow tool to evaluate in the current healthcare setting due to availability of equipment, skills and knowledge. With newer, more sensitive tests developed to detect MRD, different disease states past CR/sCR can now be defined. This study aimed to use a validated next-generation flow cytometry technique to evaluate the MRD status of participants who have undergone ASCT or induction therapy and determine the prognostic utility of this technique. Alongside bone marrow evaluation, the same technique was also to detect CTPCs in the same participants. It is important when determining the use of biomarkers to consider both laboratory and patient experience. The use of peripheral blood assays has many advantages over bone marrow assays and it is important to fully explore these options as an alternative to a bone marrow MRD technique.

4.4.1 Bone marrow MRD analysis

Unfortunately, due to several reasons (discussed in section 4.4.4), not enough viable bone marrow samples were received to power significant results. However, trends in the data can be identified and agreement with previous studies assessed. All of the pre-ASCT bone marrow

samples received in the study were MRD positive. These are samples taken after induction therapy but prior to ASCT. Given the sensitivity of the flow cytometry assay (10^{-5} to 10^{-6}) it is not surprising that residual MM plasma cells can still be detected. In a much larger study of 320 patients there was a pre-ASCT MRD negativity rate of 35%, increasing to 54% after ASCT (Oliva *et al.,* 2017). In this study 56% of participants became MRD negative after ASCT, matching the larger study closely. This also fits with previous reports of 42% to 58% of patients becoming MRD negative after ASCT (Rawstron *et al.,* 2002, deTute *et al.,* 2007, Rawstron *et al.,* 2013). One study determined MRD at the pre-ASCT time point and the study authors concluded that being MRD positive at this time point was associated with a non-significantly (p=0.1) inferior PFS (Rawstron *et al.,* 2013). Paiva *et al.,* (2008) found that both PFS and OS were significantly influenced by MRD status before transplant. These studies highlight how an early response to therapy i.e. MRD negativity after induction therapy and pre-ASCT gives a survival advantage and this likely reflects how the individuals disease is more responsive to therapy overall. None of the participants in this study reached MRD negativity at the pre-ASCT stage but a further look was taken into the pre-ASCT samples, splitting the MRD positive samples by disease levels. However, this did not lead to any significant survival differences and may indicate that it is just the presence of disease at this point which leads to a poorer prognosis. The level of MRD disease (% abnormal plasma cells) also did not have a significant influence on eventual post-ASCT MRD status, although 6/7 of the bottom 50% of pre-ASCT MRD levels went on to have a negative post-ASCT MRD status.

For post-ASCT status there was a trend towards longer PFS in MRD negative participants compared to MRD positive, though this was not significant (p=0.051). The 31% progression rate by 24 months follow up also closely matches previous studies (Rawstron *et al.,* 2013, Oliva *et al.,* 2017, Mohan *et al.*, 2021). The low sample numbers and relatively short follow up time of this study have made it difficult to reach the power needed to obtain a clear result. However, the result of no clinical disease progression in the MRD negative group suggests that the technique and study participants of this study match other reports of the strong association between MRD status and survival (Munshi *et al.,* 2017, Munshi *et al.,* 2020).

Despite not being significant, MRD status showed a greater prognostic potential than using IMWG criteria in the same participants (p=0.051 vs p=0.51). There was also not a significant relationship between MRD status and those participants classified as having a complete response or any of the other clinical parameters used in routine MM evaluation. In this study 44% of MRD negative participants had reached CR at 100 days post-ASCT, which correlates well

with other recent MRD studies by NGF and ASO-PCR (Huhn *et al.,* 2017, Medina *et al.,* 2020). This study by Medina *et al.* (2020) also found no significant associations between MRD status and laboratory parameters except for CR status. Previous studies have shown that response criteria do not always correlate well with MRD status; Rawstron *et al.,* (2013) found that 26% of 246 patients who were MRD-negative by 6 colour flow cytometry did not reach CR by standard response criteria. 12% of the MRD negative patients did not even reach VGPR. Interestingly, those MRD negative patients who did not reach CR had a similar PFS and OS as MRD positive patients. A similar outcome was found in another study showing that those achieving CR has longer PFS than those achieving less than CR but this was found to be specifically due to MRD status, with MRD positive CR patients having similar survival outcomes to MRD positive VGPR and PR patients (Lahuerta *et al.,* 2017). This current study is not large enough to be able to sub-group the MRD negative participants according to IMWG response status but a longer follow up of these participants may show if being CR alongside MRD negativity has a survival advantage. Previous studies have confirmed that MRD status is a stronger predictor for progression and/or survival than the CR status (Paiva *et al.,* 2008, Rawstron *et al.,* 2013). This supports an objective for therapy to be aiming towards MRD negativity, rather than a CR response. The data in this chapter, and the non-significance of using the IMWG criteria at post-ASCT to predict PFS in the cohort, highlights the need to investigate other sensitive methods of disease detection.

4.4.2 CTPC analysis at pre- and post-ASCT

CTPC analysis has already shown use in MGUS, SMM and newly diagnosed MM patients. As CTPC numbers are relatively low compared to bone marrow plasma cells, the sensitive NGF assay designed for bone marrow evaluation is the perfect tool to also sensitively evaluate peripheral blood. In this part of the study the suitability of CTPC quantification as a surrogate marker for MRD in the bone marrow was tested.

At the pre-ASCT time point 45% of samples were CTPC negative. This immediately suggests that the CTPC assay is not as sensitive as the bone marrow MRD assay, with all bone marrow samples being positive at pre-ASCT. Despite the survival analysis not showing significance between the CTPC positive and negative groups, there was an even split between the groups (11 and 9 samples respectively) and none of the CTPC negative participants relapsed. This therefore implies that there does exist a role for CTPC analysis at this time point. All CTPC positive samples at pre-ASCT had a corresponding positive bone marrow sample, suggesting that when CTPC is positive there will be disease present in the bone marrow. This information could be used at this

important therapy time point to indicate when a bone marrow test can be avoided as there is certain to be disease present in the bone marrow.

When evaluating day 100 post-ASCT samples specifically, 4/7 (57%) participants who were positive for bone marrow MRD were negative for the presence of CTPCs. This is a high rate of discordance between the two assays, however it is lower than a larger study which found only a third of MRD positive patients had matching CTPC positive peripheral blood samples (Sanoja-Flores *et al.,* 2019) and is close to a more recent study showing that 40% of patients displaying MRD in the bone marrow were negative in the peripheral blood (Burgos *et al.,* 2020). There was also no significant difference in PFS survival between CTPC positive and CTPC negative participants when evaluated at day 100 post-ASCT. These findings again indicate that CTPC analysis is a less sensitive MRD marker in treated MM than bone marrow MRD analysis. However, it is possible that the CTPC result is telling us something different than the bone marrow MRD result. For example, it could indicate the ability of MM plasma cells to spread to extra-medullary locations which will have an effect on the clinical outcomes of the patient.

Despite CTPC status alone not having prognostic utility, it is important to determine whether it could add value to the MRD bone marrow result. When comparing these groups (Figure 4.7) there was significant survival difference between all 3 groups, however when comparing the two MRD positive groups there was not a significant survival difference and therefore the addition of CTPC positivity on top of bone marrow MRD positivity did not have an effect on progression free survival. This highlights that bone marrow MRD status is the best predictor of survival here and MRD positivity would overcome any benefit from being CTPC negative. All participants who were CTPC positive were also bone marrow MRD positive. This suggests that although CTPC analysis cannot be a surrogate marker for bone marrow negativity, it has the potential to be able to show the persistence of MRD in the bone marrow. Therefore, rather than being a marker of good response it may have a use as a predictor of persistent disease and longer-term poor prognosis.

4.4.3 CTPC analysis at best response

The final analysis was of CTPC status at the time of best response in the treated participants, as determined by IMWG criteria. It was shown that CTPC status had a significant effect on progression free survival. Interestingly, 40% of the participants were CTPC positive at best response and had a significant association with IMWG response i.e. CTPC negative participants were significantly more likely to be classified as CR over CTPC positive participants. The reason

for this high level of CTPC positivity could be due to the study population. The majority of participants included in this part of the analysis are Group 2, non-ASCT participants. These participants are often older, as shown by the significant age difference between the two CTPC status groups (Table 4.4), and undergo milder treatment (Pulte *et al.,* 2011, Rosko *et al.,* 2017). These participants are therefore less likely to reach the deepest levels of response. The relatively short follow up time for some of these participants (minimum of 2 months) could also mean that these participants have the possibility of reaching CR at a later time point with further maintenance treatment.

The finding that 7% of participants with CTPCs present were in CR is lower than a larger study using NGF to detect CTPC in treated MM patients which found that 17% of CTPC positive patients were classified as CR/sCR (Sanoja-Flores *et al.,* 2019), but is higher than earlier studies using less sensitive flow cytometry (Gonsalves *et al.,* 2014, Chakraborty *et al.,* 2016). The finding that 54% of CTPC negative participants still had detectable M-protein by SPEP (i.e. not in CR) could be due to the prolonged half-life of Ig (~23 days) and the time for complete clearance of immunoglobulin (~29 weeks). The use of monoclonal antibody therapy (e.g. daratumumab) could also contribute to the presence of an M-protein on SPEP, however these participants were excluded from IMWG analysis if the origin of the M-protein on SPEP could not be determined.

CTPC status was well correlated with several other disease markers, showing significant associations with both clinical disease markers such as haemoglobin and calcium and other markers of disease activity such as involved immunoglobulin level and free light chain ratios. Due to these associations it was assessed whether the addition of CTPC testing is useful when using current response criteria or whether it is merely a reflection of what is already tested. When comparing the survival of CR participants by CTPC status, a significant survival advantage was shown for those that are both in CR and CTPC negative, over those who are CTPC positive. This suggests that CTPC analysis can help determine a further depth of response past CR and improve response assessment and prognosis of progression. However, CTPC was only positive in a very small number of CR patients (3/34, 9%) and therefore this is only likely to be useful in a small subset of patients.

4.4.4 Limitations

This study was limited by low sample numbers. There are several reasons for this including the shielding of MM patients and cancelling of bone marrow transplants during the course of this

study due to the COVID-19 pandemic, the receipt of haemodiluted and clotted blood samples and the receipt of samples too old to be analysed. Poor sample quality is a known limitation in bone marrow studies, especially when the study sample is not prioritised as the first draw. With larger sample numbers MRD positive samples could be evaluated more closely and participants sub-grouped by level of MRD in order to see the impact on survival times. Rawstron *et al.,* (2016) assessed 397 patients taking part in a clinical trial and found that the log level of MRD correlated with both PFS and OS. The follow up time period is relatively short, with larger studies only showing significant relapse rates at 5-10 years post treatment, and MRD negative patients having PFS times of even longer (Mateos *et al.,* 2010, Paiva *et al.,* 2016, Sanoja-Flores *et al.,* 2019).

4.4.5 Conclusion

These results are based on real-world MM patients who are not part of a specific clinical trial and who adhere to routine follow-up procedures. This study shows a trend towards the negative impact that the presence of MRD at the post-ASCT time point has on survival, as has been shown in multiple larger studies (Munshi *et al.,* 2020). However, it could not show that peripheral blood analysis was better, or even as good as, bone marrow analysis at this time point. It does however show the negative prognostic impact of CTPCs at best response on PFS and how this can be used to further sub-group patients classified as being in complete response, although this would only be useful in a small subset of patients. Although CTPC detection in the peripheral blood is less sensitive than bone marrow MRD detection it might still be useful as an independent tool to predict the presence of residual disease and/or impending progression. The presence of CTPCs is a promising, non-invasive biomarker for disease, which can be useful in determining disease activity in the bone marrow until CTPC is negative. Therefore it cannot stand alone as a marker for MRD status, but a positive CTPC result can indicate the presence of disease in the bone marrow. A larger cohort will be needed to assess this in a prospective study but it is important to continue to look for peripheral blood alternatives to bone marrow evaluation.

Chapter 5 - Immunoglobulin heavy/light chain analysis in the measurement of monoclonal proteins in multiple myeloma patients

5.1 Introduction

MM is characterised by the proliferation of clonal plasma cells in the bone marrow, usually leading to the secretion of a monoclonal intact immunoglobulin (M-protein) and/or free light chains (FLC) into the serum. In MM the M-protein acts as a biomarker which is used in diagnosis, prognosis and monitoring of the disease. Due to the complex nature of MM and the M-protein, no single laboratory test is used for diagnosis and monitoring; instead, a testing strategy must be defined by the laboratory to effectively detect this disease in line with current IMWG guidelines (Kumar *et al.,* 2016). In the era of novel agents and clinical trials, these traditional techniques are becoming incapable of monitoring patients achieving deeper responses due to their limited sensitivity and usability (Willrich *et al.,* 2016). A newer technique to measure the M-protein is the analysis of immunoglobulin heavy/light chains (HLC) using the Hevylite™ assay (The Binding Site, UK). Both the recent IMWG response and MRD assessment criteria and the NICE MM guidelines (Kumar *et al.,* 2016, NICE 2016) recommend further investigation into the potential role for HLC in deep response evaluation, particularly in those patients achieving CR post-treatment.

Previous studies have shown a use for HLC analysis in diagnosis of MM (Lopez-Anglada *et al.,* 2018), monitoring treated patients (Kastritis *et al.,* 2014, Batinic *et al.,* 2015) and more recently the detection of MRD (Michallet *et al.,* 2018, Miyazaki and Suzuki 2018). As the HLC assay has not been established in current guidelines it is important that the test characteristics are defined and how it compares to currently used methods for M-protein detection is assessed. This will help to determine whether MM patients would benefit from adding the HLC assay into routine disease assessment and whether there is a role for this assay in the detection of early disease relapse and minimal residual disease detection.

5.1.1 Aims

1. To assess the accuracy and sensitivity of the HLC assay in comparison to currently used methods for immunoglobulin and M-protein measurement.
2. To determine the use of the HLC assay in detecting relapse in treated MM patients.
3. To determine the prognostic utility of HLC measurements at best response and MRD timepoints.

4. To assess the relationship between BM NGF measurements and the HLC measurements at MRD timepoints.

5.2 Methods

5.2.1 Participants

84 consecutively recruited participants with confirmed MM were prospectively followed between April 2019 and May 2021: 61 with IgG and 23 with IgA M-protein isotype. All participants had been previously treated and were at different treatment stages and response categories during recruitment and follow up. The median follow up time was 18 months (range 3–26 months). Participants were recruited into 2 different groups depending on the stage of their disease and/or their treatment plan. Group 1 are ASCT eligible patients; Group 2 are non-ASCT eligible chemotherapy only patients. This was a prospective study and the investigation schedule reflects the routine evaluation of MM at the respective centres. Bone marrow aspirates and peripheral blood samples were taken during standard patient visits to clinic. See Methods Chapter 2, section 2.3 for further details.

5.2.2 Sample analysis

The tests performed on each samples routinely included total immunoglobulins (G, A and M), SPEP, IFE to confirm negative SPEP and FLC analysis. HLC analysis was additionally carrier out on all samples. See Methods Chapter 2, sections 2.6.2 and 2.6.3 for further details on sample analysis.

5.3 Results

A total of 216 different samples from 84 individual participants were tested. Participants were followed for a median of 18 months from the date the baseline sample was obtained. The characteristics of the participant population are shown in Table 5.1.

Table 5.1: Baseline participant characteristics

	All Participants	IgG Participants	IgA Participants
Number of Participants	84	61	23
Number of Samples	216	154	62
Age (years)	67 (37-88)	66 (37-85)	78 (48-88)
Sex (male/female)	54/30	41/20	13/10
Total involved Ig g/l (standard method)	10.4 (0.13-63.6)	12 (2-63.6)	1.21 (0.13-40.7)
sHLC g/l (HLC method)	7.8 (0.02-63.6)	9.13 (0.02-63.6)	0.96 (1.63-47.7)
I/U HLCr	7.15 (0.53-3741.9)	7.47 (0.53-463.9)	6.37 (0.65-3741.9)
FLCr (k/l)	1.66 (0.001-700.9)	1.98 (0.01-700.9)	1.02 (0.001-165.6)
Haemoglobin g/l	125 (71-166)	127 (93-166)	121 (71-156)
Serum Calcium mmol/l	2.34 (1.97-2.88)	2.33 (1.99-2.83)	2.38 (1.97-2.88)
Serum Creatinine µmol/l	73 (36-227)	75 (42-189)	69 (36-227)
CR (number of samples)	58	29	29
VGPR (number of samples)	53	43	10
PR/MR/PD (number of samples)	78/24/3	62/19/1	16/5/2

Values for measurements are median values with ranges shown in parenthesis.

5.3.1 Accuracy of HLC test

As the HLC test is not currently widely used it is important to initially ensure its accuracy in quantifying M-protein and total immunoglobulin compared to the gold standard methods of total immunoglobulin measurement (nephelometry or turbidimetry). The correlation between the total immunoglobulin of the involved isotype (IgG or IgA) and the summated HLC of the corresponding isotype (i.e. iHLC+uHLC) of 216 samples from 84 different participants was firstly assessed. Correlation analysis showed good correlation (r^2=0.85, p<0.0001) between summated HLC and total immunoglobulin across IgG (154/216) and IgA (62/216) samples (Figure 5.1).

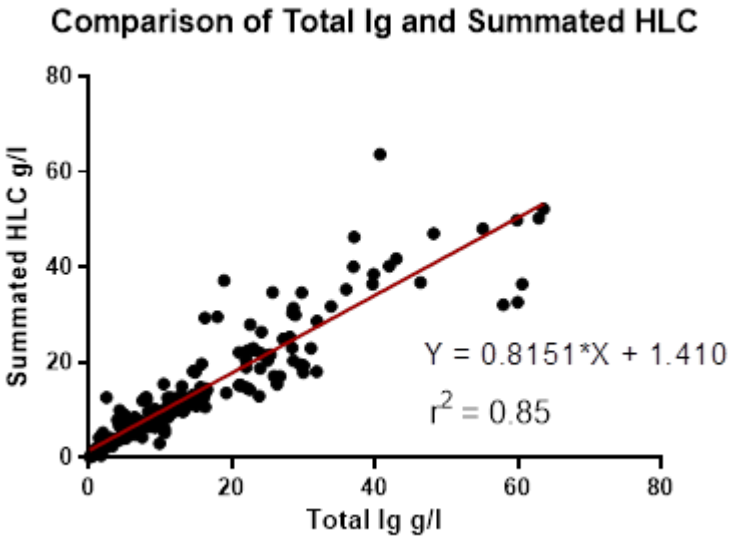


Figure 5.1: Correlation analysis between summated HLC and total immunoglobulin across all isotypes.

Analysis showed a good correlation between summated HLC and total immunoglobulin. When analysed separately IgA samples showed a slightly stronger correlation (r^2=0.90) than IgG samples (r^2=0.89). The Pearson correlation for all samples is r = 0.92 with p< 0.0001. This suggests that the correlation is significant.

In order to further define agreement between the two measurement methods a Bland-Altman analysis was carried out (Figure 5.2).

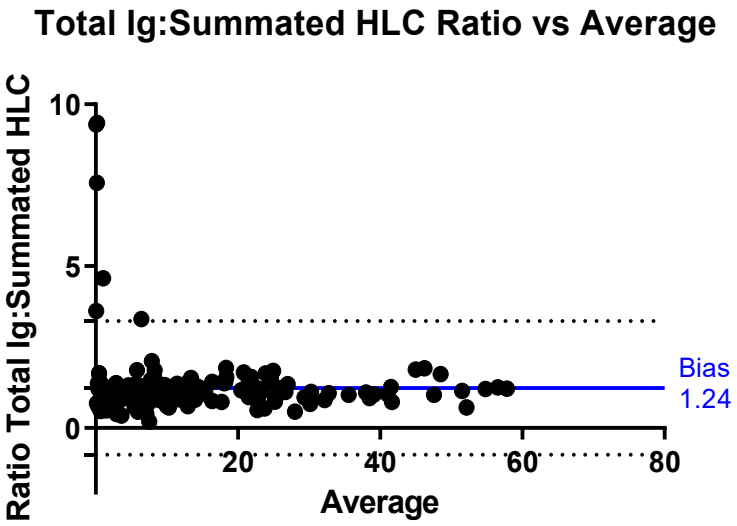


Figure 5.2: Bland-Altman correlation and agreement for total Ig for all isotypes vs summated HLC for all isotypes.

Bias (blue line) and 95% limits of agreement (dotted lines) are shown.

The bias of 1.24 is constant across the measuring range except for those at the lowest levels below 0.1. There is a slight systemic difference between the methods, with total immunoglobulins measured by turbidimetry showing a consistently higher value than summated HLC values.

Bland-Altman analysis identified a systemic bias of 1.24 with 95% limits of agreement of -0.83 – 3.31. This bias seen is likely due to the different analysers and reagents used to measure total immunoglobulin and HLCs. Although both analysers are turbidimeters, they use different reagents with different specificities to detect the target. Taking these analyses together, total Ig and summated HLC display satisfactory overall agreement, however to ensure the results would not be biased by an underestimation by the HLC assay the next compared measurements were of the M-protein itself.

The next analysis compared M-protein measurement using SPEP measurement (which utilises total protein values and densitometry) and iHLC. M-protein was measurable in 80 IgG M-protein samples and 11 IgA M-protein samples (Figure 5.3). Correlation analysis showed moderate agreement between the two methods (r^2=0.81, p<0.0001), with IgA M-protein samples showing better correlation than IgG M-protein samples (r^2=0.95 and 0.85 respectively), although far fewer IgA M-proteins are quantifiable by SPEP due to migrating to the beta region. When assessing by Bland-Altman analysis there was a slight positive bias (0.90) for M-protein

quantification by SPE over iHLC but this is not considered to be a significant bias which will affect the measurements of the M-protein (Figure 5.4).

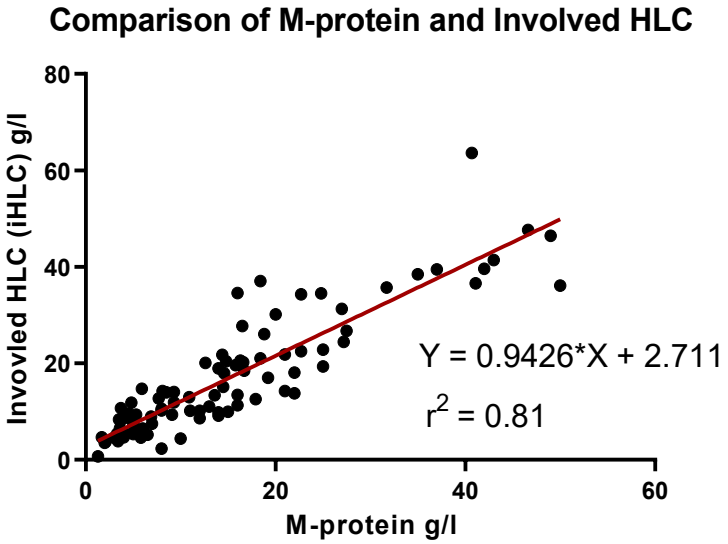


Figure 5.3: Correlation analysis between involved HLC and M-protein measurement by SPEP across all isotypes.

Analysis showed a moderate correlation between involved HLC and M-protein SPEP measurement. When analysed separately IgA samples showed a stronger correlation (r^2=0.95) than IgG samples (r^2=0.85). The Pearson correlation for all samples is r = 0.90 with p< 0.0001. This suggests that the correlation is significant.

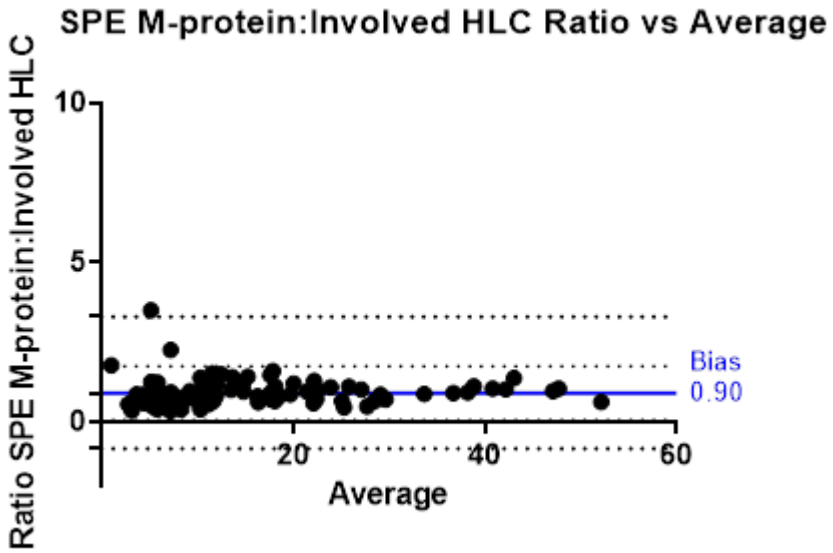


Figure 5.4: Bland-Altman correlation and agreement for SPEP M-protein vs involved HLC for all isotypes.

Bias (blue line) and 95% limits of agreement (dotted lines) are shown.

The bias of 0.90 is constant across the measuring range. There is a slight systemic difference between the methods, with M-protein measurement by SPEP showing a higher value than involved HLC values.

5.3.2 Ability of HLC to detect monoclonal protein

The ability of the HLC variables to show the presence of an M-protein compared to conventional techniques (total immunoglobulin, SPEP and FLC analysis) used in monoclonal gammopathy screening was assessed next. Table 5.2 summarises the detection of the M-protein by different methods in 216 different samples from 84 individual participants (154 IgG M-protein samples and 62 IgA M-protein samples. Table 5.2 displays the number of samples deemed abnormal by the different methods or calculations. Of the conventional methods, total immunoglobulin was the least sensitive for detecting M-protein with FLCr showing the highest proportion of abnormal results for both IgG and IgA M-protein samples. HLCr was abnormal in a high number of samples despite a normal iHLC. This led to the assessment of the level of abnormality in uHLC values, as in the case of a normal iHLC it is likely this uHLC value which contributes to the abnormal HLCr. uHLC showed the highest proportion of abnormal results across all the M-protein testing which is interesting as currently in routine practice iHLC and HLCr values are more commonly used in decision making.

Table 5.2: Summary of conventional and HLC methods for detecting M-protein

Method	IgG M-protein (n=154)	IgA M-protein (n=62)	Total (n=216)
Total Ig Abnormal	56 (36%)	23 (37%)	79 (37%)
SPE Abnormal	103 (67%)	30 (48%)	133 (62%)
FLCr Abnormal	122 (79%)	36 (58%)	158 (73%)
iHLC Abnormal	76 (49%)	24 (39%)	100 (46%)
HLCr Abnormal	113 (73%)	41 (66%)	154 (71%)
uHLC Abnormal	117 (80%)	55 (89%)	172 (80%)

5.3.3 Concordance of iHLC/HLCr with conventional methods

The concordance between the conventional methods and HLC was then analysed.

5.3.3.1 Total Ig vs HLC

As a total Ig of the involved M-protein isotype above the normal range is being defined as abnormal, the total Ig will be compared to both iHLC and HLCr values as these are measurements of the abnormality of the involved M-protein. Both an increased total Ig and iHLC above normal ranges indicate monoclonality, as does a skewed HLCr. Although uHLC showed a high level of abnormality, a low level of uHLC does not directly indicate monoclonality and therefore uHLC will not be used in this comparison of methods. All samples with abnormal total Ig levels also had abnormal iHLC (Table 5.3). However, in IgG M-protein samples 57/98 (58%) of those with normal total Ig values had either an abnormal iHLC or abnormal ratio (or both), signifying that the M-protein was still present despite the total Ig being within normal range. In IgA M-protein samples all samples with abnormal total Ig levels also had abnormal iHLC (Table 5.3). However, 18/39 (46%) of those with normal total Ig values had either an abnormal iHLC or abnormal ratio (or both), signifying that the M-protein was still present. IgA M-protein samples had a higher concordance (71%) than IgG M-protein samples (63%), giving an overall concordance between total Ig and iHLC/HLCr of 65%.

Table 5.3: Concordance between total Ig and iHLC/HLCr samples for normal or abnormal results

IgG M-protein samples	HLC (HLCr or iHLC) Normal	HLC (HLCr or iHLC) Abnormal
Total Ig Normal	41	57
Total Ig Abnormal	0	56
IgA M-protein samples	**HLC (HLCr or iHLC) Normal**	**HLC (HLCr or iHLC) Abnormal**
Total Ig Normal	21	18
Total Ig Abnormal	0	23

5.3.3.2 SPEP vs HLC

SPEP showed an abnormal result in 103/154 (67%) IgG M-protein samples compared to the 73% of samples abnormal according to iHLC and/or HLCr. Of the 103 samples abnormal by SPEP the M-protein could not be quantified in 23 samples (22%). The majority of these unquantifiable M-proteins were in the beta region (17/23) but other reasons include the presence of high background polyclonality (1), the proximity of 2 IgG M-proteins to each other in a bi-clonal patient (1) and the M-protein being too small to accurately quantify on SPEP (4). 3 samples from 3 individual participants had an abnormal SPEP result but normal iHLC/HLCr values (Table 5.4). 1 participant out of these 3 experienced disease relapse/progressions during the course of this study, interestingly the sample for this participant had an abnormal uHLC value alongside the normal iHLC/HLCr whilst the other 2 samples had normal uHLC values. SPEP showed an abnormal result in 30/62 (48%) IgA M-protein samples compared to the 66% of samples abnormal according to iHLC and/or HLCr. Of the 30 abnormal SPEP samples the M-protein could not be quantified in 19 samples (63%). The majority of these unquantifiable M-proteins were in the beta region (17/23) but other reasons include the presence of high background polyclonality (1) and the M-protein being too small to accurately quantify on SPEP (5). All IgA participants with abnormal SPEP also had abnormal HLC values. IgG M-protein samples had a higher concordance (90%) than IgA M-protein samples (84%), giving an overall concordance between SPEP and iHLC/HLCr of 88%, which is higher than with total Ig.

Table 5.4: Concordance between SPEP and iHLC/HLCr samples for normal or abnormal results

IgG M-protein samples	HLC (HLCr or iHLC) Normal	HLC (HLCr or iHLC) Abnormal
SPEP Normal	38	13
SPEP Abnormal	3	100
IgA M-protein samples	**HLC (HLCr or iHLC) Normal**	**HLC (HLCr or iHLC) Abnormal**
SPEP Normal	22	10
SPEP Abnormal	0	30

5.3.3.3 HLC vs FLC

The FLC kappa/lambda ratio (FLCr) results complemented the HLCr results but in some cases were discordant (Table 5.5). Concordance between HLCr and FLCr for IgG M-protein samples was 82% with a lower concordance of 63% for IgA M-protein samples. This gave an overall concordance between HLCr and FLCr of 76%. 25 of the mismatched samples were from 17 individual participants in CR. Follow up of these patients for the duration of the study showed that none of the 13 participants with an abnormal FLCr but normal HLCr relapsed during the course of the study. 1 out of 4 participants with a normal FLCr but abnormal HLCr relapsed during the study.

Table 5.5: HLCr vs FLCr results

	IgG samples	IgA samples	IgG + IgA samples
HLCr + FLCr Normal	23	12	35 (16%)
HLCr + FLCr Abnormal	103	27	130 (60%)
HLCr Abnormal + FLCr Normal	9	14	23 (11%)
HLCr Normal + FLCr Abnormal	19	9	28 (13%)

5.3.4 Relationship between HLC measurements and clinical response status

Clinical response status is used to help standardise the assessment of treatment response across clinical trials and health services. When assessing HLC values across the different clinical responses IgG and IgA M-protein samples were compared separately due to large differences in normal and abnormal concentration values. HLC values (iHLC, uHLC, dHLC, I/U HLCr) were correlated with the corresponding clinical response status at the time of sampling (Figure 5.5 for IgG patients).

