National Collaborating Centre for Cancer

Myeloma in adults:

diagnosis and management

NICE Guideline

Full guideline

August 2015

Draft for Consultation

Commissioned by the National Institute for Health and Care Excellence

Disclaimer
Healthcare professionals are expected to take NICE clinical guidelines fully into account when exercising their clinical judgement. However, the guidance does not override the responsibility of healthcare professionals to make decisions appropriate to the circumstances of each patient, in consultation with the patient and/or their guardian or carer.

Funding
Funded to produce guidelines for the NHS by NICE

Contents

1

Key research recommendations

1. Diagnostic investigations to predict treatment outcomes

A prospective randomised multi-centre trial of different treatment strategies should compare the prognostic value of the HevyLite assay and ratio with other prognostic factors and tests, including the serum-free light-chain assay and fluorescence in situ hybridisation (FISH), in people with newly diagnosed myeloma who are starting treatment. Outcomes of interest are overall response, complete response, minimal residual disease, progression-free survival, overall survival and resource use.

2. Imaging investigations for newly diagnosed myeloma

Observational studies should be carried out, comparing the effectiveness of whole-body MRI, fluorodeoxyglucose positron emission tomography CT (FDG PET-CT) and whole-body low-dose CT in detecting lesions that may determine the start of treatment for people with newly diagnosed myeloma. Outcomes of interest are lesion detection, sensitivity and specificity for myeloma-related bone disease, patient acceptability, incremental upstaging, radiation exposure, risk of second primary cancer, the impact of additional information on predicting progression-free survival, overall survival and skeletal-related events.

3. Management of smouldering myeloma

A randomized multi-centre prospective trial should be carried out for patients with newly diagnosed smouldering myeloma (as defined by the International Myeloma Working Group 2014 classification) to:

- identify which combinations of FISH, molecular technologies, bone marrow plasma cell percentage, whole-body imaging, immunophenotype, serum-free light-chain levels or ratio, HevyLite, paraprotein levels, immunoparesis, and International Staging System (ISS) are most effective at risk stratification for people with smouldering myeloma.
- compare fixed duration treatment (with or without bone-directed therapy), continuous treatment (with or without bone-directed therapy) and no treatment (with or without bone-directed therapy).

Outcomes of interest are time to biochemical and/or clinical progression, overall survival, adverse events, quality of life and resource use.

4. Allogeneic stem cell transplantation

Research is needed into the effectiveness of combined autologous-allogeneic stem cell transplantation compared with autologous stem cell transplantation, plus consolidation and maintenance treatment in chemosensitive patients at first response or first relapse. Outcomes of interest are progression-free survival, overall survival, transplant-related mortality, quality of life, early and late toxicity including graft-versus-host-disease (GvHD) and resource use. This research should be included as an option in appropriate mainstream clinical trials for myeloma.

5. Bisphosphonates for the prevention of bone disease

A randomised controlled trial should be carried out, comparing monthly zoledronic acid indefinitely with zoledronic acid for fixed duration in patients with myeloma. Outcomes of interest are skeletal-related events, progression-free survival, overall survival, utility of bone biomarkers, incidence of osteonecrosis