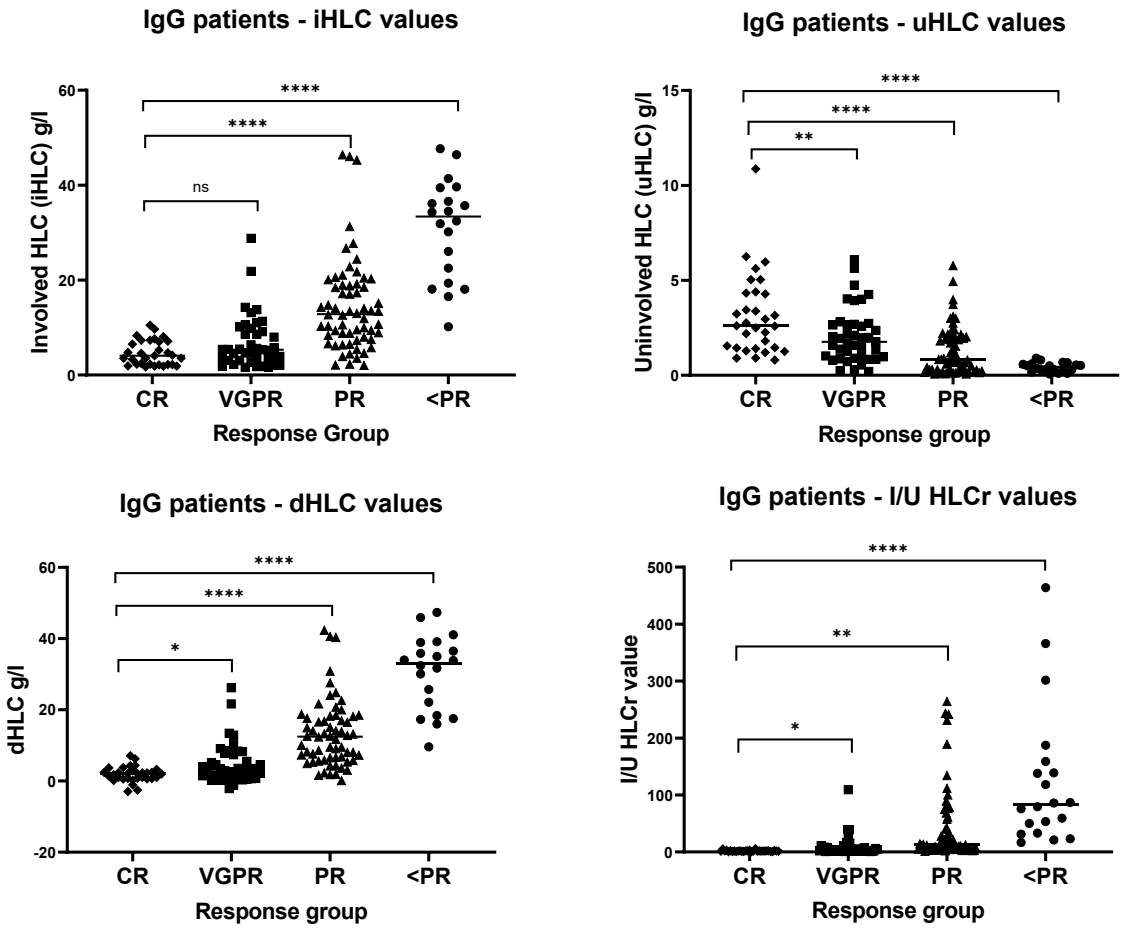


Figure 5.5: Clinical response status of IgG participants compared with iHLC, uHLC, dHLC and I/U HLCr values at the time of sampling. Horizontal lines represent median values. Significance between clinical response groups in shown. I/U HLCr is being analysed here rather than HLCr as the direction of abnormality it is not affected by M-protein isotype.

Apart from iHLC values between CR and VGPR participants there is a significant difference in HLC values between those participants in CR and those in non-CR (p=<0.0001 for CR vs non-CR for all HLC measurements). These results are reflected in IgA M-protein samples also (data not shown). As expected, iHLC, dHLC and I/U HLCr values are higher in those with a higher disease burden (non-CR) and uHLC values are lower in non-CR patients.

Figure 5.5 shows that HLC values follow conventional M-protein measurements in distinguishing between different clinical response status. As response status takes into account M-protein size, detection by SPEP and detection by IFE, with CR patients being IFE negative, it can be seen that the HLC values follow the pattern of the conventional methods. However, the uHLC values in the CR group especially are widely spread. To look further into the measurement of HLC across

the clinical responses, the percentage of samples within each clinical response group which had normal or abnormal HLC results were compared, excluding dHLC and I/U HLCr values which do not have normal ranges and have previously been used by calculating percentage increases and decreases (Michallet *et al.,* 2018) (Figure 5.6).

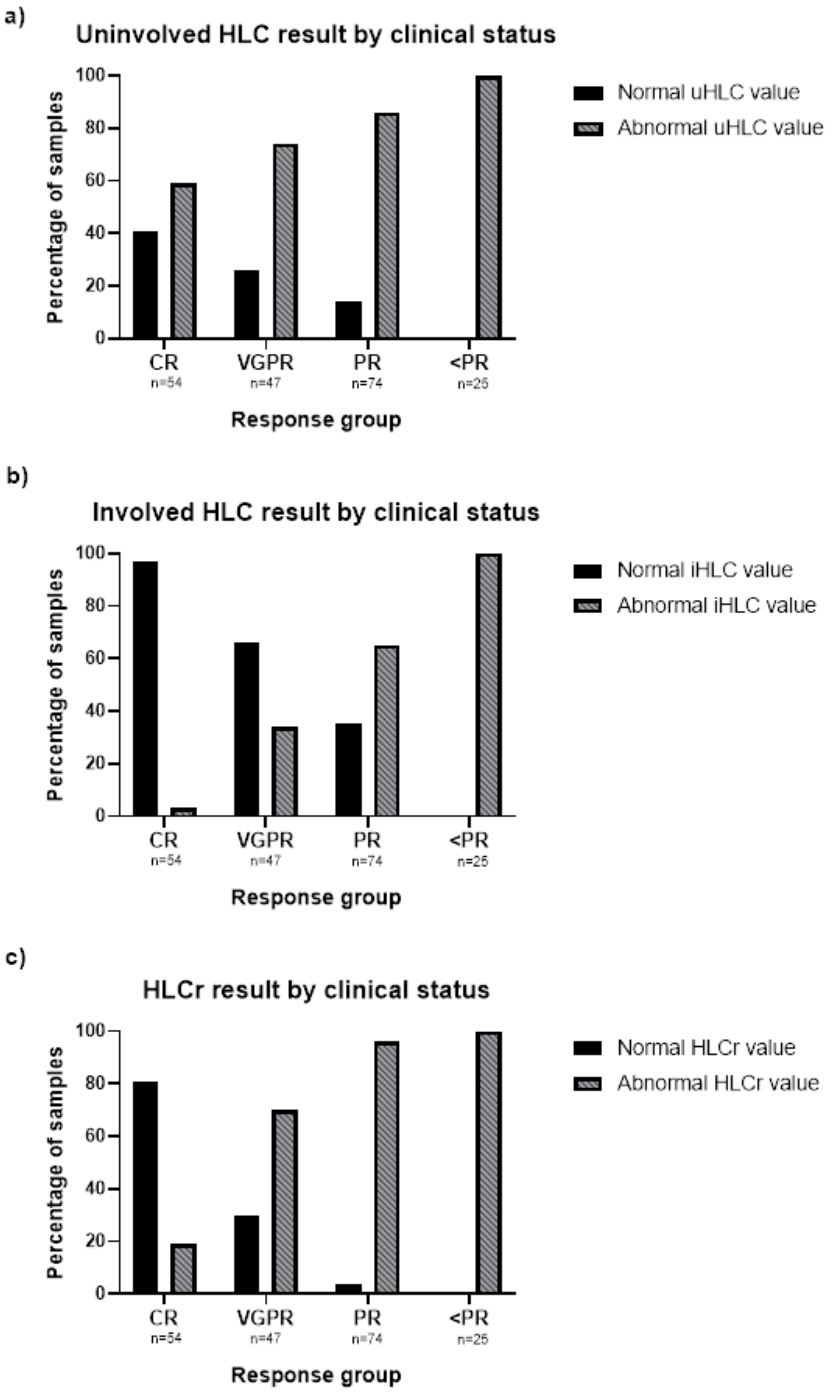


Figure 5.6: Percentage of samples in each clinical response status group with normal or abnormal a) uHLC b) iHLC c) HLCr Fisher's exact test p=<0.0001 for all values comparing CR vs non-CR.

Figure 5.6 shows the percentage of samples with either normal or abnormal HLC values according to the corresponding clinical response status at the time of sampling. The percentage of samples with normal uHLC, iHLC and HLCr levels increased with greater depth of response, with those in <PR having only abnormal HLC values. Interestingly, whilst the vast majority of CR and VGPR samples have normal iHLC, this is in contrast to uHLC values where the majority of samples in each clinical status, including CR and VGPR, have reduced (i.e. abnormal) levels of uHLC. This suggests that those participants currently grouped as CR may be able to be further subdivided according to uHLC level and this could have a prognostic value in the monitoring and treatment of these patients.

5.3.5 HLC values at best response vs relapse

To assess whether the HLC assay has a role in the early detection of relapse, those participants who relapsed during the course of this study were analysed. Best response samples were obtained either after induction therapy (for non-ASCT participants) or after ASCT. A paired comparison of M-protein and disease activity markers was performed at best response versus prior to clinically confirmed relapse (median time of pre-relapse sample prior to confirmed relapse is 57 days) (IgG M-protein participants Table 5.6, IgA M-protein participants Table 5.7). To ensure any differences in values was due to the relapse of disease a paired comparison of best response versus final sample values was performed for those participants who remained in CR throughout the course of the study. Due to low study numbers this comparison of CR participants was only possible with IgG M-protein patients (Table 5.6).

Table 5.6: Disease markers at best response and pre-relapse for relapsed IgG M-protein participants and best response and final timepoint for non-progressing IgG M-protein participants.

IgG Participants (n)				
Relapsed Participants (17)	**Normal reference ranges**	**Best response Samples (median +range)**	**Pre-relapse Samples (median +range)**	**P value**
Total IgG (g/l)	6-16 g/l	11.60 (3.17-57.94)	15 (2.86-62.93)	0.42
IFE status*	Negative	35%	11%	0.12
HLCr	0.98-2.75	4.08 (0.01-187.5)	7.07 (0.007-463.9)	0.19
I/U HLCr	N/A	4.52 (1.11-187.5)	13.29 (1.18-463.9)	0.08
iHLC (g/l)	IgGκ: 4.03-9.78 IgGλ: 1.97-5.71	6.57 (1.7-39.46)	9.97 (1.63-45.26)	0.36

uHLC (g/l)	See above	1.44 (0.17-6.08)	0.6 (0.07-4.94)	0.027 *
dHLC (g/l)	N/A	5.18 (0.19-38.94)	8.56 (0.25-41.09)	0.35
FLCr	See Campbell *et al.,* 2020	3.62 (0.43-174)	6.55 (0.03-67.21)	0.17
iFLC (mg/l)	See above	34.11 (6.44-704)	45.13 (7.19-732.5)	0.99
uFLC (mg/l)	See above	6.3 (1-30.2)	7.03 (1-57.23)	0.46
I/U FLCr	N/A	3.05 (0.1-174)	9.32 (0.04-67.21)	0.24
Hb (g/l)	115-180	121 (93-154)	125 (95-149)	0.89
Calcium (mmol/l)	2.2-2.6	2.3 (2.12-2.46)	2.33 (2.10-2.49)	0.06
Creatinine (umol/l)	49-104	68 (52-105)	70 (45-149)	0.41
Retained Complete Response (7)		**Best response Samples (median +range)**	**Final Samples (median +range)**	**P value**
Total IgG (g/l)	6-16 g/l	6.99 (2.92-14.6)	7.58 (3.23-13.9)	0.33
IFE status*	Negative	100%	100%	>0.99
HLCr	0.98-2.75	1.96 (1.59-2.81)	1.82 (1.33-1.90)	0.16
I/U HLCr	N/A	1.96 (0.63-2.81)	1.53 (0.53-1.90)	0.15
iHLC (g/l)	IgGκ: 4.03-9.78 IgGλ: 1.97-5.71	3.54 (1.86-9.68)	4.17 (1.88-7.35)	0.43
uHLC (g/l)	See above	2.96 (0.9-5.62)	2.59 (1.25-6.25)	0.18
dHLC (g/l)	N/A	2.1 (-1.1-6.24)	1.98 (-2.9-2.3)	0.25
FLCr	See Campbell *et al.,* 2020	1.29 (0.02-1.8)	1.2 (0.52-1.5)	0.69
iFLC (mg/l)	See above	13.8 (4.12-67.75)	16.44 (6.3-52.03)	0.44
uFLC (mg/l)	See above	10.11 (2.6-25.26)	10.98 (1.39-26.91)	0.81
I/U FLCr	N/A	1.42 (0.77-4.18)	1.39 (0.83-1.93)	0.26
Hb (g/l)	115-180	140 (116-153)	137 (118-166)	0.76
Calcium (mmol/l)	2.2-2.6	2.31 (2.2-2.4)	2.25 (2.16-2.42)	0.12
Creatinine (umol/l)	49-104	81 (54-121)	76 (59-123)	0.54

* For categorical variables the % of samples which are negative is stated.

Table 5.7: Disease markers at best response and pre-relapse for relapsed IgA M-protein participants.

IgA Participants (n)				
Relapsed Participants (10)	**Normal reference ranges**	**Best response Samples (median +range)**	**Pre-relapse Samples (median +range)**	**P value**
Total IgA (g/l)	0.8-3.0	1.17 (0.15-28.86)	2.66 (0.28-40.70)	0.50
IFE status*	Negative	80%	30%	0.07
HLCr	0.91-2.42	2.0 (0.01-133.5)	9.54 (0.001-311.9)	0.09
I/U HLCr	N/A	2.52 (0.67-140.6)	108 (0.94-3742)	0.27
iHLC (g/l)	IgAκ: 0.59-2.98 IgAλ: 0.43-2.04	0.84 (0.02-29.65)	4.08 (0.36-63.61)	0.39
uHLC (g/l)	See above	0.25 (0.02-1.64)	0.034 (0.01-0.19)	0.043 *
dHLC (g/l)	N/A	0.48 (0.001-29.41)	4.06 (0.54-63.60)	0.39
FLCr	See Campbell *et al.,* 2020	1.07 (0.01-8.45)	1.3 (0.01-135)	0.33
iFLC (mg/l)	See above	21.77 (0.6-163)	50.87 (1.85-308)	0.40
uFLC (mg/l)	See above	10.32 (0.5-15.5)	8.5 (0.6-104)	0.37
I/U FLCr	N/A	1.97 (0.7-111.8)	2.52 (0.02-1530)	0.38
Hb (g/l)	115-180	119 (93-139)	122 (71-134)	0.82
Calcium (mmol/l)	2.2-2.6	2.36 (2.08-2.55)	2.37 (2.14-2.88)	0.24
Creatinine (umol/l)	49-104	74 (49-169)	84 (50-174)	0.18

* For categorical variables the % of samples which are negative is stated.

For those IgG patients who remained in CR throughout the study there is no significant difference in the values of any of the M-protein measurements or disease activity markers at best response versus final samples, showing the stability of these markers in CR patients. Interestingly, for both IgG and IgA relapsed patients the only marker to show a significant difference between best response and pre-relapse samples is the uHLC value (p=0.027 for IgG, p=0.043 for IgA) suggesting that this is the most accurate factor to predict progression in treated MM patients. For other M-protein variables such as total Ig, IFE status and FLC values there was an increase in levels or abnormality between best response and pre-relapse but these were not significant.

5.3.6 Progression free survival (PFS) analysis at best response

To further assess the clinical relevance of the HLC measurements the PFS of participants grouped by their normal or abnormal HLC status at time of best response was determined (Figure 5.7). 67 participants were initially analysed due to having complete follow up data, this included patients whose clinical response was PR or better. For iHLC and HLCr there was no significant difference in PFS between those with normal and abnormal results, although HLCr did approach significance (p=0.08). There is a significant difference between the survival curves for those with normal versus abnormal uHLC (p=0.005), showing that those with an abnormal uHLC value have shorter PFS time. This remained true when restricting the uHLC analysis to those in VGPR or better (p=0.001) and, importantly, in CR only (p=0.002). iHLC and HLCr did not show a significant survival difference in these restricted response groups.

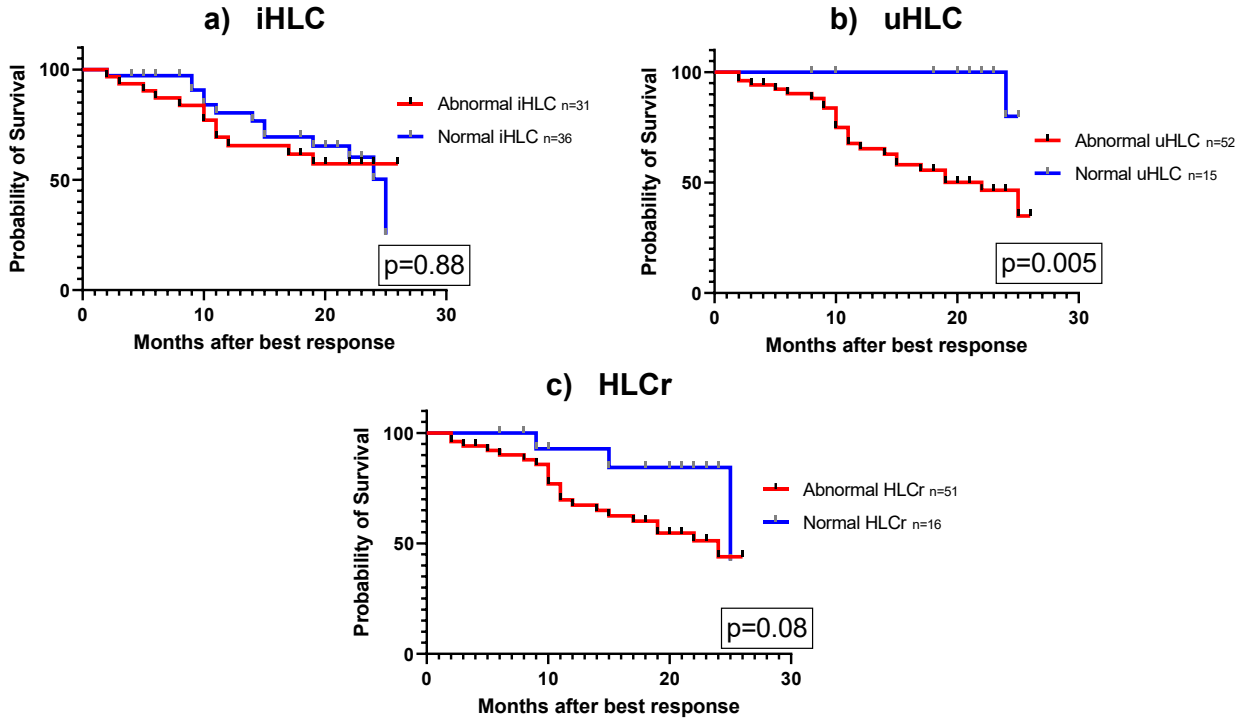


Figure 5.7: PFS in the study population at best response stratified into abnormal/normal groups according to: **a) iHLC values** (median survival is not reached vs 25 months, p=0.88) **b) uHLC values** (median survival is 22 months vs not reached, p=0.005) **c) HLCr values** (median survival is 24 months vs 25 months, p=0.08)

Due to previous reports suggesting a role for FLCr in prognosis (Rhee *et al.,* 2017, Kapoor *et al.,* 2012) its role in PFS both alone and alongside HLCr was assessed to see if it increased the utility of HLCr in survival analysis (Figure 5.8). As shown above there is some discordance between FLCr

and HLCr results with the potential for these assays to complement each other in disease assessment. The next assessment carried out was PFS in patients where only one of FLCr or HLCr needed to be abnormal to be classified as abnormal (Figure 5.8b) and where both needed to be abnormal to be classified as abnormal (Figure 5.8c). Survival analysis demonstrated that the addition of FLCr abnormality to an abnormal HLCr did not lead to statistically significant survival outcomes between the patient groups. This was true for participants in all clinical responses from PR or better but also when restricted to CR patients only. Addition of an abnormal FLCr to an abnormal uHLC did not improve the significance of uHLC in predicting PFS (p=0.03). Due to the previous significance seen using uHLC status to predict PFS it was also important to assess whether using iFLC or uFLC alone would lead to significant survival differences as this assay is already routinely in use in response assessment. However, in this cohort both iFLC and uFLC did not show a significant survival difference between those patients who had normal or abnormal values at best response (p=0.83 and p=0.14, respectively).

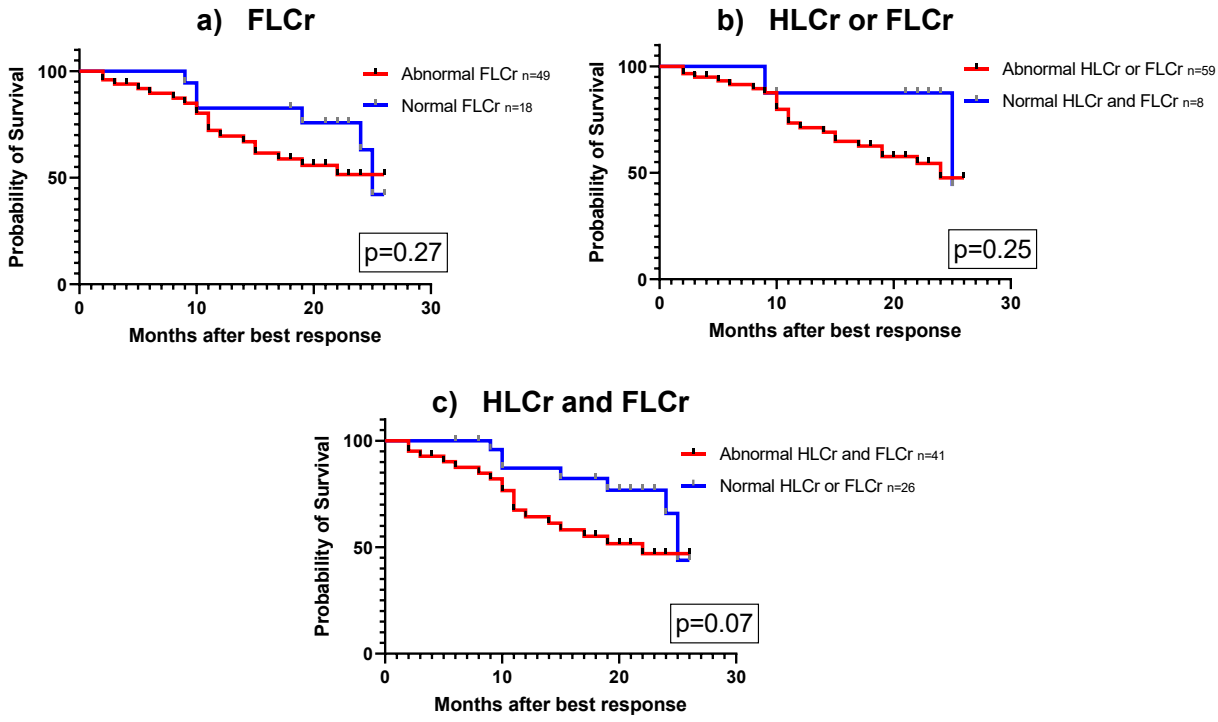


Figure 5.8: PFS in the study population at best response stratified into abnormal/normal groups according to: **a) FLCr** (median survival is not reached vs 25 months, p=0.27) **b) FLCr and/or HLCr (either) abnormal** (median survival is 24 months vs 25 months, p=0.25) **c) FLCr and HLCr (both) abnormal** (median survival is 22 months vs 25 months, p=0.07)

5.3.7 Comparison of HLC results with bone marrow NGF at MRD testing time points

The current gold standard method for measuring MRD is bone marrow MFC/ NGF. In order to determine if a serum assay could replace a bone marrow assay a comparison these tests using concordance, sensitivity and survival analyses was carried out.

An initial comparison performed was between all available bone marrow NGF results (diagnostic, pre-ASCT and post-ASCT time points) and HLC results taken at the same time points, considering agreement to be a positive bone marrow NGF result and an abnormal HLC result or a negative and normal result respectively. Out of 49 matching bone marrow/serum samples agreement between bone marrow NGF and iHLC was 51%, agreement with uHLC was 90% and agreement with HLCr was 84%. For iHLC the majority of discordance is due to a BM positive sample matching with a normal iHLC value. When considering bone marrow NGF to be the gold standard method for sensitively detecting M-protein the sensitivity and specificity of iHLC (Table 5.8), uHLC (Table 5.9) and HLCr (Table 5.10) can be seen below.

Table 5.8: iHLC - Sensitivity, specificity, PPV, NPV and LR of iHLC for detection of M-protein when BM NGF is the gold standard

Sensitivity	Specificity	PPV	NPV	Likelihood Ratio
0.45	0.78	0.90	0.24	2.03

Cohen's kappa: 0.122 SE 0.092 95% -0.058 – 0.303 = Very slight agreement

Table 5.9: uHLC - Sensitivity, specificity, PPV, NPV and LR of uHLC for detection of M-protein when BM NGF is the gold standard

Sensitivity	Specificity	PPV	NPV	Likelihood Ratio
0.93	0.78	0.95	0.70	4.16

Cohen's kappa: 0.674 SE 0.135 95% 0.409 – 0.938 = Substantial agreement

Table 5.10: HLCr - Sensitivity, specificity, PPV, NPV and LR of HLCr for detection of M-protein when BM NGF is the gold standard

Sensitivity	Specificity	PPV	NPV	Likelihood Ratio
0.80	1.00	1.00	0.53	N/A

Cohen's kappa: 0.595 SE 0.120 95% 0.360 – 0.830 = Moderate agreement

We can see that uHLC has the highest sensitivity whereas the specificity of HLCr shows that when HLCr was abnormal the bone marrow NGF was certain to be abnormal too. Cohen's kappa shows that uHLC has a substantial agreement with the bone marrow NGF results, followed by HLCr, reflecting the concordance data.

5.3.7.1 Pre-ASCT bone marrow NGF and serum HLC samples

The next analysis was restricted to HLC values at specific MRD time points. At the pre-ASCT time point there were 25 matching bone marrow and serum samples. All 25 bone marrow NGF samples were positive. The HLC values which were abnormal, and therefore in agreement with NGF, were as follows: iHLC 9/25 (36%), uHLC 24/25 (96%) and HLCr 19/25 (76%). The 1 participant who was BM NGF positive but uHLC negative was clinically classified as being in CR and did not relapse during the 23 month follow up of the study. Of the 6 mismatching HLCr participants, 1 relapsed during study follow up.

5.3.7.2 Post-ASCT bone marrow NGF and serum HLC samples

At the post-ASCT time point 16 participants had matching bone marrow and serum samples. Agreement between BM NGF and HLC values were as follows: iHLC 9/16 (56%), uHLC 13/16 (81%) and HLCr 14/16 (88%). None of the uHLC or HLCr mismatched samples relapsed during the course of the study. Table 5.11 shows the distribution of normal (BM negative) and abnormal (BM positive) samples at the post-ASCT time point. Fisher's exact test shows a significant agreement between uHLC/HLCr and BM NGF.

Table 5.11: Distribution of BM NGF and HLC post-ASCT results.

Post-ASCT	**BM NGF Status n (%)**		**P value**
Serum iHLC Status	**Negative**	**Positive**	
Normal	7	5	>0.99
Abnormal	2	2	
Serum uHLC Status			
Normal	7	1	0.04*
Abnormal	2	6	
Serum HLCr Status			
Normal	9	2	0.005**
Abnormal	0	5	

Sensitivity and specificity of the HLC variables at the post-ASCT time point are shown below, when using BM NGF as the gold standard method (Table 5.12, 5.13, 5.14). As with the samples at all bone marrow sampling time points, uHLC shows the highest sensitivity whilst HLCr has the highest specificity. Both uHLC and HLCr show substantial agreement with the bone marrow NGF method whilst iHLC cannot be used to indicate bone marrow status. As the IFE method is the current most sensitive serum method used routinely the sensitivity and specificity of this

method was calculated when using BM NGF as the gold standard method to see if the HLC variables have an advantage over this technique (Table 5.15). uHLC showed increased sensitivity and specificity over IFE, whilst HLCr showed increased specificity.

Table 5.12: iHLC - Sensitivity, specificity, PPV, NPV and LR of iHLC for detection of M-protein at post-ASCT when BM NGF is the gold standard

Sensitivity	Specificity	PPV	NPV	Likelihood Ratio
0.28	0.78	0.50	0.58	1.29

Cohen's kappa: 0.067 SE 0.231 95% -0.039 – 0.519 = Very slight agreement

Table 5.13: uHLC - Sensitivity, specificity, PPV, NPV and LR of uHLC for detection of M-protein at post-ASCT when BM NGF is the gold standard

Sensitivity	Specificity	PPV	NPV	Likelihood Ratio
0.86	0.78	0.75	0.88	3.86

Cohen's kappa: 0.625 SE 0.194 95% 0.245 – 1.00 = Substantial agreement

Table 5.14: HLCr- Sensitivity, specificity, PPV, NPV and LR of HLCr for detection of M-protein at post-ASCT when BM NGF is the gold standard

Sensitivity	Specificity	PPV	NPV	Likelihood Ratio
0.71	1.00	1.00	0.82	N/A

Cohen's kappa: 0.738 SE 0.167 95% 0.410 – 1.00 = Substantial agreement

Table 5.15: IFE - Sensitivity, specificity, PPV, NPV and LR of IFE for detection of M-protein when BM NGF is the gold standard at the post-ASCT time point.

Sensitivity	Specificity	PPV	NPV	Likelihood Ratio
0.83	0.55	0.55	0.83	1.86

Cohen's kappa: 0.359 SE 0.218 95% -0.068 – 0.786 = Fair agreement

5.3.8 Progression free survival analysis of HLC results at MRD time points

Next, survival analysis using serum samples taken at the pre-ASCT and post-ASCT time points was performed. Analysis was initially performed using those serum samples which had matching BM samples (analysed in Chapter 4, section 4.3.2) in order to provide a direct comparison to the survival outcomes of these matched patients. Survival analysis was then repeated using all available serum samples at the pre-ASCT and post-ASCT time points, regardless of whether a matching bone marrow sample was available. This gave more participants to investigate and lent more power to the analysis.

5.3.8.1 Pre-ASCT

At the pre-ASCT time point survival analysis was initially performed on 25 serum samples with matching bone marrow NGF results. All of the pre-MRD BM samples were positive and when grouping participants by MRD levels no significant difference in survival was found (see Chapter 4 Figure 4.2). Analysis of iHLC, uHLC and HLCr on the serum of these 25 participants, showed a significant survival advantage for having normal uHLC levels versus abnormal uHLC levels (p=0.04). This was not seen when comparing normal versus abnormal iHLC (p=0.57) or HLCr (p=0.37) levels.

An additional 3 serum samples without matching BM samples could be added to the analysis at this time point (Figure 5.9). When performing survival analysis on 28 pre-ASCT HLC samples there was again a significant survival difference between participants with a normal uHLC at the pre-ASCT time point versus those with an abnormal uHLC (p=0.03, median survival 21 months vs undefined). For iHLC and HLCr status there was no significant survival difference.

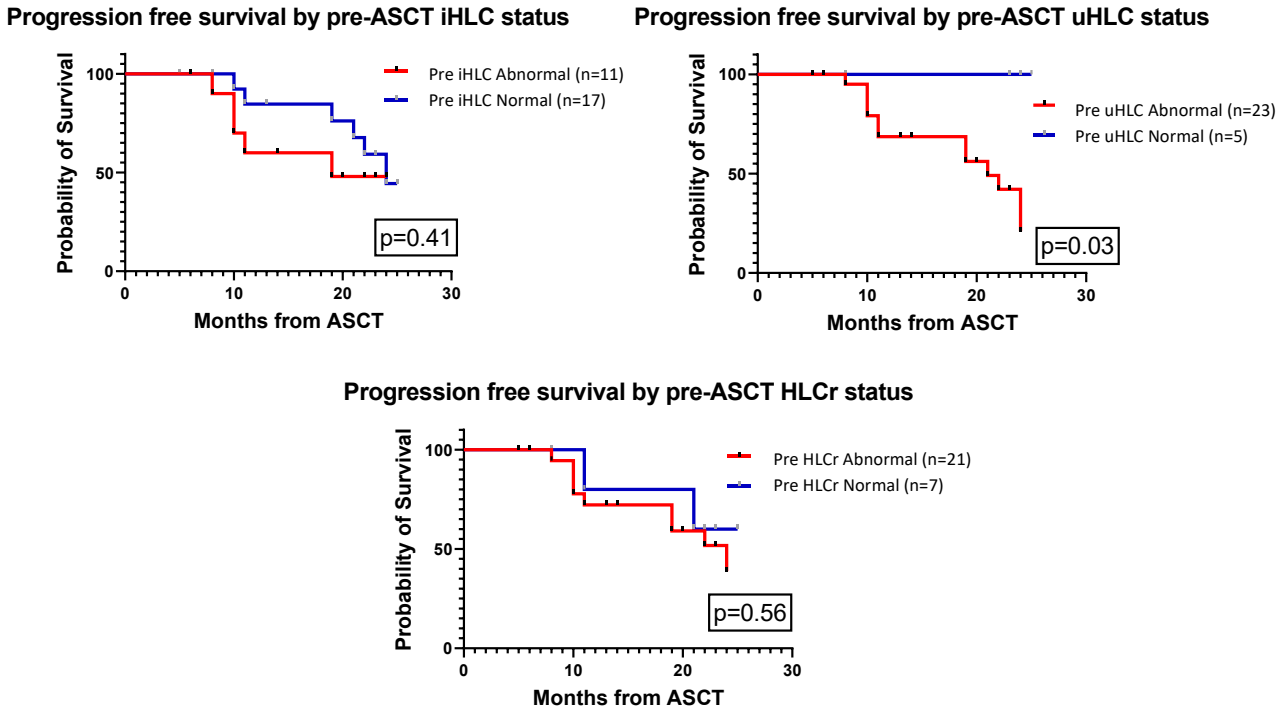


Figure 5.9: PFS of pre-ASCT HLC abnormal vs HLC normal participants for a) iHLC (HR, 1.6; 95% CI, 0.46-5.60; p=0.41) Median survival is 19 months vs 24 months. **b) uHLC** (HR, 4.7; 95% CI, 1.15-19.3; p=0.03) Median survival is 21 months vs undefined. **c) HLCr** (HR, 1.56; 95% CI, 0.40-6.02; p=0.56) Median survival is 24 months vs undefined.

5.3.8.2 Post-ASCT

Progression free survival analysis of the same 16 post-ASCT participants seen in Chapter 4 Figure 4.3, showed that having a normal uHLC or HLCr gave a significant survival advantage over those with abnormal values at this time point (p=0.04 for both variables). Interestingly, performing survival analysis on these same samples using IFE status (positive or negative) did not produce a significant PFS difference between participants (p=0.21), with IFE negative participants being the equivalent to complete responders (CR).

An additional 11 serum samples at this post-ASCT time point which did not have a matching bone marrow sample could then be added to the analysis (Figure 5.10). At this time point, iHLC was abnormal in 37% of participants, uHLC was abnormal in 52% and HLCr was abnormal in 41% of participants. There was a significant survival difference between participants with an abnormal uHLC or HLCr at the post-ASCT time point versus those with a normal uHLC or HLCr (p=0.01, p=0.02 respectively). For iHLC there was no significant survival difference. In order to thoroughly check whether there was an advantage in using HLC values in response assessment survival analysis using IFE status and FLC values was also performed. The iFLC (p=0.52), uFLC (p=0.23) and FLCr (p=0.72) values at this time point were all not predictive of PFS. In these same samples there was also no difference in PFS between IFE positive and IFE negative (equivalent to CR) participants (p=0.44) (Figure 5.11).

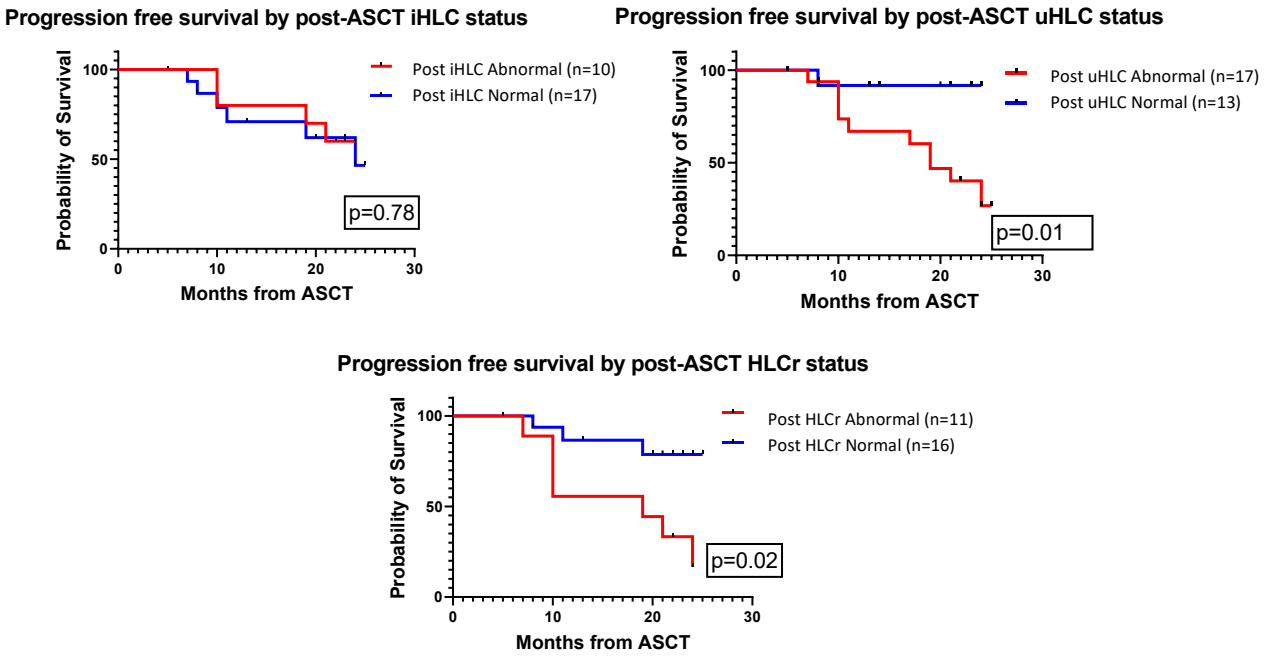


Figure 5.10: PFS of post-ASCT HLC abnormal vs HLC normal participants for a) iHLC (HR, 0.83; 95% CI, 0.24-2.92; p=0.78) Median survival is undefined vs 24 months. **b) uHLC** (HR, 7.67; 95% CI, 2.32-25.26; p=0.01) Median survival is 21 months vs undefined. **c) HLCr** (HR, 4.43; 95% CI, 1.20-16.35; p=0.02) Median survival is 19 months vs undefined.

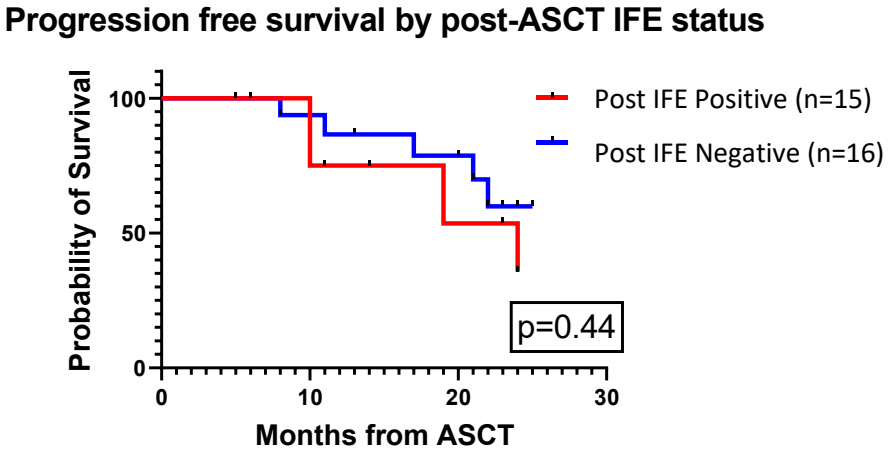


Figure 5.11: PFS survival of post-ASCT IFE positive vs IFE negative patients.
(HR, 1.6; 95% CI, 0.48-5.2; p=0.44). Median survival is 24 months vs undefined.

5.4 Discussion

Accurate quantification of the M-protein is very important for monitoring and response analysis in treated MM patients. Despite the evidence of more sensitive next generation sequencing and flow cytometry assays for disease detection, these are still not widely available to the majority of laboratories and require a bone marrow sample which is not optimal for patients. Therefore, serum measurement of the M-protein remains the most accessible and important measurement of disease response in MM. These measurements are used not only to define the presence of the M-protein but also to determine immune reconstitution in participants undergoing chemotherapy and/or ASCT.

5.4.1 HLC vs conventional methods

Traditionally, M-protein is quantified and characterised with SPEP and IFE respectively. However limitations to this include lack of sensitivity and a relative subjectivity and lack of standardisation to quantification. An advantage of HLC measurement over total immunoglobulin is the ability to detect whether an increase in immunoglobulin is clonal or reactive by determining the concentrations of iHLC and uHLC and how they interact (HLCr and I/U HLCr).

When comparing the accuracy of the HLC method in M-protein quantification to conventional total immunoglobulin and SPEP measures there was moderate to good correlation between methods. Previous studies have also reported a good to moderate agreement between summated HLC and total Ig results for MM participants (Paolini *et al.,* 2015, Ting *et al.,* 2019). Comparing iHLC values to M-protein values determined by SPEP densitometry showed slightly poorer correlation. IgA M-proteins had a better correlation compared to IgG which is in contrast to results found by Ting *et al.* (2019). However, Ting *et al.* (2019) included M-proteins which had been quantified in the beta region and therefore the quantification would have been less accurate due to the presence of other beta region proteins. The quantification of M-proteins by SPEP can be subjective and will always include small amounts of other serum proteins, as shown by the positive bias between M-protein quantification by SPEP and iHLC. When comparing HLC and SPEP measurements, it highlighted the need for alternative M-protein quantification especially in IgA patients, with 77% being unquantifiable by SPEP. All intact M-proteins can be quantified by HLC, showing the advantage of using this method in this subgroup of patients in particular. Currently the IMWG criteria recommends the use of total IgA for monitoring IgA MM patients after therapy, however these results suggest the use of iHLC and HLCr would be more sensitive and specific in detecting a monoclonal versus polyclonal process. Across all samples an

abnormal HLCr was found in 71% of samples. This compares well with Ting *et al.* (2019) who found abnormal HLCr values in 75% of treated participants. The higher proportion of patients with abnormal HLC values versus SPEP/total immunoglobulin values demonstrates the importance of using the HLC assay not only in those participants with beta region migrating M-protein but also potentially in participants who do not have a detectable M-protein by conventional techniques. An added benefit of using the HLC assay to detect M-proteins is that it can overcome interference problems from therapeutic antibodies such as daratumumab. These antibodies can cause false positives on SPEP and IFE (McCudden *et al.,* 2016) and therefore may affect the determination of CR and early relapse, however this has not been investigated this further here. FLC and HLC values were complementary but not always concordant, with 24% of samples showing discordance between the assays. This compares well with previous studies showing 29% and 25% discordance between FLCr and HLCr (Ludwig *et al.,* 2013, Batinic *et al.,* 2015). These assays measure two different aspects of the clonal process and long-term survival studies will determine their individual versus combined utility in response measurement and prognostication.

5.4.2 HLC versus clinical response

Progression and classification of MM is mostly based on the level of M-protein and, as expected, HLC values all showed a correlation with depth of response. Higher absolute iHLC levels, lower absolute uHLC levels and a greater difference between the two (dHLC) correlated to a worse clinical status in these treated MM patients. 19% of study participants in CR showed an abnormal HLCr which is slightly lower than the 26-28% found in similar studies (Ludwig *et al.,* 2013, D'Auria *et al.,* 2017). When examining iHLC and uHLC separately it can be seen that uHLC has a greater contribution to the HLCr abnormalities seen in CR patients. It is possible that uHLC can be a useful tool to further separate CR patients into groups with differing clinical outcomes, as the current way of defining this group is not sensitive enough to predict those who will relapse (Jakubowiak *et al.,* 2012, Kristinsson *et al.,* 2014, Kumar *et al.,* 2014, Landgren *et al.,* 2016). All CR patients were confirmed using IFE and therefore HLC has also shown increased sensitivity over this technique.

5.4.3 HLC in relapsed vs stable disease states

HLC values were examined both at best response and either at the end of the study or prior to clinically confirmed relapse. The time point of the sample prior to that confirming relapse was chosen to determine if the HLC values would have use for early relapse detection. Of all the

clinically relevant disease markers evaluated it was only uHLC which proved to be significantly different between best response and pre-relapse samples. The paired comparison suggests that monitoring an individual using HLC could help to detect relapse earlier than conventional methods. The significance of this result was confirmed by showing that those patients in stable disease did not experience a significant change in uHLC levels. A study by Espino *et al.* (2017) did not investigate uHLC levels in baseline vs pre-relapse samples, however, the study did find I/U HLCr to be significantly different between the two time points in IgG patients (p=0.01) but not IgA patients (p=0.45). This difference could potentially be due to a decreased uHLC value. Chae *et al.* (2018) found that in some clinical cases a decrease in uHLC, leading to an abnormal HLCr, was observed both before the reappearance of the M-protein on SPEP/IFE and also before iHLC became abnormally raised. The reappearance of an M-protein on SPEP/IFE or an absolute increase in M-protein size are both markers of disease progression and clinical relapse from complete response or MRD negative states (Kumar *et al.,* 2016). The importance of uHLC in pre-relapsed participants shown here, over usual markers, could help in both the redefinition of relapse and, more importantly, in detecting and confirming disease progression earlier to allow more effective intervention and reducing clinical complications.

5.4.4 Survival analysis at best response

The effect of HLC status on PFS across all clinical response groups was initially analysed, and then restricted analysis to ≥VGPR and ≥CR was performed. The significance of uHLC across all clinical response groups and its effect on PFS means that a normal uHLC status confers longer PFS despite IFE/clinical response status. These results again suggest that there is a possibility to further divide CR patients into different outcome groups. A larger cohort of CR patients is needed to fully elucidate the use of uHLC in this group.

These results suggest that it is uHLC and not iHLC/HLCr which has the most prognostic utility. Unlike in this current study, Ludwig *et al.* (2013) found that normalisation of HLCr at the time of best response was associated with a better outcome when looking at a whole patient cohort but not when restricted to those achieving ≥VGPR. A previous larger study found HLCr to be prognostically significant at diagnosis but not after treatment and did not find severe suppression of uHLC to have an influence on either PFS or OS (Lopez-Anglada *et al.,* 2018). However, conflicting with these results, Michallet *et al.* (2018) showed HLCr to be significantly associated with survival times alongside FLCr and also showed an association with HLC suppression. This supports a previous study by Harutyunyan *et al.* (2016) who also found that suppression of uHLC was significantly correlated with PFS in a heterogeneous group of treated

MM patients. A study looking at HLC and overall survival (OS) found that those with an abnormal HLCr had significantly shorter OS. Longer term follow up to ascertain the use of uHLC in OS will be needed but post-therapy uHLC suppression has also been shown to effect OS previously (Batinic *et al.,* 2015). Further work to stratify the uHLC cut off levels which might be able to further predict survival should be carried out to see if these will be more useful than the current normal ranges.

It was not found that FLCr normalisation at best response was associated with a better PFS when analysing all clinical responses and those in CR alone. This is in contrast to a small study by D'Auria *et al.* (2017) which found a significant association of FLCr normalisation with longer PFS in CR patients (p=0.049) but not OS. However, D'Auria *et al.* (2017) also found that the addition of HLCr to FLCr did not affect survival outcomes. This difference in results could be due to the smaller sample size D'Auria *et al.* (2017) analysed (25 patients). FLCr has previously been found to be prognostic in PFS and OS (van Rhee *et al.,* 2017, Lopez-Anglada *et al.* 2018) though not always (Batinic *et al.,* 2015). None of the FLC variables (iFLC, uFLC or FLCr) had a significant prognostic utility in PFS and the addition of an abnormal FLCr to an abnormal HLCr or uHLC did not improve the PFS significance. A potential reason the results of this study may disagree with previous analyses could be the use of locally produced reference ranges in this study which differ from the manufacturer reference ranges used by most studies to determine normal and abnormal FLCr result (Campbell *et al.,* 2020). The ranges used in this study were produced from a much larger cohort of over 4000 samples compared to the commonly used manufacturer ranges derived from 282 samples.

5.4.5 HLC in comparison to bone marrow analysis at MRD time points

One of the current gold standard methods for measuring minimal residual disease in treated MM patients is bone marrow NGF. In their 2016 consensus criteria the IMWG highlighted the potential role for the HLC assay in patients achieving complete response (CR) and therefore being investigated for the presence of MRD (Kumar *et al.,* 2016). When using bone marrow NGF as the gold standard method it is uHLC which has the highest agreement with this method when looking across all bone marrow sampling times, followed by HLCr. iHLC proved to be a poor predictor for bone marrow status which may be explained by the long half-life of Ig leading iHLC to not be a true reflection of bone marrow activity. At the post-ASCT time point HLCr showed the best agreement with BM NGF status, with both uHLC and HLCr showing significant agreement with BM NGF. The HLCr/NGF agreement values closely match those found by Michallet *et al.* (2018) in a larger cohort of 327 patients. When using BM NGF as the gold

standard method, uHLC showed the highest sensitivity for predicting a positive bone marrow result, whilst HLCr showed the highest specificity, meaning that an abnormal HLCr result was always accompanied by a positive BM NGF result. This has important implications for the monitoring of myeloma patients where the HLC variables could be used to determine whether a bone marrow biopsy is needed on an individual basis.

In an early study into the prognostic utility of HLC in MM patients HLCr abnormality at diagnosis was shown to have an association with shorter overall survival (Ludwig *et al.,* 2013). This was followed by studies showing that HLCr abnormality at the different post-therapy time point was associated with a shorter PFS (Batinic *et al.,* 2015, Garcia de Veas Silva *et al.,* 2016). From this study it appears to be the impact of a suppressed uHLC on the ratio which is contributing most to the HLCr influence on survival at both pre-ASCT and post-ASCT time points. uHLC proved to be prognostic for progression free survival at both the pre-ASCT and post-ASCT points. Comparing this to the results for BM NGF, at the post-ASCT time point uHLC and HLCr were both more significant in predicting progression free survival than BM NGF (see Chapter 4 section 4.3.2), both when restricting analysis to the 16 matched samples and when analysing 27 serum samples at this time point. Giving a significant survival advantage when the gold standard BM assay did not, and also providing more analysable samples at each time point shows the advantages of using a serum based method to evaluate disease status. Through this cohort it has been shown again that there is an increased sensitivity of the HLC assay over the IFE technique and therefore it has potential to improve current response criteria.

5.4.6 Significance of uHLC

In this study the significance of uHLC at the pre-relapse, pre-ASCT, post-ASCT and best response time points has been shown. Immunoparesis is a hallmark of MM as the expansion of clonal plasma cells over normal plasma cells leads to a decreased production of polyclonal immunoglobulins. In both newly diagnosed and relapsed patients the presence of immunoparesis (i.e. a decrease in IgA/IgM levels in an IgG M-protein patient) is associated with poor outcomes (Kastritis *et al.,* 2014, Heaney *et al.,* 2018, Chakraborty *et al.,* 2020). An important feature of the HLC assay is its ability to detect clonality better than conventional immunoglobulin measurements as it can distinguish between involved and uninvolved immunoglobulins and therefore show immunoparesis within a single isotype. In treated patients it is possible that the M-protein is suppressed to low levels within the normal range but within the bone marrow the reconstitution of normal plasma cells is still being suppressed by the clonal population. In this way, uHLC absolute levels may give a better indication of the disease

pathology than total Ig, SPEP, iHLC or HLCr. Time-dependent immune reconstitution after induction treatment has been shown to be significantly associated with prolonged progression free and overall survival (Klein *et al.,* 2021) and the most common cause of death during progressive disease is an infection caused by immunoparesis (Mai *et al.,* 2021). Optimal time points of assessment across different treatment groups needs to be clarified to determine the use of uHLC and HLCr in determining both immune suppression and immune reconstitution.

5.4.7 Limitations

Using the BM NGF assay as the gold standard method to compare with the HLC assay is an imperfect comparison due to the inherent limitations of the BM technique as outlined in Chapter 1 section 1.3.2. However, it was also ensured that the HLC assay was compared to the gold standard serum technique i.e. IFE and other response assessment variables to ensure a thorough evaluation. This study lacks defined follow up points for participants as the response assessment schedule was decided by each participants treating clinician. Although all participants were either on or post treatment, they were in different clinical stages of disease and had received different therapies, the effect of this on immunoparesis, specifically affecting the uHLC is unknown. Further analysis of relapsing patients will allow the plotting of time courses for individual patients to determine exact sampling windows which will be most useful clinically and can refine the use of the HLC assay variables regarding percentage changes in values and absolute changes as currently used in response criteria (Kumar *et al.,* 2016). This analysis also only evaluated intact IgG and IgA MM patients, excluding oligosecretory and light chain only MM patients. These are smaller cohorts which need larger centres and clinical trials to determine the usefulness of the HLC assay in these patient groups. The small IgA cohort in this study, 23 participants, has also limited the ability to determine the usefulness of the HLC assay over conventional techniques where both the monoclonal and polyclonal levels remain low throughout treatment and monitoring. Despite these limitations, the clinical relevance of the HLC assay has been shown in this study regardless of time since diagnosis, treatment received and time since treatment.

5.4.8 Conclusion

For this study it was an objective to not only determine the ability of the HLC assay to detect the M-protein but also determine whether it has clinical relevance for the monitoring and prognosis of participants. The HLC assay correlated well with current methods and showed enhanced sensitivity for M-protein detection over total Ig and SPEP. The addition of HLC measurements to current response assessment appears to add value, especially for those where the M-protein is

unmeasurable by conventional laboratory techniques. With the IMWG recommendation that total IgA is measured for monitoring IgA MM response to therapy the use of HLC will add significant value to the monitoring for these participants by proving a precise measurement (Kumar *et al.,* 2016).

Current criteria for determining the progression or relapse of disease is primarily based on elevations in the M-protein and/or FLCs. However, MRI/PET scanning with bone marrow plasma cell detection is considered to be more sensitive (Kumar *et al.,* 2016). Accessible tests are needed to support the continuation, changing, stopping or restarting of treatment. The measurement of uHLC has been shown to be the significant marker in this study, for sensitively detecting the M-protein, giving further clarification of response status and predicting early relapse.

uHLC has shown to be prognostic for PFS across all clinical statuses at time of best response and at both pre- and post-ASCT time points. HLC measurements indicate both the presence of residual MM tumour cells through M-protein production and the reconstruction of the immune system through the return of paired polyclonal immunoglobulin production. uHLC measurement is not affected by the extended half-life of low levels of M-protein as seen in deep responses as uHLC is the result of healthy plasma cells rather than the clonal population which could be eradicated whilst the M-protein is still detectable in the serum (Kim *et al.,* 2007). HLC can therefore be used to give a full overview on the immune system in treated MM patients. These results suggest that the HLC assay can be used as a surrogate marker for the presence of residual disease in the bone marrow of MM patients and therefore has the potential to reduce or delay bone marrow aspirations during the monitoring of MM patients.

The potential clinical utility and increased sensitivity of HLC in both the detection of M-protein at single time points and the early detection of relapse has also been shown. Serial measurements using the HLC assay are not currently part of the IMWG recommendations but the significance of uHLC both in pre-relapse samples and in PFS prognosis at MRD time points suggest that this would be useful to add to current monitoring criteria and has the ability to reduce the number of bone marrow aspirations required of MM patients.

Chapter 6 - Mass Spectrometry in the measurement of monoclonal proteins in multiple myeloma patients

6.1 Introduction

As multiple myeloma is a bone-marrow based disease, the majority of MRD studies have focused on bone-marrow based assays. However, to effectively monitor patients this would require sequential invasive procedures at frequent or semi-frequent intervals. The procedure of bone marrow aspiration is painful and inconvenient for patients and is therefore not suitable for monitoring purposes. Sensitive, non-invasive, blood-based assays which can detect the presence of MRD are needed to optimize the management of MM patients. Peripheral blood tests are appealing due to the ease and rapidity of sample collection and testing, the opportunity for more frequent sampling and longitudinal monitoring of disease burden, ease of standardisation across laboratories, and reduced costs at all stages of the testing procedure. For the aforementioned reasons, the IMWG recommends the development of a blood-based MRD monitoring tool as the ultimate goal in MRD testing (Kumar *et al.,* 2016).

In current IMWG guidelines, SPEP and IFE remain the gold standard M-protein tests despite several limitations including sensitivity issues and labour demands (see Table 1.4). Capillary zone electrophoresis is currently used in the majority of laboratories but the associated immunosubtraction method is not sensitive enough to identify low level M-proteins – therefore the more laborious serum IFE technique is still needed to determine the isotype of some M-proteins. The lack of sensitivity of the SPEP system also makes monitoring low level M-proteins, either in MGUS or treated MM patients, difficult. The use of the HLC assay in sensitively detecting the M-protein in treated patients has been explored in Chapter 5. Recently a novel method involving mass spectrometry (MS) has been explored for the detection of M-protein, with the potential to replace SPEP and IFE (Murray *et al.,* 2021). It is both a qualitative and, more recently, quantitative method which can also perform isotyping and can recognise both intact and light chain only M-proteins. This technique has the potential to overcome the limitations of bone marrow based assays and current serum based assays. Mass spectrometry (MS) based platforms are currently used in clinical laboratories for a number of protein biomarkers (Jannetto *et al.,* 2016). MS methods are capable of detecting very low levels of M-protein in the blood, with detection limits approximately 100 times lower than that of IFE (Barnidge *et al.,* 2014, Bergen *et al.,* 2016).

Mass spectrometry is currently being investigated for its clinical utility at different points in the monoclonal gammopathy pathway, from screening to minimal residual disease (MRD) detection (Murray *et al.,* 2019, Eveillard *et al.,* 2020). In a previously reported blind study, QIP-MS had a greater sensitivity for the detection of M-proteins than both SPEP and IFE (Kohlagen *et al.,* 2020). The most recent IMWG report (Murray *et al.,* 2021) recommends the further collection of data for the use of mass spectrometry for MRD detection in peripheral blood and its use in guiding the timings for bone marrow tests.

6.1.2 Aims

1. To compare the sensitivity of mass spectrometry with other serum based M-protein detection methods at active disease, best response and MRD timepoints.
2. To compare the sensitivity of mass spectrometry with bone marrow flow cytometry for the detection of MRD.
3. To determine the prognostic utility of mass spectrometry M-protein status at best response and MRD timepoints.

6.2 Methods

6.2.1 Participants

Participants were recruited into 2 different groups depending on the stage of their disease and/or their treatment plan. Group 1 were ASCT eligible patients; Group 2 were non-ASCT eligible chemotherapy only patients. This was a prospective study and the investigation schedule reflected the routine evaluation of MM at the respective centres. Bone marrow aspirates and peripheral blood samples were taken during standard patient visits to clinic. See Methods Chapter 2 for further details.

6.2.2 Study measurement schedule

For Group 1 patients bone marrow MRD analysis was planned to occur at around 1 month prior to ASCT (after induction therapy) and around 3 months post ASCT. For Group 2 patients bone marrow MRD analysis was planned for after induction chemotherapy and prior to the commencement of maintenance therapy. Serum analysis was also performed at these timepoints. For all patients, serial serum measurements were performed during regular follow up according to the clinician determined schedule (see Figure 6.1). Please see Methods Chapter 2 for further methods.

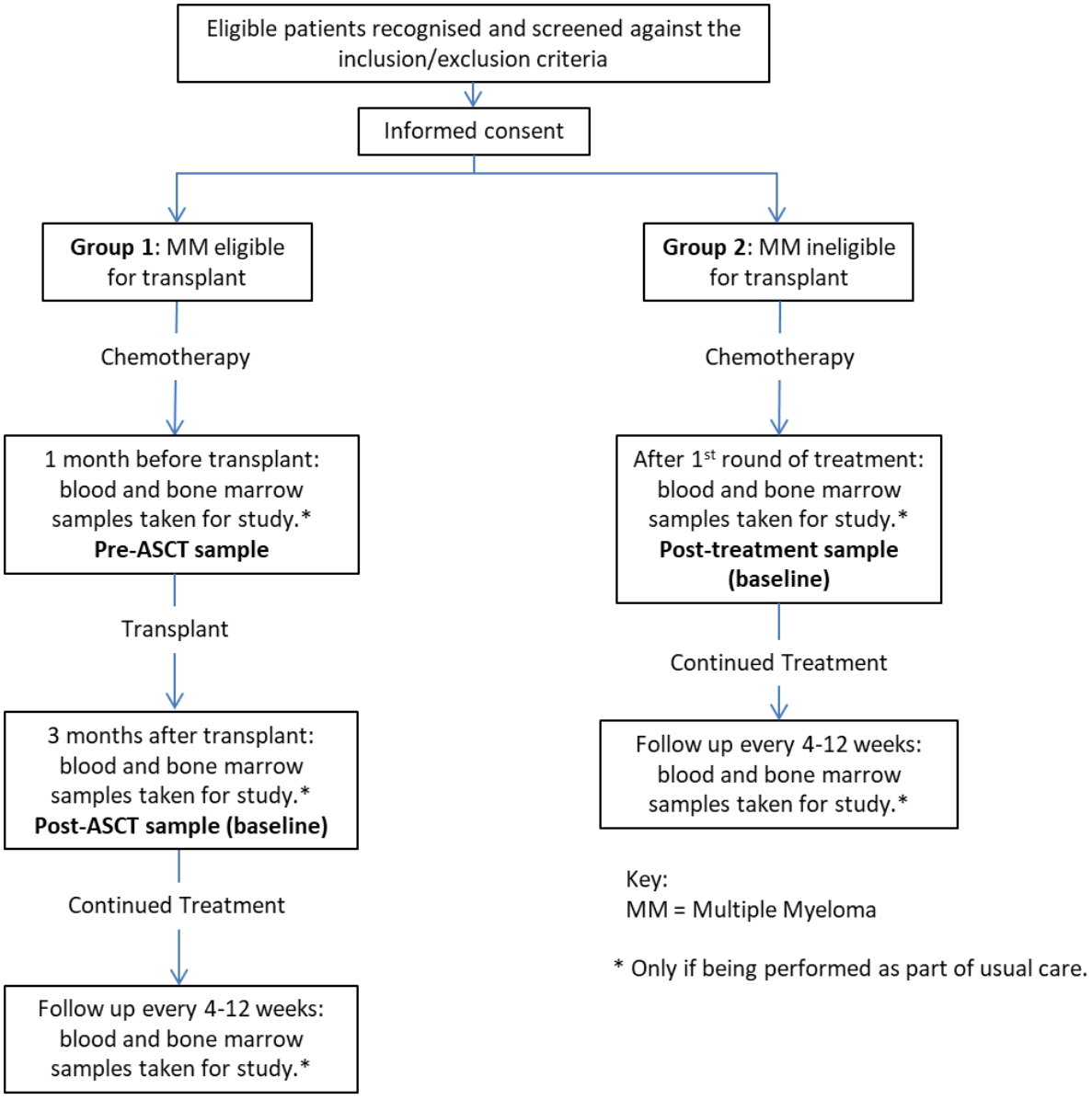


Figure 6.1: Flow diagram of the study measurement schedule.

See Chapter 2 section 2.6.4 for further information of mass spectrometry methodology.

6.3 Results

6.3.1 Patient characteristics

221 different serum samples from 87 individual participants were tested for the presence of M-protein (see Table 6.1). Not all samples had the full range of MS, SPEP, IFE, FLC results available (see individual results below for test numbers). The average number of samples received per participant was 3, with a median number of 2 (range 1 – 7 samples per participant). The median follow up time for participants was 18 months (range 3–26 months).

Table 6.1: Participant and sample characteristics

Characteristics	N (%) except where stated
Patients	87
Samples	221
Age in years (median (range))	66 (35-88)
Female	28 (32%)
Treatment	
ASCT	41 (47%)
Non-ASCT	46 (53%)
Reported diagnostic isotype	
GK/GL	36/25
AK/AL	9/5
MK/ML	1/0
Free K/free L	4/4
Biclonal	3
Oligoclonal	1
Response at time of testing	
≥CR	63
VGPR	66
PR	74
<PR	18
MS result (GAMKL method)	
Positive	186 (84%)
Negative	35 (16%)
MS result (FKFL method)	
Positive	162 (73%)
Negative	59 (27%)

ASCT = autologous stem cell transplant, CR = complete response, VGPR = very good partial response, PR = partial response

6.3.2 Comparison of MS with conventional methods for M-protein detection

The concordance of the MS method results (detection of an M-protein) with conventional methods for M-protein detection (SPEP, IFE and FLC) was evaluated to ascertain if the previous claims of MS being more sensitive are true for the study samples. The GAMKL MS method results were initially used for comparison and it is highlighted where the FKFL MS method results were also evaluated.

6.3.2.1 Serum protein electrophoresis (SPEP)

There was 68% concordance between the two testing methods, with MS detecting M-protein in 86% of samples tested but SPEP only detecting M-protein in 56% of samples (Table 6.2). The Cohen's kappa value shows a fair agreement (0.306 SE=0.055, 95%CI 0.198-0.414) between them. The majority of this discordance is due to the MS method detecting a monoclonal protein which is not detected by the SPEP method. 60 samples from 29 individual participants had an M-protein detected by MS but not by SPEP. Out of these 29 participants, 7 experienced clinical relapse (24%) during the course of this study.

Table 6.2: Agreement between SPEP and MS results for detecting M-protein in 195 samples

	MS DETECTED	**MS UNDETECTED**
SPEP DETECTED	107	2
SPEP UNDETECTED	60	26

Interestingly, there are 2 samples from 2 individual participants (see Table 6.3) where an M-protein was detected by SPEP but not MS. These samples were further investigated and it was found that including the FKFL MS trace in analysis detected a free light chain M-protein of an agreed isotype in sample 2 in Table 6.3 below. Neither of these patients who were SPEP+/MS- relapsed during the course of this study with a follow up time of 7 months and 18 months for the participants of samples 1 and 2 respectively.

Table 6.3: Samples where an M-protein was detected by SPEP but not MS GAMKL.

Sample	**Original Isotype**	**SPEP result**	**MS GAMKL result**	**MS FKFL result**	**Clinical response**	**FLCr result**	**HLCr result**	**uHLC result**
1	IgGL	Detected M-protein too small to quantify	Oligoclonal peaks	Oligoclonal peaks	VGPR – no progression during study	Abnormal	Abnormal	Abnormal
2	IgGK	Detected M-protein = 1.7g/l	Oligoclonal peaks	Free kappa	VGPR – no progression during study	Normal	Abnormal	Abnormal

Original isotype determined by IFE or CZE with immunosubtraction

Taking both GAMKL and FKFL MS method results together, concordance between SPEP and MS rises slightly to 69% and the Cohen's kappa value suggests a fair agreement of 0.316. The 1 sample which is SPEP/IFE positive but MS negative is a post-ASCT sample. This sample showed multiple oligoclonal peaks in different specificities.

6.3.2.2 Serum immunofixation (IFE)

There was 79% concordance between MS and IFE, with MS detecting M-protein in 85% of samples tested and IFE detecting M-protein in 67% of samples (Table 6.4). The Cohen's kappa value shows a moderate agreement, better than between SPEP and MS (0.462 SE=0.067, 95%CI 0.331-0.592).

Table 6.4: Agreement between IFE and MS results for detecting M-protein in 190 samples

	MS DETECTED	**MS UNDETECTED**
IFE DETECTED	125	2
IFE UNDETECTED	37	26

The 2 discordant samples which are IFE+/MS- are the same samples seen in Table 6.3 above. Taking the FKFL MS results into account also, the concordance increases to 80% and Cohen's kappa agreement value to 0.473, giving a moderate agreement between the two methods.
37 samples from 19 individual participants had an M-protein detected by MS but not by IFE. Out of these 19 participants, 7 experienced clinical relapse during the course of this study (37%) with the positive MS samples ranging from 22 months to 5 months pre-relapse. Of these 37 MS+/IFE- samples, 2 showed a difference between the MS determined M-protein isotype and the isotype determined by IFE upon diagnosis prior to this study. 1 sample was originally IgA kappa by IFE, with only free kappa detected by MS and the 2nd sample was IgG lambda by IFE but free lambda only by MS. It is possible that at very low levels of M-protein the MS assay is more sensitive at detecting free light chains over heavy chains, but this cannot be concluded from the small numbers seem here and without the m/z values of the original M-protein to determine if the original M-protein is being detected.

When considering IFE to be the gold standard method for sensitively detecting M-protein the sensitivity and specificity of MS can be seen in Table 6.5, showing a high sensitivity but relatively low specificity, reflecting the concordance data.

Table 6.5: Sensitivity, specificity, PPV, NPV and LR of MS for detection of M-protein when IFE is the gold standard

Sensitivity	Specificity	PPV	NPV	Likelihood Ratio
0.99	0.41	0.77	0.96	1.7

Isotyping agreement between IFE and MS

In all samples where IFE and MS were concordant in detecting an M-protein (MS+/IFE+) there was agreement between MS and IFE isotyping of the main M-protein, except for 3 samples (1, 2 and 3 in Table 6.6). In several patients the MS technique picked up additional peaks to the IFE method (samples 4-9 in Table 6.6).

Table 6.6: Additional M-proteins detected by the MS method in comparison to SPEP/IFE.

Sample	IFE result	MS result	Comment
1	Free kappa	IgG kappa	Patient is being treated with daratumumab which can be detected as a monoclonal IgG kappa in some patients. An IgG kappa M-protein was detected by SPEP/IFE 5 months after this sample, unable to confirm origin at this time.
2	No M-protein	IgA lambda	Patient has had no M-protein found by SPEP or IFE during diagnosis or treatment and was classified as non-secretory MM.
3	Free kappa	IgA kappa	IgA kappa M-protein found by GAMKL MS method and free kappa found by FKFL MS method and IFE method.
4	IgG lambda	IgG lambda and free kappa	An additional free kappa M-protein was detected by both GAMKL and FKFL MS.
5	IgG lambda	IgG lambda and free kappa	An additional free kappa M-protein was detected by FKFL MS only.
6	IgG kappa	IgG kappa x2	Additional IgG kappa detected by GAMKL MS.
7	IgG kappa	IgG kappa x3	Additional IgG kappa x2 detected by GAMKL MS.
8	IgA kappa	IgA kappa and IgG lambda	Additional IgG lambda detected by GAMKL MS.
9	IgG kappa	IgG kappa and IgG lambda	Additional IgG lambda detected by GAMKL MS.

Sample 1 in Table 6.6 had an original isotype of free kappa by IFE at diagnosis but is showing IgG kappa by MS. When looking into this participant it was discovered that they were being treated with daratumumab, an IgG kappa monoclonal antibody therapy. It is possible that either the MS technique is identifying daratumumab as an M-protein in this participant or the participant has a genuine myeloma associated IgG kappa M-protein. The FLCr for this patient was abnormally raised, suggesting the definite presence of a kappa associated monoclonality, but HLCr was normal. Without an original diagnostic MS sample it cannot be determined whether this IgG kappa is related to the original clone, is a therapeutic antibody or is a new clone. Since running this analysis the MS technology has now improved to be able to identify monoclonal antibodies used for therapy and this sample can be investigated further in the future.

6.3.2.3 Serum free light chains (FLC)

We looked at MS and FLC agreement, initially separating analysis by the two MS methods which can detect monoclonal free light chains (GAMKL MS and FKFL MS) (Table 6.7 and 6.8). An abnormal FLC ratio (FLCr) suggests the presence of a monoclonal free light chain and is used for both diagnosis and monitoring of MM patients.

Table 6.7: Agreement between FLCr and GAMKL MS results for detecting M-protein in 190 samples.

	MS DETECTED	**MS UNDETECTED**
FLCr Abnormal	133	13
FLCr Normal	26	18

The concordance between GAMKL MS and FLCr is 80%, with GAMKL MS being positive in 84% of samples and FLCr being abnormal in 77%. The Cohen's kappa value is 0.357, showing fair agreement.

Table 6.8: Agreement between FLCr and FKFL MS for detecting M-protein in 190 samples

	MS DETECTED	**MS UNDETECTED**
FLCr Abnormal	120	26
FLCr Normal	25	19

The concordance between FKFL MS and FLCr is 73%, with FKFL MS being positive in 76% of samples and FLCr being abnormal in 77% of samples. Cohen's kappa agreement for these method is 0.252, showing slight agreement.

Evaluating both MS methods indicates that GAMKL MS is more sensitive and a more useful measure of monoclonality than FKFL MS in detecting both intact and light chain M-proteins, however analysing both types of MS together gives the highest sensitivity.

6.3.2.4 Comparison of MS vs conventional method algorithm

As SPEP, IFE and FLC analyses are often performed together in suspected MM patients, the utility of using MS alone was compared to an algorithm of SPEP/IFE/FLC tests, where if any of the conventional tests showed abnormality this was considered positive. Using this scenario, MS outperforms the conventional tests in the 190 samples evaluated - with 7 samples being MS positive only and 1 sample being negative by MS but positive by all SPEP/IFE/FLC (sample 2 in Table 6.3).

6.3.3 Comparison of MS method with HLC measurements

This study has previously shown the enhanced sensitivity of the HLC assay in detecting M-protein over SPEP and IFE, and how it complements the FLC assay. The sensitivity of MS was therefore compared with HLC values, first looking at HLCr (Table 6.9) and then at uHLC (Table 6.10), which was found to be the most important HLC value when performing survival analyses. Due to earlier analysis showing that using GAMKL MS and FKFL MS together gives the best sensitivity, merged results from both MS methods will be used together in ongoing analysis.

Table 6.9: Agreement between HLCr and MS in 185 samples.

	MS DETECTED	**MS UNDETECTED**
HLCr Abnormal	114	5
HLCr Normal	43	24

The agreement between HLCr and MS is 75%, with HLCr abnormal in 64% of samples and MS positive in 85% of samples. The Cohen's kappa value shows a fair agreement between methods (0.361, SE=0.067, 95%CI 0.230-0.492).

There were 5 samples from 4 individual participants which had an abnormal HLCr but no M-protein detected by the MS method. The abnormal HLCr was due to an abnormal iHLC in 1 of the samples and an abnormal uHLC in all 5 samples. 1 out of the 4 participants relapsed during the course of the study, at 8 months after the sample was obtained. 43 samples from 23 individual participants had a normal HLCr but an M-protein detected by the MS method. Of these 23 participants, 1 relapsed during the course of the study at 5 months after this sample

was obtained. This IgA kappa M-protein sample had an abnormal uHLC at this time point but the HLCr value only became abnormal at the point of clinical relapse.

Table 6.10: Agreement between uHLC and MS in 185 samples

	MS DETECTED	MS UNDETECTED
uHLC Abnormal	129	15
uHLC Normal	28	13

The agreement between uHLC and MS is 76%, with uHLC abnormal in 77% of samples and MS positive in 85%. The Cohen's kappa value showed a fair agreement (0.24 SE=0.084, 95%CI 0.076 -0.404).

There were 15 samples from 12 individual participants where the uHLC was abnormal but the MS method did not detect an M-protein. 2 out of these 12 participants experienced disease relapse during the course of this study (8 months and 12 months after these samples were obtained). 28 samples from 16 individual participants had a normal uHLC value but an M-protein detected by the MS method. Of these 16 participants, 1 relapsed during the course of the study, 19 months after the MS sample was obtained.

6.3.4 MS sensitivity in different clinical response groups

IMWG clinical response criteria is currently used to help aid treatment decisions. To evaluate how MS results might be different across difference response groups a comparison of the percentage of samples within each clinical response group which had a positive or negative MS result was made (Figure 6.2). For this analysis only MS samples were considered to be positive when the monoclonal protein seen on the mass spectra matched the original isotype for that participant. This was to try to exclude oligoclonal proteins, often seen after intense treatment, which do not represent a true M-protein and therefore would be considered false positive. As expected, the percentage of samples with a negative MS result increased with greater depth of response whilst those in PR or <PR had no negative MS samples. However, even in the deepest response groups (CR, VGPR) the majority of samples were MS positive. This suggests that those patients currently grouped as VGPR, and especially CR, may be able to be subdivided according to MS result and this could have a prognostic value in the monitoring of these patients. One participant who had a CR sample which was also MS negative went on to relapse 21 months after the MS sample, with no samples closer to the relapse time available for MS analysis.

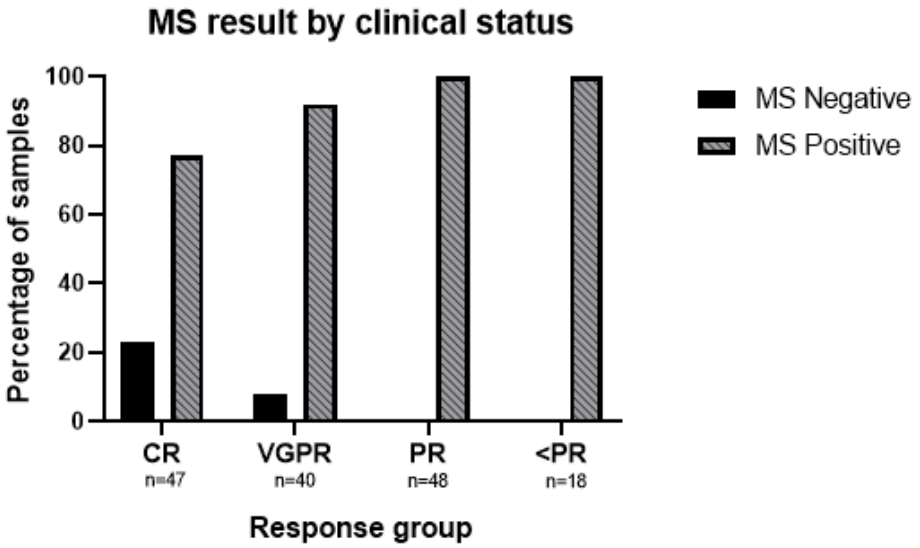


Figure 6.2: Percentage of samples in each clinical response status group with a positive or negative MS result at the time of sampling. Fisher's exact test p=0.0004 comparing CR vs non-CR samples.

6.3.5 Comparison of MS results with NGF at bone marrow testing time points

The current most sensitive method for detecting disease, and especially MRD, is bone marrow NGF. In order to determine if a serum assay can replace the use of bone marrow then the sensitivity of serum and bone marrow assays must be compared through concordance, sensitivity and survival analyses.

Initially a comparison between all available bone marrow NGF results (diagnostic, pre-ASCT and post-ASCT time points) and MS results from samples taken at the same time points was performed (Table 6.11). Out of 53 matching bone marrow/serum samples, 42 were in agreement between the two methods, giving 79% agreement. BM NGF was positive in 83% of samples and MS was positive in 81% of samples. The Cohen's kappa showed a fair agreement (0.3, SE=0.16, 95%CI -0.02-0.61). When considering BM NGF to be the gold standard method for sensitively detecting M-protein the sensitivity and specificity of MS can be seen in Table 6.12. The sensitivity is of MS is relatively high, though not as high as uHLC (0.93) when assessing across all time points. However, the specificity of this assay and the negative predictive value (NPV), i.e. how well a negative MS result can predict a negative BM NGF result, are poor.

Table 6.11: Agreement between BM NGF and MS results at diagnostic, pre-ASCT and post-ASCT time points in 53 samples.

	MS DETECTED	MS UNDETECTED
BM NGF Positive	38	6
BM NGF Negative	5	4

Table 6.12: Sensitivity, specificity, PPV, NPV and LR of MS for detection of M-protein when BM NGF is the gold standard

Sensitivity	Specificity	PPV	NPV	Likelihood Ratio
0.88	0.40	0.86	0.44	1.5

6.3.5.1 Pre-ASCT bone marrow NGF and serum MS samples

Further evaluation of the utility of MS at MRD specific timepoints was then performed. At the pre-ASCT time point 25 participants had BM NGF and serum MS results available from the same time point. All BM NGF samples were positive whilst 22/25 MS samples were positive, giving an agreement of 88%. Of the 3 discordant samples 1 patient relapsed during the course of the study. This patient had been clinically classified as being in CR, however abnormal HLCr was found at the same time point.

6.3.5.2 Post-ASCT bone marrow NGF and serum MS samples

At the post-ASCT time point 16 participants had BM NGF and serum MS results available from the same point. The agreement at the post-ASCT time point between the two methods was only 56% (Table 6.13). In 4 samples MS was positive whilst BM NGF was negative. None of these MS+/NGF- participants experienced clinical relapse during the study. Of the 3 samples which were MS negative and BM NGF positive, 1 participant was in CR and 2 in VGPR at the time of sampling and none of the participants relapsed during the study. 2 out of 3 MS-/NGF+ samples showed the presence of oligoclonal peaks on the MS trace of different specificities and without the diagnostic sample the presence of a low-level M-protein could not be determined. The MRD levels in these 3 participants ranged from 0.0006% to 0.009%, which are all close to the limit of sensitivity of the BM NGF assay (0.0005%). Oligoclonal peaks of different specificities were also seen in 3 out of the 4 samples which were negative by both the NGF and MS methods. When using BM NGF as the gold standard method for disease detection the sensitivity of MS at the post-ASCT time point is lower than the sensitivity at all time points (Table 6.14) and also lower than the sensitivity of IFE at the post-ASCT time point (Table 6.15).

Table 6.13: Agreement between BM NGF and MS results at the post-ASCT time points in 16 samples.

	MS DETECTED	MS UNDETECTED
BM NGF Positive	5	3
BM NGF Negative	4	4

Table 6.14: Sensitivity, specificity, PPV, NPV and LR of MS for detection of M-protein when BM NGF is the gold standard at the post-ASCT time point.

Sensitivity	Specificity	PPV	NPV	Likelihood Ratio
0.57	0.55	0.5	0.62	1.3

Table 6.15: Sensitivity, specificity, PPV, NPV and LR of IFE for detection of M-protein when BM NGF is the gold standard at the post-ASCT time point.

Sensitivity	Specificity	PPV	NPV	Likelihood Ratio
0.83	0.55	0.55	0.83	1.86

6.3.6 Progression free survival (PFS) analysis of mass spectrometry results

When performing survival analysis on the pre-ASCT MS samples there was no significant PFS difference between MS positive and MS negative participants (p=0.51). Survival analysis was performed for the above (Table 6.13) 16 post-ASCT participants to allow direct comparison to the BM NGF survival analysis (Chapter 4, p=0.051), finding no significant difference between PFS of the post-ASCT MS positive and MS negative participants (p=0.16). However, when taking into account all available post-ASCT MS samples, including those without a matching BM, survival analysis can be performed on 30 samples and a significant difference in survival is seen between the two groups (p=0.02) (Figure 6.3).

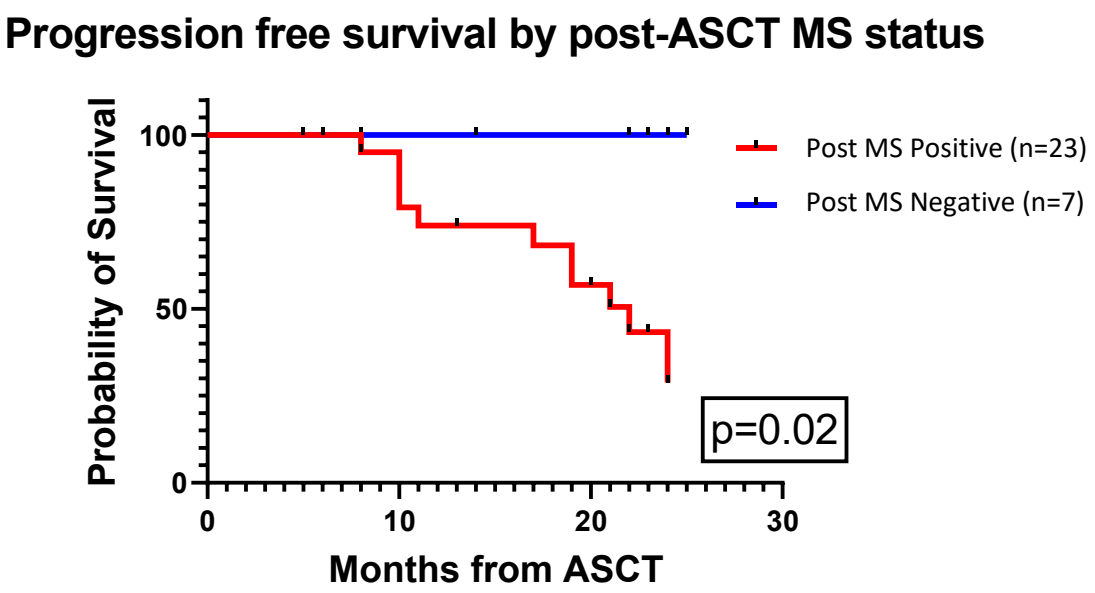


Figure 6.3: PFS survival of post-ASCT MS positive vs MS negative patients. (HR, 4.6; 95% CI, 1.3-17.5; p=0.02). Median survival is 22 months vs undefined.

In order to see the improvement of using MS of the current response criteria survival analysis was performed on this same set of either BM matched 16 serum samples or the extended 30 serum samples using the IFE status (positive or negative). IFE status did not produce a significant difference in survival between groups for the 16 (p=0.21) or the 30 (p=0.44) samples (see Figure 6.4 for the 30 samples).

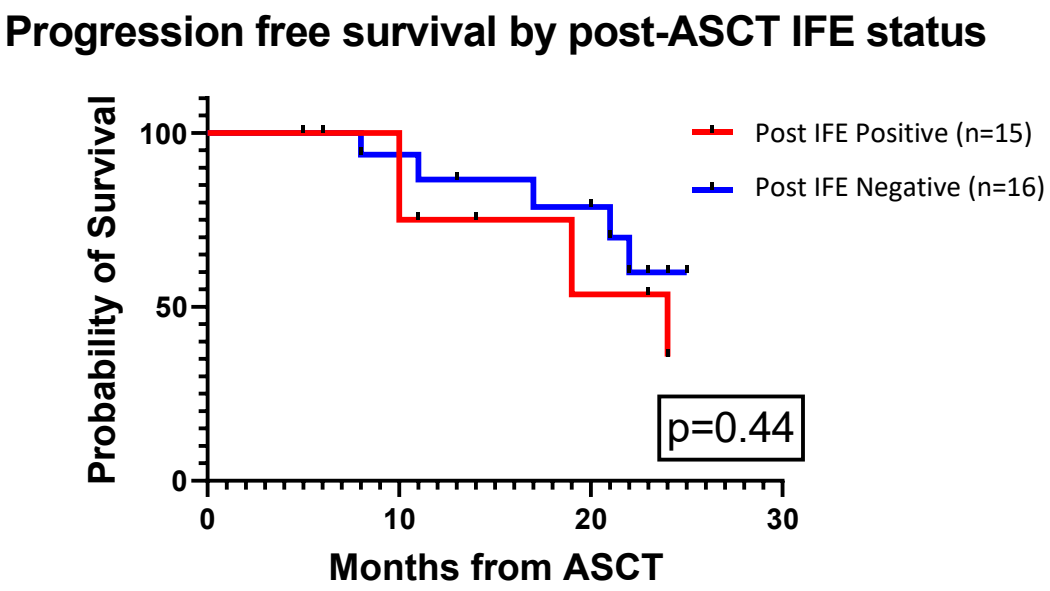


Figure 6.4: PFS survival of post-ASCT IFE positive vs IFE negative patients. (HR, 1.6; 95% CI, 0.48-5.2; p=0.44). Median survival is 24 months vs undefined.

Next survival analysis was performed using MS results at the time of best response as determined by IMWG criteria. Best response samples from 87 participants were analysed, resulting in 66 MS positive samples and 21 MS negative samples. There was a significant PFS difference between MS positive and MS negative participants (p=0.007), with median survivals of 20 months vs 26 months respectively (Figure 6.5).

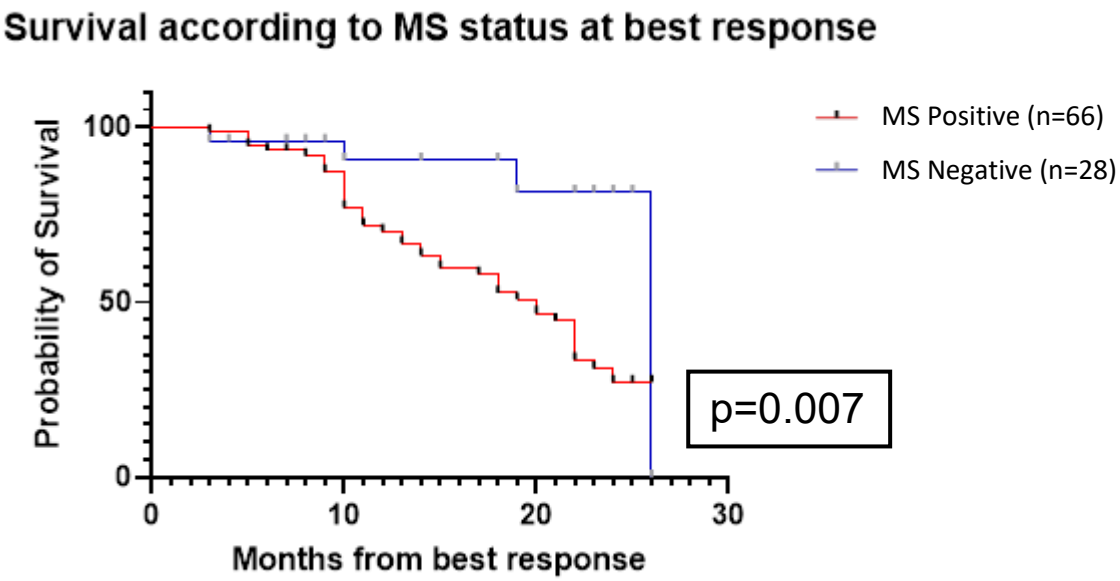


Figure 6.5: PFS survival of best response MS positive vs MS negative patients. (HR, 3.67; 95% CI 1.9-5.1; p=0.007). Median survival is 20 months vs 26 months.

6.4 Discussion

This study aimed to evaluate the viability of using MS to identify M-proteins in treated MM patients. Due to its reported sensitivity MS has the potential to give more information on disease status in MM patients across different treatments, times since diagnosis and clinical response states. Sensitively measuring for the presence of disease is vital for appropriate patient care, aiding decisions such as switching, stopping or increasing treatments. As deeper responses are associated with better progression free and overall survival it is important to appropriately stratify patients by level of response (Avet-Loiseau *et al.,* 2020). MS is an important technique to investigate as its high sensitivity brings the possibility of replacing bone marrow assays with serological assays closer to realisation.

As a pre-commercial MS system was used, two methods to identify monoclonality were carried out, the intact immunoglobulin light chain based FKFL method and the newer immune-enrichment intact immunoglobulin GAMKL method which allows identification of the heavy chain. The GAMKL method is the preferred method due its ability to performing complete immunotyping and therefore having the potential to replace IFE at the diagnosis stage, however its sensitivity when compared to the original FKFL method needed to be evaluated. From overall

positivity rates the GAMKL MS method has improved sensitivity over FKFL MS and when comparing the two MS methods with conventional techniques it was concluded that the highest sensitivity came from using both methods together.

6.4.1 Comparison of the MS method with conventional techniques

Currently, the M-protein is identified and characterised with SPEP and IFE respectively. Despite its limited sensitivity, SPEP is the main technique used for monitoring M-proteins during the course of the disease due to the more labour intensive and qualitative nature of IFE. MS can identify and characterise M-proteins in one step. In this study MS showed increased sensitivity over SPEP and IFE, as expected from previous studies (Mills *et al.,* 2016, Kohlhagen *et al.,* 2020, Eviellard *et al.,* 2020, Dispenzieri *et al.,* 2022). The highest kappa agreement value between techniques was between MS and IFE, with only 1 sample being MS negative but IFE positive. When looking further into this sample it was a day 100 post-ASCT sample and the MS trace showed multiple oligoclonal peaks of different specificities. There is the potential that if a diagnostic or pre-treatment sample for this participant was available then the original M-protein mass spectra would help to identify the M-protein peak against this oligoclonal background. This participant did not relapse during the course of the study but had a short follow up time of 7 months post-ASCT.

When considering IFE to be the gold standard M-protein detection technique the MS method showed very high sensitivity. The relatively low specificity here is due to a high number of samples being MS positive and IFE negative. When evaluating the 37 samples from 19 individual participants which were MS positive and IFE negative it was found that 7/19 (37%) participants in this group relapsed during the course of the study. This is a high number, considering overall 17/87 (20%) participants in the whole study experienced relapse. 41% of relapsed patients being in this MS positive/IFE negative group shows that the high sensitivity of the MS assay is clinically relevant in the monitoring of treated MM patients. Further longitudinal studies will help to identify the use of MS in predicting relapse and the implications this will have on clinical decisions regarding the use and continuation of maintenance therapy. Overall concordance between the MS method and FLC measurement was the same as between MS and IFE. Overall the MS method had a higher number of positive results than any of the conventional methods, and outcompeted the algorithm of SPEP/IFE/FLCr in detecting the presence of M-protein. As all the participants included in this study have a confirmed MM diagnosis it can be concluded that the MS method is both highly sensitive and specific for detecting the M-protein in this cohort. The MALDI-TOF MS system has been shown to be comparable to IFE in terms of turn around

time and ease of interpretation but with the benefit of reduced labour input (Murray *et al.,* 2021).

Isotyping agreement between the MS method and IFE was high in those samples where the M-protein was detected by both techniques. This shows that even in treated MM samples, where the M-protein is usually at a lower concentration, MS is a valid alternative to IFE for isotype characterisation. In one sample the MS method detected a heavy chain (IgA kappa) in a sample characterised as free kappa only by the IFE technique. This could have potential implications for the monitoring and prognosis of patients, with light chain only myeloma patients having a worse prognosis than intact myeloma patients (Drayson *et al.,* 2006). Another sample showed an IgG kappa M-protein by MS which had originally been isotyped by IFE as a free kappa light chain only. As this patient was being treated with a monoclonal antibody it could not be determined at this time whether this was the original M-protein, a new M-protein or the monoclonal antibody. However, the MS software technology has now advanced to allow detection of monoclonal antibodies and their differentiation from patient M-proteins (Murray *et al.,* 2021). In addition to detecting the main M-protein, the MS method also detected some additional monoclonal peaks. In post-ASCT patients it is expected to see an oligoclonal pattern during reconstitution of the bone marrow, although they often remain undetected by SPEP. Oligoclonal bands post-ASCT have been associated with a better prognosis, however the clinical significance of these peaks at other times in the disease course is unknown (Silva *et al.,* 2017). An interesting participant was one classified as non-secretory which showed an IgA lambda M-protein by the MS method. The identification of a detectable M-protein in patients classified as non-secretory or oligo-secretory would have a big positive impact on the monitoring of these patients and reduce the need to define these patients according to bone marrow analysis.

6.4.2 Comparison of the MS method with HLC measurements

It has been previously shown that the HLC assay correlated well with conventional methods (Chapter 5) and showed enhanced sensitivity for M-protein detection, adding value to current response criteria. The MS method may represent a further sensitive measurement of the M-protein and therefore its performance to the HLC assay was compared, specifically the more commonly used HLCr value and the uHLC value which was shown to be significant in prognosis. There was a slightly higher concordance between the MS method and uHLC than with HLCr, although kappa agreement was higher with HLCr. Agreement between the MS method and HLC values overall were lower than with IFE. A higher number of samples were uHLC abnormal/MS negative than HLCr abnormal/MS negative. Of the 12 participants who showed a uHLC

abnormal/MS negative pattern, 2 relapsed during the course of the study. These relapses occurred at 8 months and 12 months after the MS negative result and there is not an MS result from samples closer to the relapse date. In these participants the uHLC value remained abnormal up to and including the time of clinical relapse but IFE remained negative until relapse. 1 participant relapsed 19 months after a uHLC normal/MS positive result. In this participant the uHLC remained normal up to relapse but the HLCr was abnormal throughout follow up and matched the MS result. These results show the importance of looking at all the variables from the HLC assay, which are produced from a single sampling. As with the conventional methods, MS was positive in a higher number of samples than either HLCr or uHLC and therefore represents a more sensitive method for M-protein detection.

6.4.3 MS in clinical response groups

Progression of MM and the classification of the response is mostly based on the presence of M-protein and its level. Whilst currently this study has not been able to use MS to quantify the M-protein, its detection of the M-protein across different clinical response groups can be evaluated. As expected, a correlation was seen between the detection of M-protein by MS and depth of response. 77% of samples classified as correlating to CR were MS positive. All CR patients had previously been confirmed using IFE and therefore this high number of MS positive patients indicates the increased sensitivity of MS over IFE and also the uHLC results presented in the previous chapter. This supports data shown by Dispenzieri *et al.* (2022) who demonstrated that the Mass-Fix MS method outperformed both the IMWG criteria and IFE response in survival analysis.

6.4.4 Comparison with BM flow MRD

Assessment of MRD in the bone marrow has been shown to be prognostic for survival and there is great interest in refining and standardising MRD techniques to give the most clinical benefit for both clinicians and patients (Munshi *et al.,* 2020). Current MRD techniques focus on bone marrow evaluation, however serological evaluation has multiple benefits for the clinician, laboratory and patients. A NGF MRD bone marrow assay has been validated to use as a gold standard method with which to compare serological assays to determine their relative sensitivity (Chapter 3).

When comparing available NGF MRD results with MS result across the disease course it was found that both methods assigned a positive response to a similar number of samples, 44 and

43 samples respectively but the concordance was only fairly high at 79% with a fair kappa agreement value. Concordance at the pre-ASCT time point between methods was high at 88%, matching a previous study (Puig *et al.,* 2019). However, when evaluating post-ASCT samples only, at which MRD status is thought to be most useful, this concordance dropped to only 56%. This is similar to previous studies which found a 62% and 64% concordance between methods at this time point (Eviellard *et al.,* 2020, Foureau *et al.,* 2021). It was calculated that IFE had a better sensitivity at post-ASCT than MS when using NGF as the gold standard method, this is in contrast to a recent larger study finding that MS had a better sensitivity than IFE compared to bone marrow MRD at all time points (Dispenzieri *et al.,* 2022). However, the sample numbers in this study are very small so a conclusion cannot be drawn on the comparable sensitivities of these techniques. This increased sensitivity of IFE over MS also did not translate to the significant survival advantage seen with the MS status. At the post-ASCT time point 44% of participants were MS negative which correlated well with a previous study showing 41% MS negativity at this time point (Dispenzieri *et al.,* 2022). When comparing MS to the uHLC/HLCr results at the post-ASCT time point uHLC and HLCr both showed higher sensitivity and specificity than the MS method.

As with other serological assays the discordance between NGF and MS can be attributed to factors such as the patchiness of bone marrow disease and the lag between bone marrow resolution of disease and M-protein elimination due to its extended half-life, especially at low concentrations (Kendrick *et al.,* 2017). The presence of oligoclonal peaks in 2 out of the 3 samples which were NGF positive/MS negative might also suggest the presence of the original M-protein clone but due to not having diagnostic information on the specific mass of this M-protein it cannot be identified against the oligoclonal background.

6.4.5 Survival analysis

At the pre-ASCT time point the MS result did not have a significant effect on PFS (p=0.51). In the course of this study the only measurement which has had a survival difference at this time point (pre-ASCT) is the uHLC status (p=0.03). When directly comparing the post-ASCT NGF and MS samples neither produced a survival difference between groups (p=0.051 and p=0.16 respectively). However, when using the larger cohort of day 100 post-ASCT samples which did not necessarily have a matching bone marrow sample a significant survival benefit was seen in MS negative participants (p=0.02). Due to the better concordance of IFE measurements with NGF at this time point survival analysis on IFE status was also performed and this did not produce a significant result (p=0.44).

This study has shown MS status at post-ASCT to significantly effect survival in a univariate analysis, however with more data a multivariate analysis can be conducted to determine if it will remain prognostic. Dispenzieri *et al.* (2022) did not find MS status to be significant for survival in a multivariate analysis at either the pre- or post-ASCT time points, but did find it to be significant at 1 year post-ASCT. They also found NGF MRD measurement to still be the best predictor of survival at the post-ASCT time point and speculated that this was due to both the kinetics of the M-protein clearance and the presence of oligoclonal peaks causing false MS results. Dispenzieri *et al.* (2022) concluded that Mass-Fix and NGF were complementary measurements.

The effect of MS status on PFS across all clinical response groups at the time of best response for each patient (time after treatment where the deepest response was achieved during the course of the study) was next evaluated. The significance of MS across all clinical response groups and its effect on PFS means that a negative MS status confers longer PFS irrespective of clinical response status (p=0.007). These results also correlate with the previous analyses which showed CTPC status, uHLC status and HLCr status at best response also showing significant PFS benefits. These MS results again suggest there is a possibility to further divide CR patients into different, deeper outcome groups. A larger cohort of CR patients is needed to fully elucidate the use of MS in this group due to a high number of study participants not reaching CR during the study.

6.4.6 Limitations

As in other aspects of this study a limitation here was in the participant numbers, especially those undergoing bone marrow transplant and providing pre- and post-ASCT bone marrow samples. Larger sample numbers and longer follow up periods will allow more convincing conclusions to be drawn regarding survival benefits and prognostic use of these assays. It will be important to perform this work in a cohort with a larger number of patients in CR, which is where sensitive disease detection will be most useful. When considering the overall cohort there are many non-ASCT patients who are undergoing less intensive treatments and are therefore less likely to achieve the deepest responses. Another limitation was the lack of diagnostic sample which would allow the identification of the primary M-protein by MS and the follow-up of this specific clone during the disease course. Having this information may help to better evaluate the oligoclonal peaks seen in some patients and would help avoid both false positive and false negative MS results. There is the possibility that an oligoclonal peak was identified as the primary

M-protein but the opposite is also true where a low level M-protein was interpreted as an oligoclonal peak. There is also the possibility of a false positive result in patients on monoclonal antibody therapy. It is beyond the scope of this study but it will also be of benefit to look at the glycosylation status of the M-protein mass spectra peaks and determine if this changes during the course of the disease and what clinical impact that may have (Milani *et al.,* 2017). Finally, subsequent to the MS analysis of the cohort of samples, the MALDI-TOF technology has improved and now allows of quantification of the M-protein using analysers specific to The Binding Site company. Re-testing the samples will allow comparison and survival analyses based on MS M-protein levels alongside the presence or absence of this protein on the mass spectra.

6.4.7 Conclusion

For this study an aim was to determine the ability of the MS assay to detect M-protein and whether it has clinical relevance for the monitoring and prognosis of MM patients. It has been shown that MS has enhanced sensitivity over conventional methods and HLC measurements. There is a survival benefit for being MS negative at both post-ASCT and best response time points across a range of patients and treatments. There were a high number of participants who were MS positive throughout monitoring and it will be important to determine how such a sensitive method can be used to provide clinically useful information to clinicians to give the best balance between treatment and quality of life for patients.

Chapter 7 - Conclusions and Future Work

7.1 Overview of the study

Multiple myeloma is a complex disease, and despite great advances in both detection and treatment of the disease, there remains a vast amount of research interest in improving patient outcomes. With advanced treatments the majority of MM patients, both transplant eligible and ineligible, are now achieving very deep responses (Korde *et al.,* 2015, Mailankody *et al.,* 2015, Mateos *et al.,* 2022). Due to the need for more sensitive methods to assess response this study aimed to investigate MRD markers in treated MM patients from a laboratory perspective and also address other disease course time points where monitoring could be improved. During the course of this study, improvement in technologies and new techniques have emerged but MRD assessment is still not being used in routine practice and is confined to clinical trials due to the complex nature of the methods. In the latest published IMWG guidelines it was suggested that more data be collected to determine the role the HLC assay may have in the intact immunoglobulin MM patient pathway (Kumar *et al.*, 2016). The focus on blood-based markers has allowed this study to add to growing evidence of the use of more readily available tests in deep response evaluation and conclusions drawn here will lead to some recommendations on how these markers could be used for patient benefit. Comparing the sensitivities of bone marrow and blood-based methods is not without complications. The bone marrow NGF method counts the number of abnormal plasma cells in a single sample and the issues with using a single bone marrow sample to give an overall picture of disease state have been discussed previously (section 1.3.2), including the patchiness of bone marrow involvement and the presence of extra-medullary disease. Using the M-protein as a sensitive marker for disease status also has obstacles such as the extended half-life of the M-protein potentially leading to its detection after abnormal plasma cells have been eradicated and the unknown effect that treatment could have on the synthesis rate of the M-protein. Despite these inherent difficulties there is confidence in the recommendations made here for the extended use of M-protein assays in the MM patient pathway.

7.2 Summary of the findings

The first objective of the work described in this thesis was to validate a flow cytometric assay for the measurement of low levels of bone marrow and circulating tumour plasma cells and to standardise this for use throughout the study. This was addressed in Chapter 3. Next generation flow cytometry (NGF) is one of two currently validated highly sensitive methods for the detection of MRD in the BM, the other being next generation sequencing. A NGF assay was

validated to laboratory standards using the EuroFlow procedure guidelines and adapting these to a single-tube method (Stetler-Stevenson *et al.,* 2016, Roshal *et al.,* 2017). It was shown that this NGF assay could be used to detect plasma cells in both bone marrow and peripheral whole blood samples and that this could reach a sensitivity level up to 10^{-6}. A simple MRD analysis template was developed which is easily reproducible without using expensive or complicated software tools and used this assay and analysis template for the duration of the study.

The second objective was to evaluate whether the NGF assay had prognostic utility when assessing MRD status in the bone marrow. This has been shown by multiple previous studies (Munshi *et al.,* 2020) but was important to establish here as this study would be using this result to compare against the blood-based assays. The results presented in Chapter 4 showed a trend towards the negative impact that the presence of MRD at the post-ASCT time point has on survival (p=0.051). Unfortunately not enough viable bone marrow samples were received to power significant results. This reiterated that not only are bone marrow samples unpleasant for the patient but they also present significant problems for the laboratory in other ways. Several unusable bone marrow samples were received due to reasons such as haemodilution, low sample volume, clotted samples, samples received outside of stability, and samples received at the wrong time point. This was compounded by the COVID-19 pandemic preventing patients from coming to hospital and delaying ASCTs. However, the result of no clinical disease progression in the MRD negative group suggests that the NGF technique and study participants match other reports of the strong association between MRD status and survival. The results here also supported the objective for therapy to be aiming towards MRD negativity, rather than a CR/sCR response.

The third objective was to evaluate how blood-based assays could also be used to determine MRD status and how these compared to BM MRD analysis at specific time points in treatment. Beginning with CTPC analysis by NGF reported in Chapter 4, it was shown that at both the pre-ASCT and post-ASCT time points the CTPC assay had 100% positive predictive value (PPV) for bone marrow status, i.e. a positive CTPC result was always accompanied by a positive BM MRD result. However, it could not be shown that CTPC detection was better or even as good as BM MRD assessment. This indicates that the CTPC assay could be used as a screening assay to decide whether a bone marrow sample needs to be taken or whether it could confidently be predicted that disease is present in the bone marrow. The work reported in Chapter 5 explored the use of the HLC assay in treated MM patients. This work showed that uHLC and HLCr had significant agreement with the NGF result at the post-ASCT time point. When plotting survival curves uHLC

was shown to be prognostic for PFS at both the pre-ASCT and post-ASCT time points, which could not be shown using BM MRD status, though a larger number of samples were used for uHLC analysis. Finally, in the work presented in Chapter 6, the MS technique was evaluated and a fair concordance was found between MS and NGF BM results. MS showed a significant survival benefit at the post-ASCT time point. Importantly for the uHLC and MS results, this significant survival benefit at the post-ASCT time point was not seen when grouping by IFE status (conferring CR vs non-CR).

The fourth objective was to assess the utility of the blood-based assays during the follow up of treated patients and to see how they compare with current conventional techniques. Samples were analysed at the time of best response, as determined by the current IMWG criteria, and survival analysis performed to determine the prognostic utility. The results presented in Chapter 4 showed the negative prognostic impact on PFS of the presence of CTPCs at the time of best response, and this remained true when analysing only patients in CR. In the results presented in Chapter 5, uHLC was shown to be prognostic for PFS at the time of best response across all clinical statuses and also when restricting analysis to either those in VGPR or better and CR. HLCr and, to a greater extent, uHLC were abnormal in a high number of patients classified as CR and therefore IFE negative. In the results presented in Chapter 6, MS status also conferred survival advantages when analysed at best response and showed positivity in a high number of CR patients. Taken together these results show how the blood-based assays have increased sensitivity over current IMWG conventional blood-based assays and are able to further sub-group those patients defined as being in CR. When analysing participants who relapsed during the course of the study, it was found that of all the clinically relevant disease markers evaluated it was only uHLC which proved to be significantly different between best response and pre-relapse. This suggests that monitoring an individual using the HLC assay could help to detect relapse earlier than conventional M-protein detection methods and the monitoring of other clinical parameters.

7.3 Implementation of blood-based assays into clinical practice – useful or not?

The work described in this thesis originated from a hope to improve the MM patient experience by reducing the need for bone marrow assessment whilst maintaining a sensitive and high-quality disease evaluation process. The fifth and final major objective was to determine if and how these blood-based assays should be implemented into routine clinical practice in order to provide improved care to patients. This improved care may be through earlier detection of relapse, as explored in Chapter 5, or potentially through replacing, enhancing or delaying bone

marrow assessment. The use of blood-based assays, and the HLC assay in particular, will also have a benefit to the laboratory as sensitive MRD testing becomes more sought after by clinicians. The analyser for the HLC assay is already in place in the majority of laboratories and samples are run on a continuous loading basis with walk away capability and results can be easily integrated into current laboratory information systems. This will make it a far easier test to validate and introduce into current laboratory workflows than a flow cytometry assay which relies on greater user input, complex training and competency assessment, highly skilled users and subjective analysis.

7.3.1 During diagnosis and immediate follow up

Although not the focus of the study some of the results can be extrapolated to the diagnosis and initial disease time points. For the HLC assay, all the variables (HLCr, iHLC, uHLC, dHLC, I/U HLCr) are strongly linked to each other and are all produced from a single test. There are different ways these variables can be used depending on the stage of disease. In the work presented in Chapter 5 it was shown how the iHLC and HLCr are important for evaluating beta migrating M-proteins, suggesting that this would be useful both at diagnosis and during follow up of these patients, as proposed previously (Bradwell *et al.,* 2013, Katzmann *et al.,* 2015). Whilst the M-protein is still easily measurable by iHLC, this variable and HLCr can be used to monitor initial response to treatment. Recent advances in therapies have made them very fast acting and the long half life of immunoglobulin, especially IgG at low concentrations, will make serum measurements slower to react than changes in the BM (Kendrick *et al.,* 2017). However, serum measurements can be taken more often and if the iHLC/HLCr remains constant or increases one can surmise that the treatment has not stopped active production of the M-protein. If these variables begin to decrease soon after treatment this can indicate treatment success. The work presented in Chapter 6 showed the ability of MS to accurately isotype the M-protein when compared to IFE. A MALDI-TOF MS system is already being used in place of SPEP/IFE at the Mayo Clinic (Dispenzieri *et al.,* 2022) and as the technology improves and becomes more available this system could start to replace more conventional techniques.

7.3.2 At post-ASCT

Although CTPC detection was not prognostic for PFS at this time point it might still be useful as an independent tool to predict the presence of residual disease and determine disease activity in the bone marrow until CTPC is negative (Chapter 4). As CTPC positivity was always seen in MRD BM positive patients, CTPC positive patients could avoid a bone marrow biopsy as MRD would always be positive in these situations. However, within this group of CTPC positive

patients some were not in CR and therefore conventional techniques and/or HLC and MS assays can be used to determine response in these participants. However, it is possible that the CTPC result is telling us something different than the bone marrow MRD result. For example, it could indicate the ability of MM plasma cells to spread to extra-medullary locations which will have an effect on the clinical outcomes of the patient (Kis *et al.,* 2017, Mishima *et al.,* 2017).

In the study cohort, at the post-ASCT time point, uHLC, HLCr and MS were all prognostic for survival whilst the IMWG criteria and IFE status were not. When comparing directly with the MRD BM result an abnormal HLCr was always accompanied by a positive NGF result, whilst the uHLC value had the highest sensitivity. This has important implications for the monitoring of patients where the HLC variables, and HLCr/uHLC especially, could be used to determine whether a bone marrow biopsy is needed on an individual basis. This suggestion is strengthened by the data in Chapter 4 and Chapter 5, showing that the only significant relationship between MRD status and relevant clinical parameters was with uHLC and HLCr.

7.3.3 At best response and during plateau remission

CTPC status at the time of best or maximum response showed significant PFS benefits (Chapter 4), alongside uHLC (Chapter 5) and MS status (Chapter 6). A recent review of published HLC data suggested clinicians use HLCr to monitor patients once they had reached best response when in CR, and use the dHLC calculation for those in VGPR or below, whilst the iHLC is likely to be abnormal (Rios-Tamayo *et al.,* 2021). In this study it was found that uHLC was the most significant variable once patients reached best response and was also the only significant variable able to predict relapse in the study cohort both as a single measurement at the best response stage and during regular follow up (Chapter 5). Interestingly, whilst the vast majority of ≥VGPR samples had normal iHLC, this is in contrast to uHLC values where the majority of ≥VGPR samples and 45% of CR samples had reduced (i.e. abnormal) levels of uHLC. This suggests that those participants currently grouped as CR by IMWG criteria should be further subdivided according to uHLC level and this could have a prognostic value in the monitoring and treatment of these patients. MS also showed increased sensitivity over the IMWG CR criteria, with 77% of CR samples having a positive MS result, corresponding with recent data from the Mayo Clinic (Dispenzieri *et al.,* 2022). Both of these serum-based methods would be more useful for monitoring CR patients than the blood based CTPC method as only 3/34 (9%) CTPC positive participants were also in CR and therefore represent a very small subset of patients.

Follow up during remission after best response has been reached showed uHLC to be the most significant and earliest marker of relapse in IgG and IgA MM patients. This supports a previous study showing that a decrease in uHLC is seen prior to an increase in iHLC (Chae *et al.,* 2018). A

suggestion can therefore be made for a role for an abnormal uHLC level (below normal) being an early marker of relapse and therefore indication that therapy should be re-started, increased, or changed. Therefore, in patients considered to be in CR especially, it is a recommendation from this study that uHLC should be monitored alongside other clinically relevant parameters and this could eventually lead to the subgrouping of CR patients and the potential to introduce different monitoring schedules.

7.4 Final recommendations

Bone marrow analysis remains the gold standard for MRD evaluation in MM patients. However, the results of this study suggest that blood-based assays can be pre-emptively used to help influence decisions and improve both the patient and laboratory experience. Although CTPC analysis has been shown to be useful in diagnosis (Gonsalves *et al.,* 2014, 2017, Peceliunas *et al.,* 2012, Terpos *et al.,* 2019) and here during follow up, the technology for this technique is less accessible and the technique more cumbersome than the serum assays. Therefore the use of the CTPC assay is only recommended in centres where the NGF BM assay is already established and can be easily adapted. The MS technique is still in its relative infancy, currently only being available in the UK on a research basis. The incredibly high sensitivity of this technique means that long-term follow up and serial measurements are needed to fully elucidate how results can influence patient care in the MRD setting.

From the results of this study it is therefore the HLC assay which has been shown to have the most clinical utility at present and can easily be incorporated into current patient pathways. Although there are limitations to M-protein measurement, such as the long half-life as discussed earlier, the uHLC variable is not measuring the M-protein itself and will therefore be less affected by delayed clearance after tumour lysis. As shown in the results presented in Chapter 5, in treated patients the HLCr is more highly influenced by the uHLC level than the iHLC level, especially in patients who are in VGPR or CR where the iHLC is often normal. For this reason the uHLC and HLCr should always be evaluated together.

From the results presented in this thesis a stepwise algorithm can be suggested to be used during patient follow up post initial treatment would take the following form (see Figure 7.1):

i) If serological response of VGPR or better is reached then the HLC assay should be performed regularly.

ii) If performing serial measurements in the remission stage then if uHLC values decrease (become more abnormal) over multiple measurements then assume the patient is undergoing or about to undergo relapse.

iii) If performing checks at one time point for MRD evaluation then an abnormal HLCr OR uHLC would suggest a positive MRD status and bone marrow assessment can be delayed. If HLC values are normal and a bone marrow thought to be clinically relevant then it should be performed.

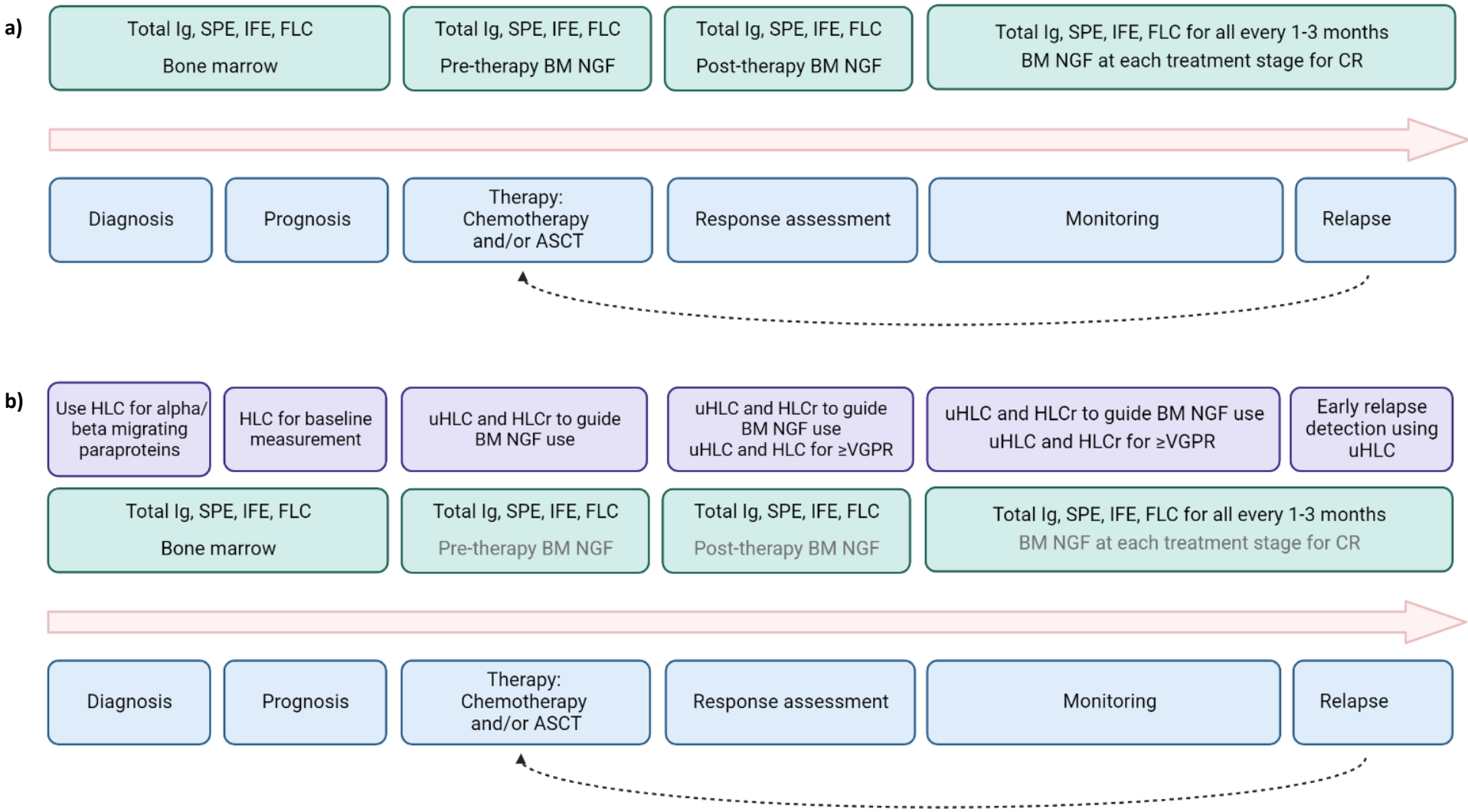


Figure 7.1: Depiction of the multiple myeloma patient pathway showing a) current basic laboratory testing in multiple myeloma and b) proposed additional serological testing using the HLC assay

In Figure 7.1b it is shown where there is the potential for bone marrow assessment to be delayed or avoided through the use of the HLC assay.

7.5 Future work

To prove an algorithm successful, it must be applied to routine clinical scenarios and potentially evaluated in prospective clinical trials. The HLC assay is not currently part of the IMWG recommendations but the significance of uHLC and HLCr in the survival analyses performed here suggest that it would be a useful addition to current monitoring criteria. Although practices will have changed over the intervening years, a 2016 international survey of over 700 laboratories found that heavy/light chain (HLC) analysis was only performed routinely by 1.3% of laboratories (Genzen *et al.,* 2018). The technology needed to run this assay is widely available and already in place in many laboratories, therefore it should not impact laboratories to introduce this test into the patient pathway. The potential uses of the HLC assay needs to be highlighted to not only clinicians and scientists but also to patients. It is a test which produces easy to understand results, and knowledge of the M-protein status is known to improve the psychological well-being of patients, as stated in a recent survey (Janssens *et al.,* 2021).

Further work to be done with follow up patients is to gather serial HLC measurements and determine exactly how a change in levels over time effects the outcome of patients and what % change can be used to signify relapse and/or reconstitution. Michallet *et al.* (2018) compared response assessment results using the current IMWG criteria vs their own HLC criteria (based on dHLC/HLCr). They found the HLC criteria gave a more accurate assessment of response when compared to MRD status, especially those in VGPR or better. This is supported by the fact that the HLC results correlated better with MRD assessment than the IMWG criteria and/or IFE results. Therefore it will be useful to create HLC criteria which can be used to influence decision making. There are currently two studies which have made clinical decisions based on MRD results, both using NGS technology, showing that adapting treatment based on MRD status has good short-term outcomes but longer term results will be published in the coming years (Derman *et al.,* 2021, Costa *et al.*, 2022,).

A population-based study found infection to be the main cause of death for patients at 6 months and 12 months after diagnosis (Rios-Tamayo *et al.,* 2018) and a later study found severe uHLC suppression to be a risk factor for infection and early death in NDMM patients. As uHLC represents both the suppression and reconstitution of a patients normal immune activity it will be interesting to correlate exact uHLC values to infection and to see if uHLC abnormal patients are suffering from infections more than their counterparts with a normal uHLC. This HLC work should be compared to analysis of total immunoglobulins, presence of immunoparesis and

functional antibody responses to vaccines. This could help decide on the type and timing of prophylactic immune therapy, such as vaccines (Sanchez-Ramon *et al.* 2016).

Further work is required to fully understand the potential of the MS technique. In this study a high number of participants were MS positive throughout monitoring and it will be important to determine how such a sensitive method can be used to provide clinically useful information to clinicians to give the best balance between treatment and quality of life for patients. This study has also highlighted the potential use of MS in patients previously thought to be oligosecretory. Please see individual results Chapters for further future work.

7.6 Limitations

This study was limited by low sample numbers, especially in the bone marrow transplant group (Group 1). There are several reasons for this including the shielding of MM patients, redistributing of resources and cancelling of bone marrow transplants during the course of this study due to the COVID-19 pandemic. Because of this, there was an over-representation of Group 2 patients in the study cohort who did not always undergo rigorous treatment and therefore did not reach a response depth where minimal residual disease evaluation was needed. The study cohort underwent heterogenous treatment regimes which may have affected outcomes and was not accounted for in analysis. However, this reflects the different cohorts one would find in a routine hospital setting so allows for the determination of the effect of these assays in a true MM cohort. This study did not have access to imaging results for the participants which would allow for the correlation of the study results to the presence of either focal lesions or extramedullary disease. Therefore this study could not draw any conclusions on how blood-based assays may help to overcome the limitations of bone marrow testing in these instances. Please see individual results Chapters for further limitations.

7.7 Conclusions

This study has shown the significance of blood-based assays at the pre-ASCT, post-ASCT, best response and pre-relapse disease stages. No single assay or means of assessment can be used during the journey of a myeloma patient for either diagnosis, prognosis or follow up. The IMWG guidelines set out which tests to use and how throughout the patient journey (Kumar *et al.,* 2016). Although the latest guidelines did not include HLC analysis they did suggest collecting more information on its use at each of the disease stages. This study has focused on treated

patients, looking at response evaluation and prognosis at specific treatment times and during follow up. This relatively small study has shown the potential use of the HLC assay in monitoring patients and determining next steps and has started a promising investigation into the use of the MS technique in a similar way. The HLC assay is a validated test, available to the majority of laboratories. Working on this study during the COVID-19 pandemic highlighted how important it is that immunocompromised patients with MM can be accurately monitored outside of clinic settings. These blood-based assays allow for close monitoring of patients away from the hospital, whilst potentially reducing their need to be subjected to unnecessary and painful procedures – benefiting both the multiple myeloma care teams and multiple myeloma patients.

References

Abe, M. (2011) 'Targeting the interplay between myeloma cells and the bone marrow microenvironment in myeloma', *Int J Hematol,* 94(4), pp. 334-343. Available at: https://doi.org/10.1007/s12185-011-0949-x.

Al Hamed, R. *et al.* (2019) 'Current status of autologous stem cell transplantation for multiple myeloma', *Blood Cancer Journal,* 9(4), pp. 44. Available at: https://doi.org/10.1038/s41408-019-0205-9.

Alapat, D. *et al.* (2012) 'Diagnostic usefulness and prognostic impact of CD200 expression in lymphoid malignancies and plasma cell myeloma', *Am J Clin Pathol,* 137(1), pp. 93-100. Available at: https://doi.org/10.1309/AJCP59UORCYZEVQO.

Arana, P. *et al.* (2018) 'Prognostic value of antigen expression in multiple myeloma: a PETHEMA/GEM study on 1265 patients enrolled in four consecutive clinical trials', *Leukemia,* 32(4), pp. 971-978. Available at: https://doi.org/10.1038/leu.2017.320.

Arroz, M. *et al.* (2016) 'Consensus guidelines on plasma cell myeloma minimal residual disease analysis and reporting', *Cytometry B Clin Cytom,* 90(1), pp. 31-9. Available at: https://doi.org/10.1002/cyto.b.21228.

Attal, M. *et al.* (1996) 'A prospective, randomized trial of autologous bone marrow transplantation and chemotherapy in multiple myeloma. Intergroupe Francais du Myelome', *N Engl J Med,* 335(2), pp. 91-7. Available at: https://doi.org/10.1056/nejm199607113350204.

Attal, M. *et al.* (2017) 'Lenalidomide, Bortezomib, and Dexamethasone with Transplantation for Myeloma', *N Engl J Med,* 376(14), pp. 1311-1320. Available at: https://doi.org/10.1056/NEJMoa1611750.

Avet-Loiseau, H. *et al.* (2015) 'Evaluation of Minimal Residual Disease (MRD) By Next Generation Sequencing (NGS) Is Highly Predictive of Progression Free Survival in the IFM/DFCI 2009 Trial', *Blood,* 126(23), pp. 191.

Avet-Loiseau, H. *et al.* (2020) 'Minimal Residual Disease Status as a Surrogate Endpoint for Progression-free Survival in Newly Diagnosed Multiple Myeloma Studies: A Meta-analysis', *Clin Lymphoma Myeloma Leuk,* 20(1), pp. e30-e37. Available at: https://doi.org/10.1016/j.clml.2019.09.622.

Barlogie, B. *et al.* (2014) 'Curing myeloma at last: defining criteria and providing the evidence', *Blood,* 124(20), pp. 3043-51. Available at: https://doi.org/10.1182/blood-2014-07-552059.

Barnidge, D. R. *et al.* (2014) 'Using mass spectrometry to monitor monoclonal immunoglobulins in patients with a monoclonal gammopathy', *J Proteome Res,* 13(3), pp. 1419-27. Available at: https://doi.org/10.1021/pr400985k.

Batinić, J. *et al.* (2015) 'Immunoglobulin heavy/light chain analysis enhances the detection of residual disease and monitoring of multiple myeloma patients', *Croat Med J,* 56(3), pp. 263-71. Available at: https://doi.org/10.3325/cmj.2015.56.263.

Bayly, E. *et al.* (2020) 'Validation of a modified pre-lysis sample preparation technique for flow cytometric minimal residual disease assessment in multiple myeloma, chronic lymphocytic leukemia, and B-non Hodgkin lymphoma', *Cytometry Part B: Clinical Cytometry,* 98(5), pp. 385-398. Available at: https://doi.org/10.1002/cyto.b.21893.

BDWG, B. D. W. G. (2001) 'Biomarkers and surrogate endpoints: preferred definitions and conceptual framework', *Clin Pharmacol Ther,* 69(3), pp. 89-95. Available at: https://doi.org/10.1067/mcp.2001.113989.

Beaumont-Epinette, M. P. *et al.* (2015) 'Heavy/light chain specific immunoglobulin ratios provides no additional information than serum proteins electrophoresis and immunofixation for the diagnosis and the follow-up of intact immunoglobulin multiple myeloma patients', *Pathol Biol (Paris),* 63(4-5), pp. 215-21. Available at: https://doi.org/10.1016/j.patbio.2015.06.001.

Bergen, H. R. *et al.* (2016) 'Clonotypic Light Chain Peptides Identified for Monitoring Minimal Residual Disease in Multiple Myeloma without Bone Marrow Aspiration', *Clin Chem,* 62(1), pp. 243-51. Available at: https://doi.org/10.1373/clinchem.2015.242651.

Blade, J. *et al.* (2011) 'Soft-tissue plasmacytomas in multiple myeloma: incidence, mechanisms of extramedullary spread, and treatment approach', *J Clin Oncol,* 29(28), pp. 3805-12. Available at: https://doi.org/10.1200/jco.2011.34.9290.

Bolomsky, A. and Young, R.M. (2022) ' Pathogenic signaling in multiple myeloma', *Semin Oncol.*, 49(1), pp. 27-40. Available at: doi: 10.1053/j.seminoncol.2022.01.005.

Boyle, E. M. *et al.* (2014a) 'Understanding the multiple biological aspects leading to myeloma', *Haematologica,* 99(4), pp. 605-612. Available at: https://doi.org/10.3324/haematol.2013.097907.

Boyle, E. M. *et al.* (2014b) 'IgA kappa/IgA lambda heavy/light chain assessment in the management of patients with IgA myeloma', *Cancer,* 120(24), pp. 3952-7. Available at: https://doi.org/10.1002/cncr.28946.

Bradwell, A. R. *et al.* (2003) 'Serum test for assessment of patients with Bence Jones myeloma', *Lancet,* 361(9356), pp. 489-91. Available at: https://doi.org/10.1016/s0140-6736(03)12457-9.

Bradwell, A. R. *et al.* (2001) 'Highly sensitive, automated immunoassay for immunoglobulin free light chains in serum and urine', *Clin Chem,* 47(4), pp. 673-80.

Bradwell, A. R. *et al.* (2009) 'Assessment of monoclonal gammopathies by nephelometric measurement of individual immunoglobulin kappa/lambda ratios', *Clin Chem,* 55(9), pp. 1646-55. Available at: https://doi.org/10.1373/clinchem.2009.123828.

Brennan, S. K. and Matsui, W. (2009) 'Cancer stem cells: controversies in multiple myeloma', *J*

Mol Med (Berl), 87(11), pp. 1079-85. Available at: https://doi.org/10.1007/s00109-009-0531-7.

Bruggemann, M., Raff, T. and Kneba, M. (2012) 'Has MRD monitoring superseded other prognostic factors in adult ALL?', *Blood,* 120(23), pp. 4470-81. Available at: https://doi.org/10.1182/blood-2012-06-379040.

Burgos, L. *et al.* (2020) 'Measurable residual disease in multiple myeloma: ready for clinical practice?', *J Hematol Oncol,* 13(1), pp. 82. Available at: https://doi.org/10.1186/s13045-020-00911-4.

Campbell, J.P. *et al.* (2017) 'Active multiple myeloma suppresses and typically eliminates coexisting MGUS', *Br J Cancer*., 117(6), pp. 835-839. Available at: https://doi:10.1038/bjc.2017.229.

Campbell, J.P. *et al.* (2017b) 'Response comparison of multiple myeloma and monoclonal gammopathy of undetermined significance to the same anti-myeloma therapy: a retrospective cohort study', *Lancet Haematol*., 4(12), pp. e584-e594. Available at: https://doi:10.1016/S2352-3026(17)30209-0.

Campbell, L. *et al.* (2019) 'Serological normalisation as a surrogate marker for minimal residual disease negativity in multiple myeloma', *Br J Haematol,* 185(4), pp. 775-778. Available at: https://doi.org/10.1111/bjh.15615.

Campbell, L. *et al.* (2021) 'Using quantitative immunoprecipitation mass spectrometry (QIP-MS) to identify low level monoclonal proteins', *Clin Biochem,* 95, pp. 81-83. Available at: https://doi.org/10.1016/j.clinbiochem.2021.05.012.

Campbell, L. *et al.* (2020) 'In-house age-specific reference ranges for free light chains measured on the SPAPlus® analyser', *Ann Clin Biochem,* 57(2), pp. 138-143. Available at: https://doi.org/10.1177/0004563219899421.

Carulli, G. *et al.* (2019) 'Assessment of response to therapy in multiple myeloma by multiparameter flow cytometry. Usefulness of an eight-color single tube with monoclonal antibodies in dried formulation', *Clin Ter,* 170(5), pp. e352-e356. Available at: https://doi.org/10.7417/ct.2019.2159.

Chakraborty, R. *et al.* (2016) 'Risk stratification in myeloma by detection of circulating plasma cells prior to autologous stem cell transplantation in the novel agent era', *Blood cancer journal,* 6(12), pp. e512-e512. Available at: https://doi.org/10.1038/bcj.2016.117.

Chakraborty, R. *et al.* (2017) 'Serial measurements of circulating plasma cells before and after induction therapy have an independent prognostic impact in patients with multiple myeloma undergoing upfront autologous transplantation', *Haematologica,* 102(8), pp. 1439-1445. Available at: https://doi.org/10.3324/haematol.2017.166629.

Chakraborty, R. *et al.* (2020) 'Characterisation and prognostic impact of immunoparesis in relapsed multiple myeloma', *Br J Haematol,* 189(6), pp. 1074-1082. Available at: https://doi.org/10.1111/bjh.16488.

Chanan-Khan, A. A. and Giralt, S. (2010) 'Importance of achieving a complete response in multiple myeloma, and the impact of novel agents', *J Clin Oncol,* 28(15), pp. 2612-24. Available at: https://doi.org/10.1200/jco.2009.25.4250.

Chatterjee, T., Mallhi, R. S. and Venkatesan, S. (2016) 'Minimal residual disease detection using flow cytometry: Applications in acute leukemia', *Med J Armed Forces India,* 72(2), pp. 152-6. Available at: https://doi.org/10.1016/j.mjafi.2016.02.002.

Chee, C. E. *et al.* (2009) 'The importance of bone marrow examination in determining complete response to therapy in patients with multiple myeloma', *Blood,* 114(13), pp. 2617-8. Available at: https://doi.org/10.1182/blood-2009-01-198788.

Chen, Y. *et al.* (2017) 'Improved survival in Medicare patients with multiple myeloma: findings from a large nationwide and population-based cohort', *Med Oncol,* 34(9), pp. 153. Available at: https://doi.org/10.1007/s12032-017-1001-7.

Child, J. A. *et al.* (2003) 'High-Dose Chemotherapy with Hematopoietic Stem-Cell Rescue for Multiple Myeloma', *New England Journal of Medicine,* 348(19), pp. 1875-1883. Available at: https://doi.org/10.1056/NEJMoa022340.

Ching, T. *et al.* (2020) 'Analytical evaluation of the clonoSEQ Assay for establishing measurable (minimal) residual disease in acute lymphoblastic leukemia, chronic lymphocytic leukemia, and multiple myeloma', *BMC cancer,* 20(1), pp. 612-612. Available at: https://doi.org/10.1186/s12885-020-07077-9.

Corso, A. and Mangiacavalli, S. (2017) 'Non-Secretory Myeloma: Ready for a new Definition?', *Mediterranean journal of hematology and infectious diseases,* 9(1), pp. e2017053-e2017053. Available at: https://doi.org/10.4084/MJHID.2017.053.

Costa, L. J. *et al.* (2022) 'Daratumumab, Carfilzomib, Lenalidomide, and Dexamethasone With Minimal Residual Disease Response-Adapted Therapy in Newly Diagnosed Multiple Myeloma', *J Clin Oncol,* 40(25), pp. 2901-2912. Available at: https://doi.org/10.1200/JCO.21.01935.

CRUK (2022) *Myeloma Statistics.* Available at: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/myeloma#heading-One (Accessed: September 2022).

D'Auria, F. *et al.* (2018) 'Heavy/light chain ratio for the assessment of minimal residual disease in myeloma patients achieving complete response', *Br J Haematol,* 181(4), pp. 550-552. Available at: https://doi.org/10.1111/bjh.14662.

Da Vià, M. C. *et al.* (2020) 'A Journey Through Myeloma Evolution: From the Normal Plasma Cell to Disease Complexity', *HemaSphere,* 4(6), pp. e502-e502. Available at: https://doi.org/10.1097/HS9.0000000000000502.

Davies, F. E. *et al.* (2001) 'The impact of attaining a minimal disease state after high-dose melphalan and autologous transplantation for multiple myeloma', *Br J Haematol,* 112(3), pp. 814-9. Available at: https://doi.org/10.1046/j.1365-2141.2001.02530.x.

de Tute, R. M. *et al.* (2007) 'A single-tube six-colour flow cytometry screening assay for the detection of minimal residual disease in myeloma', *Leukemia,* 21(9), pp. 2046-9. Available at: https://doi.org/10.1038/sj.leu.2404815.

Derman, B. A., Jakubowiak, A. J. and Thompson, M. A. (2022) 'Clinician survey regarding measurable residual disease-guided decision-making in multiple myeloma', *Blood Cancer J,* 12(7), pp. 108. Available at: https://doi.org/10.1038/s41408-022-00705-6.

Derman, B. A. *et al.* (2021) 'Measurable residual disease assessed by mass spectrometry in peripheral blood in multiple myeloma in a phase II trial of carfilzomib, lenalidomide, dexamethasone and autologous stem cell transplantation', *Blood Cancer Journal,* 11(2), pp. 19. Available at: https://doi.org/10.1038/s41408-021-00418-2.

Dimopoulos, M. A. *et al.* (2019) 'Oral ixazomib maintenance following autologous stem cell transplantation (TOURMALINE-MM3): a double-blind, randomised, placebo-controlled phase 3 trial', *Lancet,* 393(10168), pp. 253-264. Available at: https://doi.org/10.1016/s0140-6736(18)33003-4.

Dispenzieri, A. *et al.* (2022) 'Mass-Fix better predicts for PFS and OS than standard methods among multiple myeloma patients participating on the STAMINA trial (BMT CTN 0702 /07LT)', *Blood Cancer J,* 12(2), pp. 27. Available at: https://doi.org/10.1038/s41408-022-00624-6.

Dispenzieri, A. *et al.* (2009) 'International Myeloma Working Group guidelines for serum-free light chain analysis in multiple myeloma and related disorders', *Leukemia,* 23(2), pp. 215-24. Available at: https://doi.org/10.1038/leu.2008.307.

Dispenzieri, A. *et al.* (2006) 'Absolute values of immunoglobulin free light chains are prognostic in patients with primary systemic amyloidosis undergoing peripheral blood stem cell transplantation', *Blood,* 107(8), pp. 3378-83. Available at: https://doi.org/10.1182/blood-2005-07-2922.

Dold, S. M. *et al.* (2020) 'Validated single-tube multiparameter flow cytometry approach for the assessment of minimal residual disease in multiple myeloma', *Haematologica,* 105(10), pp. e523. Available at: https://doi.org/10.3324/haematol.2019.238394.

Drayson, M. *et al.* (2006) 'Effects of paraprotein heavy and light chain types and free light chain load on survival in myeloma: an analysis of patients receiving conventional-dose chemotherapy in Medical Research Council UK multiple myeloma trials', *Blood,* 108(6), pp. 2013-9. Available at: https://doi.org/10.1182/blood-2006-03-008953.

Drayson, M. *et al.* (2001) 'Serum free light-chain measurements for identifying and monitoring patients with nonsecretory multiple myeloma', *Blood,* 97(9), pp. 2900-2.

Durie, B. G. *et al.* (2006) 'International uniform response criteria for multiple myeloma', *Leukemia,* 20(9), pp. 1467-73. Available at: https://doi.org/10.1038/sj.leu.2404284.

Espiño, M. *et al.* (2017) 'Involved/uninvolved heavy/light chain index can predict progression in transplanted multiple myeloma patients', *Bone marrow transplantation,* 52(8), pp.

1206-1207. Available at: https://doi.org/10.1038/bmt.2017.97.

Eveillard, M. *et al.* (2021) 'Using MALDI-TOF mass spectrometry in peripheral blood for the follow up of newly diagnosed multiple myeloma patients treated with daratumumab-based combination therapy', *Clin Chim Acta,* 516, pp. 136-141. Available at: https://doi.org/10.1016/j.cca.2021.01.021.

Eveillard, M. *et al.* (2020) 'Comparison of MALDI-TOF mass spectrometry analysis of peripheral blood and bone marrow-based flow cytometry for tracking measurable residual disease in patients with multiple myeloma', *Br J Haematol,* 189(5), pp. 904-907. Available at: https://doi.org/10.1111/bjh.16443.

Facon, T. *et al.* (2019) 'Carfilzomib or bortezomib with melphalan-prednisone for transplant-ineligible patients with newly diagnosed multiple myeloma', *Blood,* 133(18), pp. 1953-1963. Available at: https://doi.org/10.1182/blood-2018-09-874396.

Facon, T. *et al.* (2007) 'Melphalan and prednisone plus thalidomide versus melphalan and prednisone alone or reduced-intensity autologous stem cell transplantation in elderly patients with multiple myeloma (IFM 99-06): a randomised trial', *Lancet,* 370(9594), pp. 1209-18. Available at: https://doi.org/10.1016/S0140-6736(07)61537-2.

Fermand, J. P. *et al.* (2005) 'High-dose therapy and autologous blood stem-cell transplantation compared with conventional treatment in myeloma patients aged 55 to 65 years: long-term results of a randomized control trial from the Group Myelome-Autogreffe', *J Clin Oncol,* 23(36), pp. 9227-33. Available at: https://doi.org/10.1200/jco.2005.03.0551.

Fernandez de Larrea, C. *et al.* (2014) 'Pattern of relapse and progression after autologous SCT as upfront treatment for multiple myeloma', *Bone Marrow Transplant,* 49(2), pp. 223-7. Available at: https://doi.org/10.1038/bmt.2013.150.

Ferrero, S. *et al.* (2015) 'Long-term results of the GIMEMA VEL-03-096 trial in MM patients receiving VTD consolidation after ASCT: MRD kinetics' impact on survival', *Leukemia,* 29(3), pp. 689-95. Available at: https://doi.org/10.1038/leu.2014.219.

Fitzmaurice, C. *et al.* (2017) 'Global, Regional, and National Cancer Incidence, Mortality, Years of Life Lost, Years Lived With Disability, and Disability-Adjusted Life-years for 32 Cancer Groups, 1990 to 2015: A Systematic Analysis for the Global Burden of Disease Study', *JAMA Oncol,* 3(4), pp. 524-548. Available at: https://doi.org/10.1001/jamaoncol.2016.5688.

Flanders, A., Stetler-Stevenson, M. and Landgren, O. (2013) 'Minimal residual disease testing in multiple myeloma by flow cytometry: major heterogeneity', *Blood,* 122(6), pp. 1088-9. Available at: https://doi.org/10.1182/blood-2013-05-506170.

Flores-Montero, J. *et al.* (2017) 'Next Generation Flow for highly sensitive and standardized detection of minimal residual disease in multiple myeloma', *Leukemia,* 31(10), pp. 2094-2103. Available at: https://doi.org/10.1038/leu.2017.29.

Foulk, B. *et al.* (2018) 'Enumeration and characterization of circulating multiple myeloma cells in

patients with plasma cell disorders', *British Journal of Haematology,* 180(1), pp. 71-81. Available at: https://doi.org/https://doi.org/10.1111/bjh.15003.

Foureau, D. *et al.* (2021) 'Comparison of mass spectrometry and flow cytometry in measuring minimal residual disease in multiple myeloma', *Cancer Med,* 10(20), pp. 6933-6936. Available at: https://doi.org/10.1002/cam4.4254.

Fulciniti, M., Munshi, N. C. and Martinez-Lopez, J. (2015) 'Deep Response in Multiple Myeloma: A Critical Review', *Biomed Res Int,* 2015, pp. 832049. Available at: https://doi.org/10.1155/2015/832049.

Galtseva, I. V. *et al.* (2018) 'Minimal residual disease in multiple myeloma: Benefits of flow cytometry', *Int J Lab Hematol,* 40(1), pp. 12-20. Available at: https://doi.org/10.1111/ijlh.12757.

Gandolfi, S., Prada, C. P. and Richardson, P. G. (2018) 'How I treat the young patient with multiple myeloma', *Blood,* 132(11), pp. 1114-1124. Available at: https://doi.org/10.1182/blood-2017-05-693606.

García de Veas Silva, J. L. *et al.* (2016) 'Immunoglobulin Heavy/Light Chain Ratio and Suppression of the Myeloma Non-Involved Heavy/Light Chain Pair Isotype Correlate with Poor Survival in Multiple Myeloma at Diagnosis and after Autologous Stem Cell Transplant', *Blood,* 128(22), pp. 5647.

Gay, F. *et al.* (2011) 'Complete response correlates with long-term progression-free and overall survival in elderly myeloma treated with novel agents: analysis of 1175 patients', *Blood,* 117(11), pp. 3025-31. Available at: https://doi.org/10.1182/blood-2010-09-307645.

Genzen, J. R. *et al.* (2018) 'Screening and Diagnosis of Monoclonal Gammopathies: An International Survey of Laboratory Practice', *Arch Pathol Lab Med,* 142(4), pp. 507-515. Available at: https://doi.org/10.5858/arpa.2017-0128-CP.

Goldschmidt, H. *et al.* (2000) 'Multiple myeloma and renal failure', *Nephrology Dialysis Transplantation,* 15(3), pp. 301-304. Available at: https://doi.org/10.1093/ndt/15.3.301.

Gonsalves, W. I. *et al.* (2014a) 'Quantification of clonal circulating plasma cells in relapsed multiple myeloma', *Br J Haematol,* 167(4), pp. 500-5. Available at: https://doi.org/10.1111/bjh.13067.

Gonsalves, W. I. *et al.* (2014b) 'Quantification of clonal circulating plasma cells in newly diagnosed multiple myeloma: implications for redefining high-risk myeloma', *Leukemia,* 28(10), pp. 2060-5. Available at: https://doi.org/10.1038/leu.2014.98.

González, D. *et al.* (2007) 'Immunoglobulin gene rearrangements and the pathogenesis of multiple myeloma', *Blood,* 110(9), pp. 3112-21. Available at: https://doi.org/10.1182/blood-2007-02-069625.

Goswami, C. *et al.* (2019) 'Staging System to Predict the Risk of Relapse in Multiple Myeloma Patients Undergoing Autologous Stem Cell Transplantation', *Frontiers in oncology,* 9, pp.

633-633. Available at: https://doi.org/10.3389/fonc.2019.00633.

Greaves, M. F. (2000) 'Evolution, immune response, and cancer', *Lancet,* 356(9234), pp. 1034.

Greil, C. *et al.* (2017) 'The serum heavy/light chain immunoassay: A valuable tool for sensitive paraprotein assessment, risk, and disease monitoring in monoclonal gammopathies', *Eur J Haematol,* 99(5), pp. 449-458. Available at: https://doi.org/10.1111/ejh.12958.

Gu, J. *et al.* (2018) 'Longitudinal Flow Cytometry Identified "Minimal Residual Disease" (MRD) Evolution Patterns for Predicting the Prognosis of Patients with Transplant-Eligible Multiple Myeloma', *Biol Blood Marrow Transplant,* 24(12), pp. 2568-2574. Available at: https://doi.org/10.1016/j.bbmt.2018.07.040.

Guikema, J. E. *et al.* (2003) 'CD27 is heterogeneously expressed in multiple myeloma: low CD27 expression in patients with high-risk disease', *Br J Haematol,* 121(1), pp. 36-43. Available at: https://doi.org/10.1046/j.1365-2141.2003.04260.x.

Gupta, S. *et al.* (2018) 'Flow cytometric aberrancies in plasma cell myeloma and MGUS - correlation with laboratory parameters', *Cytometry B Clin Cytom,* 94(3), pp. 500-508. Available at: https://doi.org/10.1002/cyto.b.21624.

Hanamura, I. (2022) 'Multiple myeloma with high-risk cytogenetics and its treatment approach', *International Journal of Hematology,* 115(6), pp. 762-777. Available at: https://doi.org/10.1007/s12185-022-03353-5.

Harousseau, J. L. *et al.* (2010) 'Better quality of response to lenalidomide plus dexamethasone is associated with improved clinical outcomes in patients with relapsed or refractory multiple myeloma', *Haematologica,* 95(10), pp. 1738-44. Available at: https://doi.org/10.3324/haematol.2009.015917.

Harutyunyan, N. M. *et al.* (2016) 'Levels of uninvolved immunoglobulins predict clinical status and progression-free survival for multiple myeloma patients', *Br J Haematol,* 174(1), pp. 81-7. Available at: https://doi.org/10.1111/bjh.14026.

Heaney, J. L. J. *et al.* (2018) 'Characterisation of immunoparesis in newly diagnosed myeloma and its impact on progression-free and overall survival in both old and recent myeloma trials', *Leukemia,* 32(8), pp. 1727-1738. Available at: https://doi.org/10.1038/s41375-018-0163-4.

Hemminki, K. *et al.* (2021) ‘Epidemiology, genetics and treatment of multiple myeloma and precursor diseases. *Int. J. Cancer*, 149(12), pp. 1980-1996. Available at: https:// doi:10.1002/ijc.33762

Hill, P. G. *et al.* (2006) 'Serum free light chains: an alternative to the urine Bence Jones proteins screening test for monoclonal gammopathies', *Clin Chem,* 52(9), pp. 1743-8. Available at: https://doi.org/10.1373/clinchem.2006.069104.

Holstein, S. A. *et al.* (2018) 'BMT CTN Myeloma Intergroup Workshop on Minimal Residual Disease and Immune Profiling: Summary and Recommendations from the Organizing Committee', *Biol Blood Marrow Transplant,* 24(4), pp. 641-648. Available at:

https://doi.org/10.1016/j.bbmt.2017.12.774.

Hosen, N. *et al.* (2012) 'CD48 as a novel molecular target for antibody therapy in multiple myeloma', *Br J Haematol,* 156(2), pp. 213-24. Available at: https://doi.org/10.1111/j.1365-2141.2011.08941.x.

Huhn, S. *et al.* (2017) 'Circulating tumor cells as a biomarker for response to therapy in multiple myeloma patients treated within the GMMG-MM5 trial', *Bone Marrow Transplant,* 52(8), pp. 1194-1198. Available at: https://doi.org/10.1038/bmt.2017.91.

Hutchison, C. A. *et al.* (2011) 'Early reduction of serum-free light chains associates with renal recovery in myeloma kidney', *J Am Soc Nephrol,* 22(6), pp. 1129-36. Available at: https://doi.org/10.1681/ASN.2010080857.

Hutchison, C. A. *et al.* (2008) 'Serum free light chain measurement aids the diagnosis of myeloma in patients with severe renal failure', *BMC Nephrol,* 9, pp. 11. Available at: https://doi.org/10.1186/1471-2369-9-11.

IMWG, I. M. W. G. (2003) 'Criteria for the classification of monoclonal gammopathies, multiple myeloma and related disorders: a report of the International Myeloma Working Group', *Br J Haematol,* 121(5), pp. 749-57.

Jakubowiak, A. (2012) 'Management strategies for relapsed/refractory multiple myeloma: current clinical perspectives', *Semin Hematol,* 49 Suppl 1, pp. S16-32. Available at: https://doi.org/10.1053/j.seminhematol.2012.05.003.

Janeway CA, T. P., Walport M et al. (2001) *Immunobiology: The Immune System in Health and Disease.* New York.

Jannetto, P. J. and Fitzgerald, R. L. (2016) 'Effective Use of Mass Spectrometry in the Clinical Laboratory', *Clin Chem,* 62(1), pp. 92-8. Available at: https://doi.org/10.1373/clinchem.2015.248146.

Janssens, R. *et al.* (2021) 'Patient Preferences for Multiple Myeloma Treatments: A Multinational Qualitative Study', *Front Med (Lausanne),* 8, pp. 686165. Available at: https://doi.org/10.3389/fmed.2021.686165.

Joshi, R. *et al.* (2008) 'Bone marrow trephine combined with immunohistochemistry is superior to bone marrow aspirate in follow-up of myeloma patients', *J Clin Pathol,* 61(2), pp. 213-6. Available at: https://doi.org/10.1136/jcp.2007.049130.

Kantarjian, H. *et al.* (2008) 'Monitoring the response and course of chronic myeloid leukemia in the modern era of BCR-ABL tyrosine kinase inhibitors: practical advice on the use and interpretation of monitoring methods', *Blood,* 111(4), pp. 1774-80. Available at: https://doi.org/10.1182/blood-2007-09-110189.

Kastritis, E. *et al.* (2014) 'Preserved levels of uninvolved immunoglobulins are independently associated with favorable outcome in patients with symptomatic multiple myeloma', *Leukemia,* 28(10), pp. 2075-9. Available at: https://doi.org/10.1038/leu.2014.110.

Katzmann, J. A. *et al.* (2013) 'Suppression of uninvolved immunoglobulins defined by heavy/light

chain pair suppression is a risk factor for progression of MGUS', *Leukemia,* 27(1), pp. 208-212. Available at: https://doi.org/10.1038/leu.2012.189.

Kendrick, F. *et al.* (2017) 'Analysis of a Compartmental Model of Endogenous Immunoglobulin G Metabolism with Application to Multiple Myeloma', *Front Physiol,* 8, pp. 149. Available at: https://doi.org/10.3389/fphys.2017.00149.

Kim, J. *et al.* (2007) 'Kinetics of FcRn-mediated recycling of IgG and albumin in human: pathophysiology and therapeutic implications using a simplified mechanism-based model', *Clin Immunol,* 122(2), pp. 146-55. Available at: https://doi.org/10.1016/j.clim.2006.09.001.

Kis, O. *et al.* (2017) 'Circulating tumour DNA sequence analysis as an alternative to multiple myeloma bone marrow aspirates', *Nat Commun,* 8, pp. 15086. Available at: https://doi.org/10.1038/ncomms15086.

Klein, E. M. *et al.* (2021) 'Prognostic Impact of Serum Free Light Chain Ratio Normalization in Patients with Multiple Myeloma Treated within the GMMG-MM5 Trial', *Cancers (Basel),* 13(19). Available at: https://doi.org/10.3390/cancers13194856.

Kohlhagen, M. *et al.* (2020) 'Automation and validation of a MALDI-TOF MS (Mass-Fix) replacement of immunofixation electrophoresis in the clinical lab', *Clin Chem Lab Med*. Available at: https://doi.org/10.1515/cclm-2020-0581.

Korde, N. *et al.* (2015) 'Treatment With Carfilzomib-Lenalidomide-Dexamethasone With Lenalidomide Extension in Patients With Smoldering or Newly Diagnosed Multiple Myeloma', *JAMA Oncol,* 1(6), pp. 746-54. Available at: https://doi.org/10.1001/jamaoncol.2015.2010.

Koulieris, E. *et al.* (2012) 'Ratio of involved/uninvolved immunoglobulin quantification by Hevylite™ assay: clinical and prognostic impact in multiple myeloma', *Experimental Hematology & Oncology,* 1(1), pp. 9. Available at: https://doi.org/10.1186/2162-3619-1-9.

Kraj, M. *et al.* (2012) 'Comparison of immunoglobulin free light chain (FLC), heavy chain/light chain (HLC) assays and immunofixation (IFE) in assessment of remission in multiple myeloma', *Acta Haematologica Polonica,* 43(2), pp. 122-131.

Kraj, M. *et al.* (2004) 'C-kit receptor (CD117) expression on plasma cells in monoclonal gammopathies', *Leuk Lymphoma,* 45(11), pp. 2281-9. Available at: https://doi.org/10.1080/10428190412331283279.

Kriegsmann, K. *et al.* (2020) 'Comparison of NGS and MFC Methods: Key Metrics in Multiple Myeloma MRD Assessment', *Cancers,* 12(8), pp. 2322. Available at: https://doi.org/10.3390/cancers12082322.

Kristinsson, S. Y., Anderson, W. F. and Landgren, O. (2014) 'Improved long-term survival in multiple myeloma up to the age of 80 years', *Leukemia,* 28(6), pp. 1346-8. Available at: https://doi.org/10.1038/leu.2014.23.

Kumar, S. *et al.* (2012) 'Trisomies in multiple myeloma: impact on survival in patients with high-risk cytogenetics', *Blood,* 119(9), pp. 2100-5. Available at: https://doi.org/10.1182/blood-2011-11-390658.

Kumar, S. *et al.* (2016) 'International Myeloma Working Group consensus criteria for response and minimal residual disease assessment in multiple myeloma', *Lancet Oncol,* 17(8), pp. e328-e346. Available at: https://doi.org/10.1016/s1470-2045(16)30206-6.

Kumar, S. *et al.* (2004) 'Cell proliferation of myeloma plasma cells: comparison of the blood and marrow compartments', *Am J Hematol,* 77(1), pp. 7-11. Available at: https://doi.org/10.1002/ajh.20125.

Kumar, S. K. *et al.* (2014) 'Continued improvement in survival in multiple myeloma: changes in early mortality and outcomes in older patients', *Leukemia,* 28(5), pp. 1122-8. Available at: https://doi.org/10.1038/leu.2013.313.

Kyle, R. A. *et al.* (2003) 'Review of 1027 patients with newly diagnosed multiple myeloma', *Mayo Clin Proc,* 78(1), pp. 21-33. Available at: https://doi.org/10.4065/78.1.21.

Kyle, R. A. *et al.* (2018) 'Long-Term Follow-up of Monoclonal Gammopathy of Undetermined Significance', *N Engl J Med,* 378(3), pp. 241-249. Available at: https://doi.org/10.1056/NEJMoa1709974.

Kyle, R. A. and Rajkumar, S. V. (2009) 'Criteria for diagnosis, staging, risk stratification and response assessment of multiple myeloma', *Leukemia,* 23(1), pp. 3-9. Available at: https://doi.org/10.1038/leu.2008.291.

Kyle, R. A. *et al.* (2007) 'Clinical course and prognosis of smoldering (asymptomatic) multiple myeloma', *N Engl J Med,* 356(25), pp. 2582-90. Available at: https://doi.org/10.1056/NEJMoa070389.

Kyrtsonis, M. C. *et al.* (2007) 'Prognostic value of serum free light chain ratio at diagnosis in multiple myeloma', *Br J Haematol,* 137(3), pp. 240-3. Available at: https://doi.org/10.1111/j.1365-2141.2007.06561.x.

Ladetto, M. *et al.* (2014) 'Next-generation sequencing and real-time quantitative PCR for minimal residual disease detection in B-cell disorders', *Leukemia,* 28(6), pp. 1299-307. Available at: https://doi.org/10.1038/leu.2013.375.

Lahuerta, J.-J. *et al.* (2017) 'Depth of Response in Multiple Myeloma: A Pooled Analysis of Three PETHEMA/GEM Clinical Trials', *Journal of Clinical Oncology,* 35(25), pp. 2900-2910. Available at: https://doi.org/10.1200/jco.2016.69.2517.

Lahuerta, J. J. *et al.* (2008) 'Influence of pre- and post-transplantation responses on outcome of patients with multiple myeloma: sequential improvement of response and achievement of complete response are associated with longer survival', *J Clin Oncol,* 26(35), pp. 5775-82. Available at: https://doi.org/10.1200/JCO.2008.17.9721.

Lam, W. Y. and Bhattacharya, D. (2018) 'Metabolic Links between Plasma Cell Survival, Secretion, and Stress', *Trends Immunol,* 39(1), pp. 19-27. Available at:

https://doi.org/10.1016/j.it.2017.08.007.

Landgren, O. *et al.* (2009) 'Monoclonal gammopathy of undetermined significance (MGUS) consistently precedes multiple myeloma: a prospective study', *Blood,* 113(22), pp. 5412-7. Available at: https://doi.org/10.1182/blood-2008-12-194241.

Landgren, O. and Owen, R. G. (2016) 'Better therapy requires better response evaluation: Paving the way for minimal residual disease testing for every myeloma patient', *Cytometry B Clin Cytom,* 90(1), pp. 14-20. Available at: https://doi.org/10.1002/cyto.b.21273.

Lee, J. W. *et al.* (2006) 'Fit-for-purpose method development and validation for successful biomarker measurement', *Pharm Res,* 23(2), pp. 312-28. Available at: https://doi.org/10.1007/s11095-005-9045-3.

Lidén, Y. *et al.* (2009) 'Procedure-related pain among adult patients with hematologic malignancies', *Acta Anaesthesiol Scand,* 53(3), pp. 354-63. Available at: https://doi.org/10.1111/j.1399-6576.2008.01874.x.

Lin, A. W. *et al.* (2004) 'CD56(+dim) and CD56(+bright) cell activation and apoptosis in hepatitis C virus infection', *Clin Exp Immunol,* 137(2), pp. 408-16. Available at: https://doi.org/10.1111/j.1365-2249.2004.02523.x.

Lopez-Anglada, L. *et al.* (2018) 'Prognostic utility of serum free light chain ratios and heavy-light chain ratios in multiple myeloma in three PETHEMA/GEM phase III clinical trials', *PLoS One,* 13(9), pp. e0203392. Available at: https://doi.org/10.1371/journal.pone.0203392.

Ludwig, H. *et al.* (2016) 'Suppression of the noninvolved pair of the myeloma isotype correlates with poor survival in newly diagnosed and relapsed/refractory patients with myeloma', *Am J Hematol,* 91(3), pp. 295-301. Available at: https://doi.org/10.1002/ajh.24268.

Ludwig, H. *et al.* (2013) 'Immunoglobulin heavy/light chain ratios improve paraprotein detection and monitoring, identify residual disease and correlate with survival in multiple myeloma patients', *Leukemia,* 27(1), pp. 213-9. Available at: https://doi.org/10.1038/leu.2012.197.

Ludwig, H. *et al.* (2014) 'European perspective on multiple myeloma treatment strategies in 2014', *Oncologist,* 19(8), pp. 829-44. Available at: https://doi.org/10.1634/theoncologist.2014-0042.

Mai, E. K. *et al.* (2018) 'A systematic classification of death causes in multiple myeloma', *Blood Cancer J,* 8(3), pp. 30. Available at: https://doi.org/10.1038/s41408-018-0068-5.

Mailankody, S. *et al.* (2015) 'Minimal residual disease in multiple myeloma: bringing the bench to the bedside', *Nat Rev Clin Oncol,* 12(5), pp. 286-95. Available at: https://doi.org/10.1038/nrclinonc.2014.239.

Majithia, N. *et al.* (2016) 'Early relapse following initial therapy for multiple myeloma predicts poor outcomes in the era of novel agents', *Leukemia,* 30(11), pp. 2208-2213. Available at: https://doi.org/10.1038/leu.2016.147.

Martin, T. and Huff, C. A. (2019) 'Multiple Myeloma: Current Advances and Future Directions',

Clin Lymphoma Myeloma Leuk, 19(5), pp. 255-263. Available at: https://doi.org/10.1016/j.clml.2019.03.025.

Martinez-Lopez, J. *et al.* (2015) 'Critical analysis of the stringent complete response in multiple myeloma: contribution of sFLC and bone marrow clonality', *Blood,* 126(7), pp. 858-62. Available at: https://doi.org/10.1182/blood-2015-04-638742.

Martínez-Sánchez, P. *et al.* (2008) 'Evaluation of minimal residual disease in multiple myeloma patients by fluorescent-polymerase chain reaction: the prognostic impact of achieving molecular response', *Br J Haematol,* 142(5), pp. 766-74. Available at: https://doi.org/10.1111/j.1365-2141.2008.07263.x.

Mateos, M. V., Nooka, A. K. and Larson, S. M. (2022) 'Moving Toward a Cure for Myeloma', *Am Soc Clin Oncol Educ Book,* 42, pp. 1-12. Available at: https://doi.org/10.1200/EDBK_349603.

Mateos, M. V. *et al.* (2010) 'Bortezomib, melphalan, and prednisone versus bortezomib, thalidomide, and prednisone as induction therapy followed by maintenance treatment with bortezomib and thalidomide versus bortezomib and prednisone in elderly patients with untreated multiple myeloma: a randomised trial', *Lancet Oncol,* 11(10), pp. 934-41. Available at: https://doi.org/10.1016/S1470-2045(10)70187-X.

Mateos, M. V. *et al.* (2014) 'GEM2005 trial update comparing VMP/VTP as induction in elderly multiple myeloma patients: do we still need alkylators?', *Blood,* 124(12), pp. 1887-93. Available at: https://doi.org/10.1182/blood-2014-05-573733.

McCarthy, P. L. *et al.* (2017) 'Lenalidomide Maintenance After Autologous Stem-Cell Transplantation in Newly Diagnosed Multiple Myeloma: A Meta-Analysis', *Journal of clinical oncology : official journal of the American Society of Clinical Oncology,* 35(29), pp. 3279-3289. Available at: https://doi.org/10.1200/JCO.2017.72.6679.

McCudden, C. *et al.* (2016) 'Monitoring multiple myeloma patients treated with daratumumab: teasing out monoclonal antibody interference', *Clin Chem Lab Med,* 54(6), pp. 1095-104. Available at: https://doi.org/10.1515/cclm-2015-1031.

Meads, M. B. *et al.* (2016) 'Targeting PYK2 mediates microenvironment-specific cell death in multiple myeloma', *Oncogene,* 35(21), pp. 2723-2734. Available at: https://doi.org/10.1038/onc.2015.334.

Medina, A. *et al.* (2020) 'Comparison of next-generation sequencing (NGS) and next-generation flow (NGF) for minimal residual disease (MRD) assessment in multiple myeloma', *Blood Cancer J,* 10(10), pp. 108. Available at: https://doi.org/10.1038/s41408-020-00377-0.

Melchor, L. *et al.* (2014) 'Single-cell genetic analysis reveals the composition of initiating clones and phylogenetic patterns of branching and parallel evolution in myeloma', *Leukemia,* 28(8), pp. 1705-15. Available at: https://doi.org/10.1038/leu.2014.13.

Michallet, M. *et al.* (2018) 'Heavy+light chain monitoring correlates with clinical outcome in multiple myeloma patients', *Leukemia,* 32(2), pp. 376-382. Available at:

https://doi.org/10.1038/leu.2017.209.

Migkou, M. *et al.* (2020) 'Clinical characteristics and outcomes of oligosecretory and non-secretory multiple myeloma', *Annals of Hematology,* 99(6), pp. 1251-1255. Available at: https://doi.org/10.1007/s00277-020-03984-w.

Milani, P. *et al.* (2017) 'The utility of MASS-FIX to detect and monitor monoclonal proteins in the clinic', *Am J Hematol,* 92(8), pp. 772-779. Available at: https://doi.org/10.1002/ajh.24772.

Mills, J. R. *et al.* (2017) 'High sensitivity blood-based M-protein detection in sCR patients with multiple myeloma', *Blood Cancer J,* 7(8), pp. e590. Available at: https://doi.org/10.1038/bcj.2017.75.

Mills, J. R. *et al.* (2016) 'Comprehensive Assessment of M-Proteins Using Nanobody Enrichment Coupled to MALDI-TOF Mass Spectrometry', *Clin Chem,* 62(10), pp. 1334-44. Available at: https://doi.org/10.1373/clinchem.2015.253740.

Mishima, Y. *et al.* (2017) 'The Mutational Landscape of Circulating Tumor Cells in Multiple Myeloma', *Cell Rep,* 19(1), pp. 218-224. Available at: https://doi.org/10.1016/j.celrep.2017.03.025.

Miyazaki, K. and Suzuki, K. (2018) 'Abnormal Heavy/Light Chain Ratio and Matched Pair Suppression Increase Residual Disease Detection Sensitivity in Patients With Multiple Myeloma With Deep Responses', *Clin Lymphoma Myeloma Leuk,* 18(4), pp. 293-296. Available at: https://doi.org/10.1016/j.clml.2018.01.008.

Mohan, M. *et al.* (2021) 'Persistent bone marrow minimal residual disease as a "high-risk" disease feature in multiple myeloma', *Am J Hematol,* 96(9), pp. E341-E344. Available at: https://doi.org/10.1002/ajh.26255.

Moore, L. M., Cho, S. and Thoren, K. L. (2019) 'MALDI-TOF mass spectrometry distinguishes daratumumab from M-proteins', *Clin Chim Acta,* 492, pp. 91-94. Available at: https://doi.org/10.1016/j.cca.2019.02.017.

Moreau, P. *et al.* (2017) 'Prospective Evaluation of Magnetic Resonance Imaging and [', *J Clin Oncol,* 35(25), pp. 2911-2918. Available at: https://doi.org/10.1200/JCO.2017.72.2975.

Moreau, P. *et al.* (2019) 'Bortezomib, thalidomide, and dexamethasone with or without daratumumab before and after autologous stem-cell transplantation for newly diagnosed multiple myeloma (CASSIOPEIA): a randomised, open-label, phase 3 study', *The Lancet,* 394(10192), pp. 29-38. Available at: https://doi.org/10.1016/S0140-6736(19)31240-1.

Moreaux, J. *et al.* (2006) 'CD200 is a new prognostic factor in multiple myeloma', *Blood,* 108(13), pp. 4194-7. Available at: https://doi.org/10.1182/blood-2006-06-029355.

Morgan, G. J. *et al.* (2010) 'First-line treatment with zoledronic acid as compared with clodronic acid in multiple myeloma (MRC Myeloma IX): a randomised controlled trial', *Lancet,* 376(9757), pp. 1989-99. Available at: https://doi.org/10.1016/S0140-6736(10)62051-X.

Muccio, V. E. *et al.* (2016) 'Multiple myeloma: New surface antigens for the characterization of plasma cells in the era of novel agents', *Cytometry B Clin Cytom,* 90(1), pp. 81-90. Available at: https://doi.org/10.1002/cyto.b.21279.

Munshi, N. C. *et al.* (2020) 'A large meta-analysis establishes the role of MRD negativity in long-term survival outcomes in patients with multiple myeloma', *Blood advances,* 4(23), pp. 5988-5999. Available at: https://doi.org/10.1182/bloodadvances.2020002827.

Munshi, N. C. *et al.* (2017) 'Association of Minimal Residual Disease With Superior Survival Outcomes in Patients With Multiple Myeloma: A Meta-analysis', *JAMA Oncol,* 3(1), pp. 28-35. Available at: https://doi.org/10.1001/jamaoncol.2016.3160.

Murray, D. *et al.* (2019) 'Detection and prevalence of monoclonal gammopathy of undetermined significance: a study utilizing mass spectrometry-based monoclonal immunoglobulin rapid accurate mass measurement', *Blood Cancer J,* 9(12), pp. 102. Available at: https://doi.org/10.1038/s41408-019-0263-z.

Murray, D. L. *et al.* (2021) 'Mass spectrometry for the evaluation of monoclonal proteins in multiple myeloma and related disorders: an International Myeloma Working Group Mass Spectrometry Committee Report', *Blood Cancer J,* 11(2), pp. 24. Available at: https://doi.org/10.1038/s41408-021-00408-4.

NICE (2016) 'Myeloma: diagnosis and management', *NICE Guideline,* NG35.

Oliva, S. *et al.* (2016) 'Minimal residual disease after transplantation or lenalidomide-based consolidation in myeloma patients: a prospective analysis', *Oncotarget,* 8(4).

ONS (2019) *Office for National Statistics Cancer survival in England.* Available at: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datasets/cancersurvivalratescancersurvivalinenglandadultsdiagnosed (Accessed: September 2022).

Paiva, B. *et al.* (2016) 'Minimal residual disease monitoring and immune profiling in multiple myeloma in elderly patients', *Blood,* 127(25), pp. 3165-74. Available at: https://doi.org/10.1182/blood-2016-03-705319.

Paiva, B., Merino, J. and San Miguel, J. F. (2016) 'Utility of flow cytometry studies in the management of patients with multiple myeloma', *Curr Opin Oncol,* 28(6), pp. 511-517. Available at: https://doi.org/10.1097/CCO.0000000000000331.

Paiva, B. *et al.* (2013) 'Detailed characterization of multiple myeloma circulating tumor cells shows unique phenotypic, cytogenetic, functional, and circadian distribution profile', *Blood,* 122(22), pp. 3591-8. Available at: https://doi.org/10.1182/blood-2013-06-510453.

Paiva, B. *et al.* (2008a) 'Multiparameter flow cytometric remission is the most relevant prognostic factor for multiple myeloma patients who undergo autologous stem cell transplantation', *Blood,* 112(10), pp. 4017-23. Available at: https://doi.org/10.1182/blood-2008-05-159624.

Paiva, B. *et al.* (2008b) 'Multiparameter flow cytometric remission is the most relevant prognostic factor for multiple myeloma patients who undergo autologous stem cell transplantation', *Blood,* 112(10), pp. 4017-23. Available at: https://doi.org/10.1182/blood-2008-05-159624.

Palumbo, A. and Anderson, K. (2011) 'Multiple myeloma', *N Engl J Med,* 364(11), pp. 1046-60. Available at: https://doi.org/10.1056/NEJMra1011442.

Palumbo, A. *et al.* (2014) 'Second primary malignancies with lenalidomide therapy for newly diagnosed myeloma: a meta-analysis of individual patient data', *Lancet Oncol,* 15(3), pp. 333-42. Available at: https://doi.org/10.1016/s1470-2045(13)70609-0.

Paolini, L. *et al.* (2015) 'Comparison of Hevylite™ IgA and IgG assay with conventional techniques for the diagnosis and follow-up of plasma cell dyscrasia', *Ann Clin Biochem,* 52(Pt 3), pp. 337-45. Available at: https://doi.org/10.1177/0004563214564225.

Peceliunas, V. *et al.* (2012) 'Circulating plasma cells predict the outcome of relapsed or refractory multiple myeloma', *Leuk Lymphoma,* 53(4), pp. 641-7. Available at: https://doi.org/10.3109/10428194.2011.627481.

Pellat-Deceunynck, C. and Bataille, R. (2004) 'Normal and malignant human plasma cells: proliferation, differentiation, and expansions in relation to CD45 expression', *Blood Cells Mol Dis,* 32(2), pp. 293-301. Available at: https://doi.org/10.1016/j.bcmd.2003.12.001.

Perrot, A. *et al.* (2018) 'Minimal residual disease negativity using deep sequencing is a major prognostic factor in multiple myeloma', *Blood,* 132(23), pp. 2456-2464. Available at: https://doi.org/10.1182/blood-2018-06-858613.

Puig, N. *et al.* (2019a) 'Qip-Mass Spectrometry in High Risk Smoldering Multiple Myeloma Patients Included in the GEM-CESAR Trial: Comparison with Conventional and Minimal Residual Disease IMWG Response Assessment', *Blood,* 134(Supplement_1), pp. 581-581. Available at: https://doi.org/10.1182/blood-2019-127717.

Puig, N. *et al.* (2019b) 'Qip-Mass Spectrometry in High Risk Smoldering Multiple Myeloma Patients Included in the GEM-CESAR Trial: Comparison with Conventional and Minimal Residual Disease IMWG Response Assessment', *Blood,* 134(Supplement_1), pp. 581-581. Available at: https://doi.org/10.1182/blood-2019-127717.

Puig, N. *et al.* (2014) 'Critical evaluation of ASO RQ-PCR for minimal residual disease evaluation in multiple myeloma. A comparative analysis with flow cytometry', *Leukemia,* 28(2), pp. 391-7. Available at: https://doi.org/10.1038/leu.2013.217.

Pulte, D., Gondos, A. and Brenner, H. (2011) 'Improvement in Survival of Older Adults with Multiple Myeloma: Results of an Updated Period Analysis of SEER Data', *The Oncologist,* 16(11), pp. 1600-1603. Available at: https://doi.org/10.1634/theoncologist.2011-0229.

Rajkumar, S. V. (2020) 'Multiple myeloma: 2020 update on diagnosis, risk-stratification and management', *American Journal of Hematology,* 95(5), pp. 548-567. Available at: https://doi.org/https://doi.org/10.1002/ajh.25791.

Rajkumar, S. V. *et al.* (2014) 'International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma', *Lancet Oncol,* 15(12), pp. e538-48. Available at: https://doi.org/10.1016/s1470-2045(14)70442-5.

Rajkumar, S. V. *et al.* (2011) 'Consensus recommendations for the uniform reporting of clinical trials: report of the International Myeloma Workshop Consensus Panel 1', *Blood,* 117(18), pp. 4691-5. Available at: https://doi.org/10.1182/blood-2010-10-299487.

Rajkumar, S. V. *et al.* (2005) 'Serum free light chain ratio is an independent risk factor for progression in monoclonal gammopathy of undetermined significance', *Blood,* 106(3), pp. 812-7. Available at: https://doi.org/10.1182/blood-2005-03-1038.

Rasche, L. *et al.* (2018) 'The presence of large focal lesions is a strong independent prognostic factor in multiple myeloma', *Blood,* 132(1), pp. 59-66. Available at: https://doi.org/10.1182/blood-2018-04-842880.

Rawstron, A. C. *et al.* (2013) 'Minimal residual disease assessed by multiparameter flow cytometry in multiple myeloma: impact on outcome in the Medical Research Council Myeloma IX Study', *J Clin Oncol,* 31(20), pp. 2540-7. Available at: https://doi.org/10.1200/jco.2012.46.2119.

Rawstron, A. C. *et al.* (2002) 'Flow cytometric disease monitoring in multiple myeloma: the relationship between normal and neoplastic plasma cells predicts outcome after transplantation', *Blood,* 100(9), pp. 3095-100. Available at: https://doi.org/10.1182/blood-2001-12-0297.

Rawstron, A. C. *et al.* (2016) 'Measuring disease levels in myeloma using flow cytometry in combination with other laboratory techniques: Lessons from the past 20 years at the Leeds Haematological Malignancy Diagnostic Service', *Cytometry B Clin Cytom,* 90(1), pp. 54-60. Available at: https://doi.org/10.1002/cyto.b.21271.

Rawstron, A. C. *et al.* (2015) 'Minimal residual disease in myeloma by flow cytometry: independent prediction of survival benefit per log reduction', *Blood,* 125(12), pp. 1932-5. Available at: https://doi.org/10.1182/blood-2014-07-590166.

Rawstron, A. C. *et al.* (2008) 'Report of the European Myeloma Network on multiparametric flow cytometry in multiple myeloma and related disorders', *Haematologica,* 93(3), pp. 431-8. Available at: https://doi.org/10.3324/haematol.11080.

Riebl, V. *et al.* (2021) 'Ten Color Multiparameter Flow Cytometry in Bone Marrow and Apheresis Products for Assessment and Outcome Prediction in Multiple Myeloma Patients', *Front Oncol,* 11, pp. 708231. Available at: https://doi.org/10.3389/fonc.2021.708231.

Robson, E. *et al.* (2009) 'Utility of Serum Free Light Chain Analysis When Screening for Lymphoproliferative Disorders: The Experience at a District General Hospital in the United Kingdom', *Labmedicine,* 40, pp. 325-329. Available at: https://doi.org/10.1309/M6YUPSL3EIR7KE.

Rollig, C., Knop, S. and Bornhauser, M. (2015) 'Multiple myeloma', *Lancet,* 385(9983), pp. 2197-

208. Available at: https://doi.org/10.1016/s0140-6736(14)60493-1.

Roman, E. *et al.* (2022) 'Cohort Profile Update: The Haematological Malignancy Research Network (HMRN) UK population-based cohorts', *International Journal of Epidemiology,* 51(3), pp. e87-e94. Available at: https://doi.org/10.1093/ije/dyab275.

Roshal, M. *et al.* (2017) 'MRD detection in multiple myeloma: comparison between MSKCC 10-color single-tube and EuroFlow 8-color 2-tube methods', *Blood advances,* 1(12), pp. 728-732. Available at: https://doi.org/10.1182/bloodadvances.2016003715.

Rosko, A. *et al.* (2017) 'Myeloma in Elderly Patients: When Less Is More and More Is More', *Am Soc Clin Oncol Educ Book,* 37, pp. 575-585. Available at: https://doi.org/10.1200/edbk_175171.

Royston, D. J. *et al.* (2016) 'Single-Tube 10-Fluorochrome Analysis for Efficient Flow Cytometric Evaluation of Minimal Residual Disease in Plasma Cell Myeloma', *Am J Clin Pathol,* 146(1), pp. 41-9. Available at: https://doi.org/10.1093/ajcp/aqw052.

Rustad, E. H. *et al.* (2017) 'Monitoring multiple myeloma by quantification of recurrent mutations in serum', *Haematologica,* 102(7), pp. 1266-1272. Available at: https://doi.org/10.3324/haematol.2016.160564.

Ríos-Tamayo, R. *et al.* (2018) 'Epidemiology of Multiple Myeloma', in Khalid Ahmed, A.-A. (ed.) *Update on Multiple Myeloma*. Rijeka: IntechOpen, pp. Ch. 2.

Ríos-Tamayo, R. *et al.* (2021) 'The Current Role of the Heavy/Light Chain Assay in the Diagnosis, Prognosis and Monitoring of Multiple Myeloma: An Evidence-Based Approach', *Diagnostics (Basel),* 11(11). Available at: https://doi.org/10.3390/diagnostics11112020.

Sahara, N. *et al.* (2002) 'Clinicopathological and prognostic characteristics of CD56-negative multiple myeloma', *Br J Haematol,* 117(4), pp. 882-5. Available at: https://doi.org/10.1046/j.1365-2141.2002.03513.x.

Salem, D. *et al.* (2016) 'Myeloma minimal residual disease testing in the United States: Evidence of improved standardization', *Am J Hematol,* 91(12), pp. E502-E503. Available at: https://doi.org/10.1002/ajh.24540.

Sánchez-Ramón, S., Dhalla, F. and Chapel, H. (2016) 'Challenges in the Role of Gammaglobulin Replacement Therapy and Vaccination Strategies for Hematological Malignancy', *Front Immunol,* 22(7): pp. 317. Available at: https://doi: 10.3389/fimmu.2016.00317.

San Miguel, J. F. *et al.* (2002) 'Immunophenotypic evaluation of the plasma cell compartment in multiple myeloma: a tool for comparing the efficacy of different treatment strategies and predicting outcome', *Blood,* 99(5), pp. 1853-6. Available at: https://doi.org/10.1182/blood.v99.5.1853.

Sanoja-Flores, L. *et al.* (2019) 'Blood monitoring of circulating tumor plasma cells by next generation flow in multiple myeloma after therapy', *Blood,* 134(24), pp. 2218-2222. Available at: https://doi.org/10.1182/blood.2019002610.

Schinke, C. *et al.* (2017) 'The prognostic value of the depth of response in multiple myeloma

depends on the time of assessment, risk status and molecular subtype', *Haematologica,* 102(8), pp. e313-e316. Available at: https://doi.org/10.3324/haematol.2017.165217.

Shain, K. H. and Dalton, W. S. (2009) 'Environmental-mediated drug resistance: a target for multiple myeloma therapy', *Expert Rev Hematol,* 2(6), pp. 649-62. Available at: https://doi.org/10.1586/ehm.09.55.

Silva, L. S. V. D. *et al.* (2017) 'For survival, the emergence of oligoclonal bands after multiple myeloma treatment is less important than achieving complete remission', *Rev Bras Hematol Hemoter,* 39(4), pp. 331-336. Available at: https://doi.org/10.1016/j.bjhh.2017.05.010.

Soh, K. T., Tario, J. D., Jr. and Wallace, P. K. (2017) 'Diagnosis of Plasma Cell Dyscrasias and Monitoring of Minimal Residual Disease by Multiparametric Flow Cytometry', *Clin Lab Med,* 37(4), pp. 821-853. Available at: https://doi.org/10.1016/j.cll.2017.08.001.

Soh, K. T. *et al.* (2020) 'Methodological considerations for the high sensitivity detection of multiple myeloma measurable residual disease', *Cytometry Part B: Clinical Cytometry,* 98(2), pp. 161-173. Available at: https://doi.org/10.1002/cyto.b.21862.

Sonneveld, P. (2017) 'Should minimal residual disease negativity', *Blood Adv,* 1(8), pp. 522-525. Available at: https://doi.org/10.1182/bloodadvances.2017000109.

Sonneveld, P. *et al.* (2016) 'Treatment of multiple myeloma with high-risk cytogenetics: a consensus of the International Myeloma Working Group', *Blood,* 127(24), pp. 2955-2962. Available at: https://doi.org/10.1182/blood-2016-01-631200.

Sonneveld, P. and Broijl, A. (2016) 'Treatment of relapsed and refractory multiple myeloma', *Haematologica,* 101(4), pp. 396-406. Available at: https://doi.org/10.3324/haematol.2015.129189.

Stetler-Stevenson, M. *et al.* (2016) 'Consensus guidelines for myeloma minimal residual disease sample staining and data acquisition', *Cytometry B Clin Cytom,* 90(1), pp. 26-30. Available at: https://doi.org/10.1002/cyto.b.21249.

Suehara, Y. *et al.* (2017) 'Abnormal heavy/light chain ratio after treatment is associated with shorter survival in patients with IgA myeloma', *Cancer Sci,* 108(2), pp. 187-192. Available at: https://doi.org/10.1111/cas.13125.

Takamatsu, H. *et al.* (2017) 'Prognostic value of sequencing-based minimal residual disease detection in patients with multiple myeloma who underwent autologous stem-cell transplantation', *Ann Oncol,* 28(10), pp. 2503-2510. Available at: https://doi.org/10.1093/annonc/mdx340.

Teitelbaum, A. *et al.* (2013) 'Health care costs and resource utilization, including patient burden, associated with novel-agent-based treatment versus other therapies for multiple myeloma: findings using real-world claims data', *Oncologist,* 18(1), pp. 37-45. Available at: https://doi.org/10.1634/theoncologist.2012-0113.

Terpos, E. *et al.* (2019) 'Next Generation Flow Cytometry Provides a Standardized, Highly

Sensitive and Informative Method for the Analysis of Circulating Plasma Cells in Newly Diagnosed Multiple Myeloma: A Single Center Study in 182 Patients', *Blood,* 134(Supplement_1), pp. 4338-4338. Available at: https://doi.org/10.1182/blood-2019-127989.

Thompson, P. A. *et al.* (2018) 'Serial minimal residual disease (MRD) monitoring during first-line FCR treatment for CLL may direct individualized therapeutic strategies', *Leukemia,* 32(11), pp. 2388-2398. Available at: https://doi.org/10.1038/s41375-018-0132-y.

Thoren, K. L. (2018) 'Mass spectrometry methods for detecting monoclonal immunoglobulins in multiple myeloma minimal residual disease', *Semin Hematol,* 55(1), pp. 41-43. Available at: https://doi.org/10.1053/j.seminhematol.2018.02.008.

Ting, H. Y. *et al.* (2019) 'Heavy/light chain assay in the monitoring of multiple myeloma', *Pathology,* 51(5), pp. 507-511. Available at: https://doi.org/10.1016/j.pathol.2019.04.002.

Tomlinson, R. (2018) 'Multiple myeloma: Updated approach to management in 2018', *Aust J Gen Pract,* 47(8), pp. 526-529.

Usmani, S. Z. *et al.* (2018) 'Clinical predictors of long-term survival in newly diagnosed transplant eligible multiple myeloma - an IMWG Research Project', *Blood Cancer J,* 8(12), pp. 123. Available at: https://doi.org/10.1038/s41408-018-0155-7.

van de Velde, H. J. K. *et al.* (2007) 'Complete response correlates with long-term survival and progression-free survival in high-dose therapy in multiple myeloma', *Haematologica,* 92(10), pp. 1399-1406. Available at: https://doi.org/10.3324/haematol.11534.

van der Velden, V. H. J. *et al.* (2019) 'Optimization and testing of dried antibody tube: The EuroFlow LST and PIDOT tubes as examples', *J Immunol Methods,* 475, pp. 112287. Available at: https://doi.org/10.1016/j.jim.2017.03.011.

van Dongen, J. J. M. *et al.* (2012) 'EuroFlow antibody panels for standardized n-dimensional flow cytometric immunophenotyping of normal, reactive and malignant leukocytes', *Leukemia,* 26(9), pp. 1908-1975. Available at: https://doi.org/10.1038/leu.2012.120.

van Rhee, F. *et al.* (2007) 'High serum-free light chain levels and their rapid reduction in response to therapy define an aggressive multiple myeloma subtype with poor prognosis', *Blood,* 110(3), pp. 827-32. Available at: https://doi.org/10.1182/blood-2007-01-067728.

Vij, R. *et al.* (2015) 'Serum free light chain reduction correlates with response and progression-free survival following carfilzomib therapy in relapsed/refractory multiple myeloma', *Leuk Lymphoma,* 56(10), pp. 2959-61. Available at: https://doi.org/10.3109/10428194.2015.1020801.

Wadhera, R. K. and Rajkumar, S. V. (2010) 'Prevalence of monoclonal gammopathy of undetermined significance: a systematic review', *Mayo Clin Proc,* 85(10), pp. 933-42. Available at: https://doi.org/10.4065/mcp.2010.0337.

Walker, B. A. *et al.* (2010) 'A compendium of myeloma-associated chromosomal copy number

abnormalities and their prognostic value', *Blood,* 116(15), pp. e56-e65. Available at: https://doi.org/10.1182/blood-2010-04-279596.

Walker, B. A. *et al.* (2014) 'Intraclonal heterogeneity is a critical early event in the development of myeloma and precedes the development of clinical symptoms', *Leukemia,* 28(2), pp. 384-390. Available at: https://doi.org/10.1038/leu.2013.199.

Weiss, B.M. *et al.* (2009) 'A monoclonal gammopathy precedes multiple myeloma in most patients', *Blood,* 113(22), pp. 5418-22. Available at: https:// doi:10.1182/blood-2008-12-195008

Willrich, M. A. and Katzmann, J. A. (2016) 'Laboratory testing requirements for diagnosis and follow-up of multiple myeloma and related plasma cell dyscrasias', *Clin Chem Lab Med,* 54(6), pp. 907-19. Available at: https://doi.org/10.1515/cclm-2015-0580.

Willrich, M. A. V., Murray, D. L. and Kyle, R. A. (2018) 'Laboratory testing for monoclonal gammopathies: Focus on monoclonal gammopathy of undetermined significance and smoldering multiple myeloma', *Clinical Biochemistry,* 51, pp. 38-47. Available at: https://doi.org/https://doi.org/10.1016/j.clinbiochem.2017.05.001.

Wood, B. *et al.* (2013) 'Validation of cell-based fluorescence assays: practice guidelines from the ICSH and ICCS - part V - assay performance criteria', *Cytometry B Clin Cytom,* 84(5), pp. 315-23. Available at: https://doi.org/10.1002/cyto.b.21108.

Wood, P. B., McElroy, Y. G. and Stone, M. J. (2010) 'Comparison of serum immunofixation electrophoresis and free light chain assays in the detection of monoclonal gammopathies', *Clin Lymphoma Myeloma Leuk,* 10(4), pp. 278-80. Available at: https://doi.org/10.3816/CLML.2010.n.057.

Wu, V. *et al.* (2018) 'Risk stratification of smoldering multiple myeloma: predictive value of free light chains and group-based trajectory modeling', *Blood Adv,* 2(12), pp. 1470-1479. Available at: https://doi.org/10.1182/bloodadvances.2018016998.

Zamagni, E. *et al.* (2020) 'Role of Imaging in the Evaluation of Minimal Residual Disease in Multiple Myeloma Patients', *Journal of clinical medicine,* 9(11), pp. 3519. Available at: https://doi.org/10.3390/jcm9113519.

Zingone, A. and Kuehl, W.M. (2011) 'Pathogenesis of monoclonal gammopathy of undetermined significance and progression to multiple myeloma', *Semin. Hematol,* 48(1), pp. 4-12. Available at: https:// doi: 10.1053/j.seminhematol.2010.11.003.

Zhu, S. *et al.* (2021) 'Serum Protein Electrophoresis and Immunofixation Electrophoresis Detection in Multiple Myeloma', *J Coll Physicians Surg Pak,* 30(7), pp. 864-867. Available at: https://doi.org/10.29271/jcpsp.2021.07.864.

Appendix 1 – Standard operating procedures

Standard operating procedures (SOPs) for the measurement of immunoglobulins (on c4000 analyser), SPEP (on V8 analyser), IFE (on SAS-3 analyser), FLC and HLC (both on Optilite analyser). Please note these versions are uncontrolled and are only accurate up the time of inclusion in this thesis. These SOPs are the property of the immunology laboratory at Oxford University Hospitals NHS Foundation Trust.

Oxford University Hospitals
NHS Foundation Trust

Pathology and Laboratories Directorate

The Department of Laboratory Immunology

Standard Operating Procedure (SOP)

Title: c4000 analyser

Code:	Immunology 1260 2
Document Status:	Authorised
Authors:	Dawn Simpson
Version:	1.13
Section:	Immunochemistry section
Document Category:	Analytical Process
Authorised By:	Lauren Campbell
Location Of Copy:	unknown
Authorised on:	25-May-2022
Review on or Before:	25-May-2023
Date Time Of Last Update:	26-May-2022 13:33
Date Time Printed:	PREVIEW ONLY - UNCONTROLLED
Number Of Copies:	N/A
Print History:	Copy 1: Imm main lab. Printed: 25-May-2022

This is a controlled document

The master copy of this document is held on iPassport. This document must not be copied or altered without express permission from appropriate management teams. This document will be reviewed on a regular, scheduled basis.

c4000 Analyser

Introduction

The ARCHITECT *c*4000 System is an open, fully-automated, clinical chemistry analyser based on turbidimetric methodology allowing random and continuous access, and priority processing. The processing module, robotic sample handler & System Control Centre comprise the three major parts of the instrument. The c4000 is used for the analysis of serum, urine & CSF for a variety of proteins using Abbott Architect products.

Turbidimetry measures the decrease in light transmission/scatter through a solution; the amount of scatter is inversely proportional to the amount of particles (e.g. proteins) in the solution held within the reaction cuvettes. The light source for the c4000 is a tungsten halogen lamp. Glass cuvettes are housed in a reaction carousel and are used for both sample dilution and assay reactions.

The instrument has a reaction time of up to 10 minutes.
Do not lift front cover when analyser is in Running mode (analysis will stop)
Do not place hands into body of the c4000 when analysis is taking place, damage may occur to the staff and the instrument.

Performance characteristics

Uncertainty of measurement

Analyte	**Techno 1**	**Techno 3**	**Analyte**	**Techno 1**	**Techno 3**
IgG g/l	0.23	0.35	B2MG mg/l	0.13	0.12
IgA g/l	0.04	0.07	RhF IU/ml		1.68
IgM g/l	0.03	0.04	IgGc mg/dl *	0.71	1.51
C3 mg/dl	4.42	8.36	Microalbumin * mg/dl	0.70	2.92
C4 mg/dl	0.92	1.21	Total Protein g/l	2.16	2.68
A1 antitrypsin mg/dl	3.9	4.5	Ceruloplasmin mg/dl	1.52	2.32

*Randox 1 & 2

Trueness (accuracy)

Analyte	MRVIS
IgG	43 to 24
IgA	90 to 79
IgM	29 to 22
C3	112 to 58
C4	55 to 47
Rheumatoid factor	OMIS=0
B2 microglobulin	93 to 117
Ceruloplasmin	82 to 299
A1 antitrypsin	21 to 47
Total Protein	23 to 27
IgGc	52 to 34
MicroAlbumin	42 to 39

For full performance characteristics please see individual verification documents:

Calibration and Traceability
There are 7 different calibrators to perform calibration curves for all analytes. There must be a valid curve for each test for results to be generated. The system software will alert when each Cal Status has expired. Calibrations must be performed on change of lot number & when a curve has expired (approx. 30 days). Re-calibration should also be performed if an IQC drift or bias is noted.
Each curve consists of a blank (water) and a number of standards (the number is analyte specific). Open vial stability in brackets.

- Specific Proteins Multiconstituent Calibrator; consists of human sera containing IgG, IgA, IgM, C3 & C4. It is calibrated against ERM-DA470. (30 days)
- Multiconstituent Calibrator; human-based matrix containing Total Protein (derived from human sourced albumin material). It is calibrated against NIST SRM 927. (7 days)
- Plasmaproteins Calibrator; contains plasmaproteins (ceruloplasmin) in human based serum. It is calibrated against ERM-DA470. (30 days)
- Microalbumin Calibrator; contains human serum albumin. It is calibrated against CRM 470. (6 months)
- Rheumatoid factor Calibrator; contains rheumatoid factor diluted in buffer solution containing 1% w/w bovine serum albumin. It is traceable to WHO International Standard Preparation of Rheumatoid Arthritis Serum, NIBSC 64/2.
- B2 Microglobulin Standard; contains B2 Microglobulin obtained from urine of renal tube dysfunction patients. It is calibrated against the 1st WHO International Standard for Beta-2 Microglobulin.
- Proteins Standard; is of human origin & contains alpha 1 antitrypsin. It is standardized to the International Reference Material for Measurement of 14 Human Serum Proteins (CRM 470).

Sample requirements
Serum, LH plasma, urine & CSF may be tested on the c4000. The analyte stability differs from analyte to analyte (please see relevant section in Performance Characteristics). All samples must be stored at 4oC prior to testing.

Microbially contaminated, heat-treated samples should not be used. Specimens containing visible particulates should be centrifuged before analysis.
Grossly haemolysed or lipaemic specimens should be avoided.

Interferences and cross reactions
EDTA plasma samples are contraindicated for the analysis of IgG, IgA, IgM, C3, C4 and Ceruloplasmin as increased fibrinogen levels may artificially give depressed results. Lithium heparin plasma is acceptable, though serum is preferred.

Method
Check expiry dates of any reagent before use. Use eldest reagents first

<u>Start Up (daily/as needed)</u>

Analyser must be in 'Ready' mode to perform the following;
Turn on pump of ELIX 70 Still.
Perform Maintenance procedures needed (see above).
Check amounts of liquids; system will alert when running low. To replace; Liquids>Update Supplies>update lot numbers (if new) & tick Replaced box. Choose Done to complete.

Check amounts of reagents; remove all empty segments and replace with new ones. All new reagents must be checked for bubbles before use. Visually check lot numbers beforehand. Fill dedicated saline segment with bottled 0.9% Saline. Discard remainder every Monday.
Check if any calibrations are due in Calibration Status window. If needed, check there are enough of each calibrator before requesting. If it is a new lot number, the details must be inputted whilst c4000 is in Ready mode; System > Configuration >
Put c4000 modules into Running Mode to run QC, calibrators and patient samples; Highlight both, select Run.

Calibration (as needed)
Click on Orders button, then Calibration Order.
Input first rack to be used & select analyte to be calibrated. Click on Add Order. Repeat if necessary.
Click on Order Status and select Print. Choose List Report, click Done.
Aliquot appropriate calibrant/s (min. volume given on List Report) into sample cup/s. Place in correct positions on rack/s. Date newly opened bottles.
Select both modules then choose Run. When both are in Running mode place racks onto analyser, calibration will be performed.
When finished check curve; QC-Cal>Calibration Status. Highlight calibration performed and choose Details to view standard curve. Print details (Cal curve Details Report) for each new lot number. Run IQC to check curve, print out results and attach to Cal curve Details Report. Store in appropriate folder.

Internal QC (2x daily/as needed)
Vortex and centrifuge (1 minute) all IQCs before use. Techno 1 aliquots are taken out of the -80 freezer daily. Techno 3 aliquots are taken out of the -80 freezer, dated and kept in the cold room for a week. CSF Randox 1 and 2 aliquots are dated, kept in the cold room and have a stability of 30 days once in use.

Running Samples
Check sample has been spun down (primary tubes).
All caps must be removed from the sample tubes before analysis.
Check there is sufficient specimen (>250µl) and that there are no surface bubbles.
Check the sample is not solid. Place solid samples in 37^{O}C incubator before analysis.

Bar-coded samples are placed directly into the analyser racks with the bar-code facing outwards. Sample details will be entered automatically by the host computer.
Results analysis/Interpretation/Reporting
Analyse results list after each run (see below for printing out files). The c4000 will carry out a number of dilutions to generate an absolute value (dilutions are assay specific).
Print results; Results>Results review. Place results in lab number order, select all then print

Alert/Critical values
All high immunoglobulin levels must be checked with electrophoresis (please see V8 SOP Immunology 1017 & Serum Immunofixation SOP Immunology 1013).

Abnormal/low results that could indicate a significant immunodeficiency should be flagged at bench level and investigated. Those indicating significant immunodeficiency should be highlighted as soon as possible to the person authorising or one of the Consultant Clinical Scientists. If in doubt, speak to a senior member of staff.

Reportable Interval

IgG 1.08-41.4 g/l	IgA 0.05-66.2 g/l	IgM 0.05-31.3* g/l
C3 11- 368 mg/dl	C4 2.9-63 mg/dl	RHF 7-200 IU/ml
A1AT 25-300 mg/dl	CER 2-74 mg/dl	Total protein 8-184 g/l
B2MG 0.1-94 mg/l	IgGc 1-4140 mg/dl	Albc 0.5-200 mg/dl

*further, off-line dilutions can be performed for > results.

Reference ranges

IgG	W1	5.2-18.0	W4	3.9 - 13.0	W12	2.1 - 7.7
	W24	2.4 - 8.8	Y1	3.0 - 10.9	Y2	3.0 - 13.0
	Y6	4.9 - 13.	Adult	6.0 - 16.0 g/l		

IgA	W4	0 - 0.2	W26	0.1 - 0.6	Y1	0.2 - 0.7
	Y2	0.3 - 1.2	Y6	0.4 - 2.0	Adult	0.8 – 3.0 g/l

IgM	W1	0 - 0.2	W4	0.1 - 0.4	W12	0.1 - 0.7
	W26	0.2 - 1.2	Y1	0.4 - 2.0	Adult	0.4 – 2.5 g/l

TPRO 60-80 g/l

Oxford University Hospitals
NHS Foundation Trust

Pathology and Laboratories Directorate

The Department of Laboratory Immunology

Standard Operating Procedure (SOP)

Title: V8 analyser

Code:	Immunology 1017
Document Status:	Authorised
Authors:	Dawn Simpson
Version:	2.7
Section:	Immunochemistry section
Document Category:	Analytical Process
Authorised By:	Lauren Campbell
Location Of Copy:	unknown
Authorised on:	29-Mar-2021
Review on or Before:	29-Mar-2022
Date Time Of Last Update:	03-Aug-2021 16:36
Date Time Printed:	PREVIEW ONLY - UNCONTROLLED
Number Of Copies:	N/A
Print History:	Copy 1: Immunology office. Printed: 25-Nov-2021 Copy 2: Back bench. Printed: 29-Nov-2021

This is a controlled document

The master copy of this document is held on iPassport. This document must not be copied or altered without express permission from appropriate management teams. This document will be reviewed on a regular, scheduled basis.

Introduction

The V8 Serum Protein 6-band Zoom SPE kit is intended for the separation and quantitation of serum proteins by capillary electrophoresis (CE) on the V8 Clinical Capillary Electrophoresis system. The V8 Serum Protein SPE buffer facilitates this separation, and, with the aid of the Platinum software, gives quantitative and qualitative analysis to assist in the clinical diagnosis of many disease states, including multiple myeloma and immunodeficiencies.

The V8 Serum Protein 6-band Zoom kit separates proteins within a buffer filled capillary. The charged molecules separate on a mass:charge ratio, this is achieved using a microbore, fused silica capillary filled with an appropriate electrolyte medium under high voltage. Positively charged ions are drawn through the capillary toward the cathode, the smallest ions eluting first. Electroosmosis of small ions within the buffer electrolyte draws neutral molecules through the capillary and overcomes the electrostatic attraction of negatively charged ions. This electroosmotic flow means negative ions are still drawn through the capillary. The rate of separation is also dependent upon buffer pH, buffer chemistry, strength of the electro-osmotic flow, and the voltage that is applied to the capillaries. This is known as Capillary Zone Electrophoresis (CZE). This process is primarily used to detect the presence or absence of paraproteins.

Performance characteristics

Uncertainty of measurement

Retrospective analysis of kit normal control and in house paraprotein QC over a1 month period gave the following results;

Normal Control

overall mean	11.3
SD	0.25
UOM	0.51

Paraprotein QC

overall mean	6.23
SD	0.4
UOM	0.81

Trueness (accuracy)

Retrospective analysis of samples sent by UKNEQAS as part of the Monoclonal Protein scheme from the previous year showed a constant OMIS of 0 for immunotyping and an MRVIS ranging from 75 to 69 for paraprotein quantification.

Calibration and Traceability

Calibration is performed by an engineer after capillaries have been replaced as part of the annual preventative maintenance. No further calibration is required.

There is no International Standard available for CZE.

Sample requirements

Fresh (<7 days old) serum is needed for electrophoresis. Plasma must not be used. The minimum volume is 50µl (when using sample cups). All samples must be stored at 4^{O}C prior to testing.

Microbially contaminated, heat-treated, or specimens containing visible particulates should not be used. Grossly haemolysed or lipaemic specimens should be avoided.

Samples are disposed after approximately 2 months into large sharps bins.

Interferences and cross reactions
Plasma is contraindicated as it contains a large fibrinogen band between the beta and gamma fractions which may potentially obscure a monoclonal band.
Haemolysis may cause false elevation in the alpha-2 and beta fractions which may also obscure a monoclonal band.

METHOD
Check expiry dates of any buffer or other reagent being placed on the V8 before use. Use eldest reagents first.

Serum electrophoresis

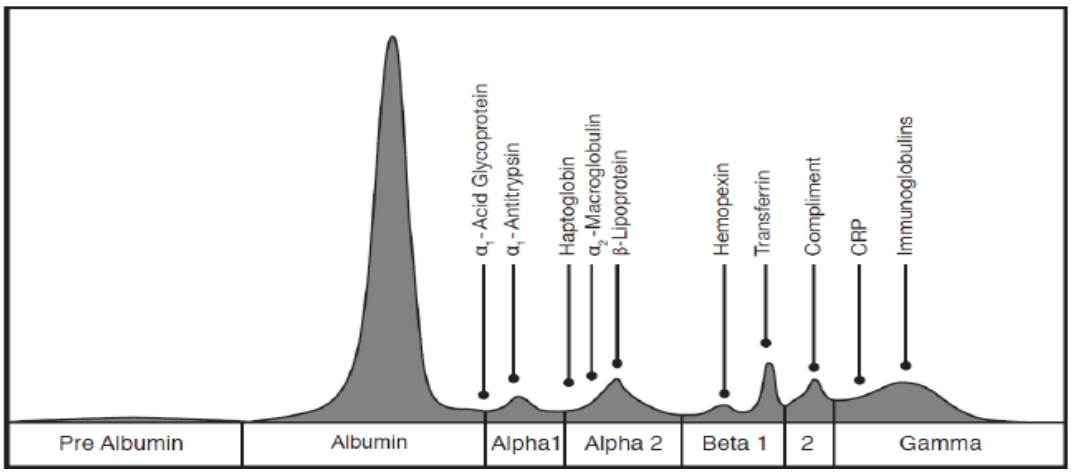
1. Place samples in racks, remembering to remove lids first. Any tiny samples must be placed in a labelled clear sample cup for analysis (use LIMS function, ABCL to print new barcode). Ensure all tubes are pushed to the bottom of the racks. Remember to fill empty positions with skip position barcoded tubes.
2. Open rack cover and place racks on left hand side of sample rack transport area. Close cover. Analysis will start automatically.
3. A square icon  in the Results menu signifies a sample has finished and the trace may be read.

Sample tubes with no barcodes or ones that have been misread will have a blank Tube ID on the worklist. These lab numbers must be entered manually after the V8 has processed the sample(s).

1. Expand patient record using  button. Tap/Click on Tube ID.
2. Enter lab number, click OK

Results analysis/Interpretation

Trace map



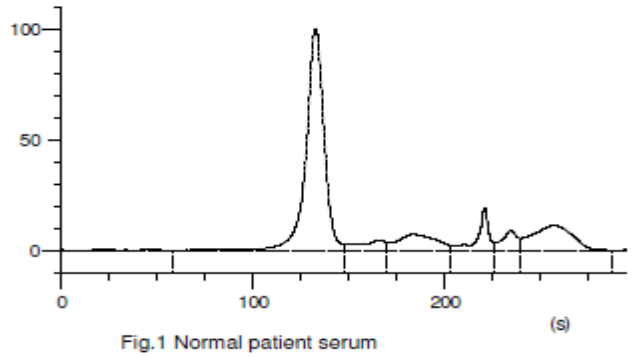


Fig.1 Normal patient serum

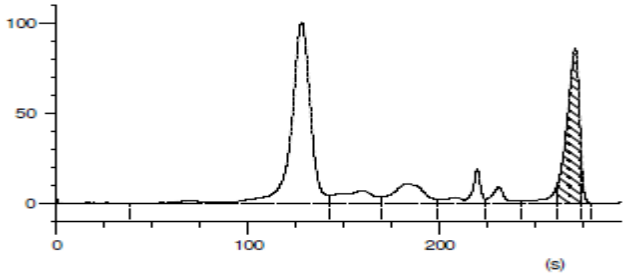


Fig.6 Abnormal patient serum. High concentration monoclonal paraprotein in the gamma region.

Each patient must be analysed across the trace, do not just look at the gamma region. Note down the amount of polyclonality seen in the gamma region on the worklist & in Patient comments box (for use with interface)

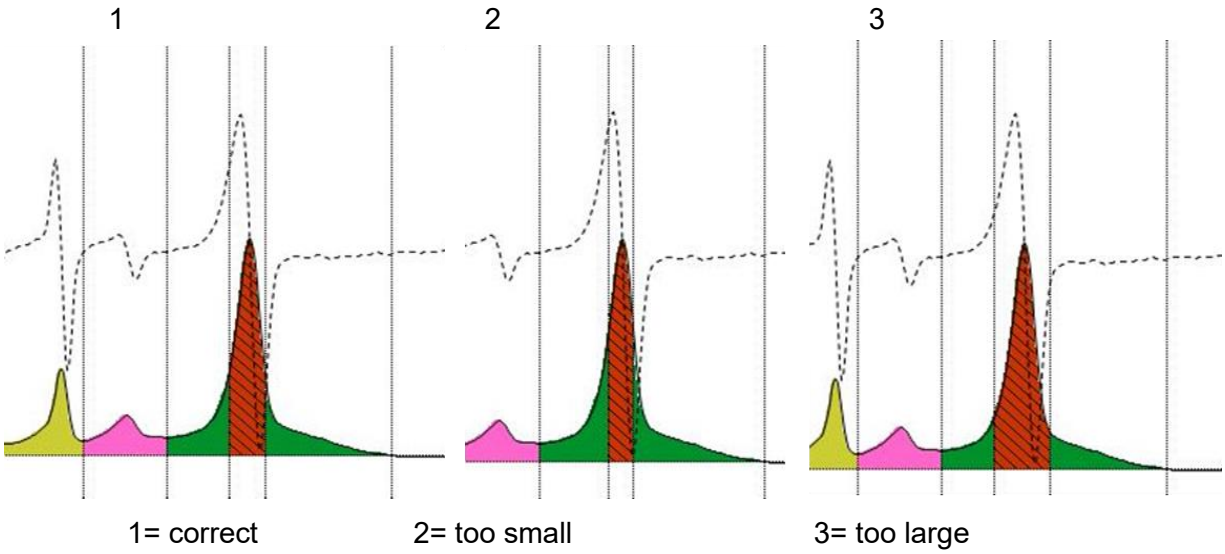
- NORMAL, HYPERGAMMA, HYPOGAMMA, ESA = Normal for age.
- Note down any monoclonal peaks detected and where, on worklist.
- Also note if peaks are missing.

Any monoclonal peaks (paraproteins) must be investigated to determine if they are new or existing. All new paraproteins must be characterised either by Immunodisplacement (see below) or Serum Immunofixation (see separate SOP, Immunology 1013).

Quantification of Paraproteins

All paraproteins in the gamma region are quantified using the same procedure as for the IQC;

1. Add total protein to patient on Results screen if needed; Expand patient record, add value in patient chemistry values.
2. Tap/Click on pencil icon then tap and hold or right click on paraprotein and select 'Add sliced M-spike'.



Alert/Critical values

The following results must be rung to the relevant clinician as soon as possible;

- IgG paraproteins >30 g/l IgA & IgM paraproteins >10 g/l
- Any paraproteins with accompanying immunoparesis
- All monoclonal light chain positive patients

If in doubt, speak to a senior member of staff.

All results rung out must have a hidden message attached to the patient's record on LIMS to signify this has occurred.

Reportable Interval

There is no maximum paraprotein result that can be reported in this assay accurately. The minimum result is 0.2 g/l. However, in reality, tiny paraproteins are reported as 'Too small to quantify'.

Oxford University Hospitals
NHS Foundation Trust

Pathology and Laboratories Directorate

The Department of Laboratory Immunology

Standard Operating Procedure (SOP)

Title: Serum Immunofixation

Code:	Immunology 1013
Document Status:	Authorised
Authors:	Dawn Simpson
Version:	1.8
Section:	Immunochemistry section
Document Category:	Analytical Process
Authorised By:	Rabiah Saleh
Location Of Copy:	unknown
Authorised on:	09-Nov-2021
Review on or Before:	09-Nov-2022
Date Time Of Last Update:	09-Nov-2021 14:31
Date Time Printed:	PREVIEW ONLY - UNCONTROLLED
Number Of Copies:	N/A
Print History:	Copy 1: Imm main lab. Printed: 09-Nov-2021 Copy 2: Immunology office. Printed: 25-Nov-2021 Copy 3: Immunology office. Printed: 26-Nov-2021

This is a controlled document

The master copy of this document is held on iPassport. This document must not be copied or altered without express permission from appropriate management teams. This document will be reviewed on a regular, scheduled basis.

Introduction

The SAS 3 is an analyser provided by Helena Biosciences to perform electrophoresis & immunofixation on agarose gels. The machine includes automated sample application, electrophoresis and automatic reagent application and spreading. It is used for the analysis of serum samples using Helena Biosciences products.

The SAS-3 IFE-9 Kit is intended for the characterisation of serum samples by agarose gel electrophoresis and Immunoprecipitation with anti-sera provided.

The assay is based on the principle of using the differing electrophoretic mobilities of proteins to separate them into discrete bands. A small amount of proteins in solution (e.g., serum) is applied to a thin sheet of agarose gel. When a charge is applied across the gel, the proteins will be attracted toward the electrode wicks. The rate of migration is determined by the size of the protein molecule and its overall charge. The proteins thus move in bands across the gel at different rates. Monoclonal bands can be characterised using specific antisera. The proteins are then stained using the SAS 4 machine to allow visualisation and qualitative interpretation.

The gel based system is considered the most appropriate method for characterising small paraproteins & light chains.

Method

Check expiry dates of any reagent before use. Use oldest reagents first.

Worklist

Use the code SFIX in IMMWF. Fill in the worklist sheet (Immunology 1622) to include lab number, patient's name, dilution used (you can leave blank if the screening dilution is to be used) and whether the test has been reflexed from a FLITE result.

Dilutions

Samples are diluted with IFE diluent. If the monoclonal isotype is known, the following table supplied by Helena can be used to determine which dilution to use:

Monoclonal conc^{n} g/l	**SP Lane**	**Volume Sample/diluent µl**	**G, A, M, K, L**	**Volume Sample/diluent µl**
25-50	1 in 3	20/40	1 in 20	10/190
10-25	1 in 3	20/40	1 in 10	20/180
3-10	1 in 3	20/40	1 in 5	40/160
<3	1 in 3	20/40	Neat	-

Lane	**Screening dilution**	**Volume Sample/diluent µl**
Total protein	1 in 3	20/40
G	1 in 7	10/60
A	1 in 3	20/40
M	1 in 2	20/20
K	1 in 3	20/40
L	1 in 4	10/30

For IgD & IgE fixes add 70 µl serum to 140 µl sample diluent (use 1:3 dilution for all lanes – TP, D, E, K, L)

If the monoclonal component is unknown, Helena recommends using either a screening dilution or to run the sample neat.

Place 3 sample dilution cups into a sample tray. If samples are diluted manually, add 35µl of diluted serum for each patient, following templates on sample tray (using the bottom smaller wells). Ensure there are no bubbles.

Immunofixation

1. Switch on SAS 3 (power switch on right hand side).
2. Open cover and press the RIGHT arrow key to move trolley assembly to the right.
3. Place the loaded sample tray on the SAS 3 using the aligning pins.
4. Press the LEFT arrow key to move trolley assembly to its home position.
5. Break off tabs and place 3 applicator blades at positions 2, 8 & 14.
6. Apply 2ml of REP Prep onto gel plate in an L shape. Avoid bubbles.
7. Remove gel from packaging (remember to remove clear overlay). WEAR GLOVES.
8. Place gel onto gel plate using alignment pins (bottom left-hand side first). Avoid bubbles underneath the gel.
9. Blot the gel with thin blotting paper (Blotter C) for 5-10 seconds. Remove paper.
10. Wipe off any excess REP Prep from around the gel.
11. Attach carbon electrodes to the outside of the magnetic posts. Close lid.
12. Select SERUM IFE test from menu (using up/down keys) then press START/STOP button to confirm. Press ENTER to start electrophoresis.
13. When finished remove electrodes and remove gel blocks with yellow gel scraper. WEAR GLOVES.
14. Place antisera template onto gel. Ensure it is centrally aligned.
15. Add 50µl of each anti-serum to the large holes in the following order, Protein fixative, G, A, M, K, L for each patient. Ensure all lanes are filled and keep pipette/bottle upright. For D & E fixes; Protein fixative, D, E, blank, K, L.
16. Close lid & press ENTER to incubate the gel.
17. Following incubation wick away excess antisera using 3 blotter combs inserted into the large holes of the template: insert combs close lid and press ENTER.
18. When finished open lid, remove combs and template. Place a thick piece of blotting paper (Blotter D) onto the gel, smooth side down. Replace antisera template to hold blotter flat. Close lid and press ENTER.
19. At beep, remove template and blotter D. Close lid and press ENTER. Dry gel (8 minutes). Press START/STOP button to silence beep and finish.
20. When finished discard applicators & sample cups. Remove gel from gel plate. Wipe off excess REP Prep from plate. Wipe plate with distilled water on tissue.
21. Switch on SAS 4 (power switch left hand side).
22. Select SERUM IFE test from menu (using up/down keys) then press START/STOP button to confirm then press ENTER. The system will perform a 3 second wash with the gel holder in the instrument.
23. When prompted attach the gel to the staining chamber holder ensuring agarose side faces the back of the machine. Then place gel in chamber of the SAS 4. Press ENTER to start staining program.
24. When finished press START/STOP button, remove the gel and wipe the back to remove any leftover stain.
25. Date gel and read (see below). All gels must be second read by a registered Healthcare Scientist prior to reporting.
26. Scan the gel on the V8. If the samples were pipetted on the V8, ensure that you put the same barcode ID used for the sample tray. This will automatically link the gel scans to the patients. If samples were pipetted manually then you must enter the Lab number for each patient manually and query LIMS to fill all the patient data.

Results analysis/Interpretation

Anti-sera are more sensitive at detecting heavy and free light chains compared with protein fixation. Heavy and free light chains will often appear to be amplified in the anti-sera channels

compared with the serum protein electrophoresis channel and should be interpreted cautiously. Any band(s) in the antisera channels should match up exactly with the corresponding band(s) in the protein channel (use a clear plastic ruler if necessary). If no monoclonal band is seen in the protein channel, bands seen in the anti-sera lane are not clinically significant and merely reflect the greater sensitivity of the anti-sera reagents. It is useful and often necessary to refer to the original CZE trace to confirm that the monoclonal bands actually match.

Alert/Critical values
The following results must be rung to the relevant clinician as soon as possible;
IgG paraproteins >30 g/l, IgA & IgM paraproteins >10 g/l
Any paraproteins with accompanying immunoparesis
All monoclonal light chain positive patients

If in doubt, speak to a senior member of staff.

Oxford University Hospitals
NHS Foundation Trust

Pathology and Laboratories Directorate

The Department of Laboratory Immunology

Standard Operating Procedure (SOP)

Title: Optilite analyser

Code:	Immunology 2190
Document Status:	Under Review
Authors:	Dawn Simpson
Version:	1.6
Section:	Immunochemistry section
Document Category:	Analytical Process
Authorised By:	Dawn Simpson
Location Of Copy:	unknown
Authorised on:	22-Dec-2020
Review on or Before:	22-Aug-2022
Date Time Of Last Update:	12-Jan-2023 12:37
Date Time Printed:	PREVIEW ONLY - UNCONTROLLED
Number Of Copies:	N/A
Print History:	Copy 1: Immunochemistry. Printed: 22-Dec-2020 Copy 2: Back bench. Printed: 15-Jul-2021

This is a controlled document

The master copy of this document is held on iPassport. This document must not be copied or altered without express permission from appropriate management teams. This document will be reviewed on a regular, scheduled basis.

Optilite Analyser

Introduction

The Optilite analyser is an automated turbidimeter provided by The Binding Site. It is used for the analysis of serum for a variety of analytes using Binding Site Optilite products. Turbidimetry measures the decrease in light transmission/scatter through a solution; the amount of scatter is inversely proportional to the amount of particles (e.g. proteins) in the solution held within the reaction cuvettes. The light source for the Optilite is a Xenon lamp. Disposable acrylic cuvettes are housed in a central loader and are placed into the reaction carousel automatically by the analyser. They are used for both sample dilution and assay reactions. Reaction times are assay dependent.

1 - Operation buttons
2 - Main switch
3 - Waste container
4 - Cuvette waste bin
5 - Water container

Calibration

All FreeLITE, HevyLITE, IgG subclass and C1 inhibitor kits contain a calibrator to perform a calibration curve. There must be a valid curve for each test for results to be generated. Calibrations must be performed on changing lot number and every month (sooner if an IQC drift or bias is noted).

FreeLite calibrators consist of human sera containing free light chains and are supplied in a stabilised liquid form. They are calibrated against an internal reference standard (>99% pure by SDS-PAGE) as there is no international reference material available.
HevyLite calibrators are pooled normal human sera in liquid form. An internal reference material is used as there is no international reference standard. Total immunoglobulin values and Hevylite values have been assigned against this. The total values are assigned against DA470K.

OVERVIEW OF START UP

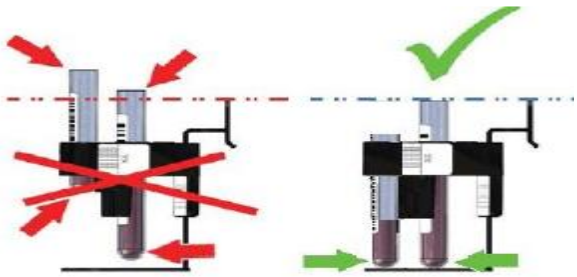
1. Perform Start Up / Daily maintenance
2. Check calibration status of assay/s
3. Review worklists to estimate amount of reagent needed
4. Add reagents appropriate for assay / workload
5. Perform calibrations & QCs where applicable

TESTING BARCODED PATIENT SAMPLES

SST tubes (12x75), LP4 tubes and short SST tubes may be placed directly into the **black** sample racks. False bottom tubes (Inesco) received from Stoke Mandeville must be placed in the **blue**

racks. 2ml tubes received from Milton Keynes etc. must be decanted into fully labelled LP4 tubes. If needed, reprint lab no barcode using LIMS transaction ABCL. **All tubes must be visually inspected for sample volume & 'solidity' to prevent sample probe problems.** Short samples must be transferred to a sample cup (see below). Solid samples (cryoprotein positive &/or viscous samples) must be placed in 37^{O}C incubator to dissolve before analysis.

Remove lids and place samples in correct rack. Ensure barcodes are visible and tubes are inserted correctly.



Press sample cover button; wait for click the open the cover and insert the rack. Close cover. Ensure analyser has read each rack before placing next one on, to prevent communication problems. Repeat for all racks needed (a maximum of 6). Ensure racks are seated correctly. Shorter samples in LP4 tubes may be raised slightly to prevent short sample error. Check tubes are seated straight and no higher than 12x75 tube.
Confirm there are enough reagents and water on board. Also check the waste containers are not full.
Click Start to begin analysis (wait for rack/s to be scanned).

HEVYLITE REQUESTS
After racks have been scanned by the analyser it is possible to pick the relevant dilutions needed (based on previous results) before analysis is started. Write dilutions, if needed, on worklist.

F2> 1 Samples. Highlight patient in list on left hand side.
Highlight test to be changed.
Use drop down menu to select new dilution. Save.
Repeat for other patients, if needed.

RESULTS ANALYSIS/INTERPRETATION
Analyse log files after each run (see below for printing out files). The Optilite will carry out a number of dilutions to generate an absolute value (dilutions are assay specific), check all results and flags make sense.

Accepting Results
Reportable patient sample results will be automatically accepted and sent to AMS. Unaccepted results are usually lower than the limit of detection. To review:
F2> 4 Results
Select Not Accepted to view results needing attention.
If applicable, highlight results and select Accept selected.
Print results: F2 > 4 Results > Choose All> Request type=Sample > Highlight results in strict order (they are printed out in the order they are selected!). Ensure the printer is switched on.

Additional checks must be performed for all **HevyLite** results to look for potential antigen excess; write total involved immunoglobulin level on results sheet along with isotype of paraprotein. Compare HLITE values with the total immunoglobulin level e.g. check total IgG against IgG kappa & IgG lambda.

Reference Ranges

FLITE (mg/L)

Age (years)	Free Kappa	Free Lambda	Ratio
20-40	7.5-16.8	9.1-20.2	0.73-1.48
41-60	9.6-21.6	9.8-22.6	0.87-1.45
61-80	11.3-27.6	10.3-24.4	0.99-1.8
81+	14.2-37	13.7-38	1.0-1.8

Results generated from 4293 individual serum free light results that were obtained on patients from primary care who had serum electrophoresis analysis requested.

HLITE (g/L)

IgG kappa = 3.84-12.07 IgG lambda = 1.91-6.74 K/L ratio = 1.12-3.21
IgA kappa = 0.57-2.08 IgA lambda = 0.44-2.04 K/L ratio = 0.78-1.94
IgM kappa = 0.19-1.63 IgM lambda = 0.12-1.01 K/L ratio = 1.18-2.74

All ranges generated by Binding Site:
The ranges were obtained by measuring the IgG kappa and IgG lambda concentrations of 129 normal sera. The ranges were obtained by measuring the IgA kappa and IgA lambda concentrations of 120 normal sera. The ranges were obtained by measuring the IgM kappa and IgM lambda concentrations of 147 normal sera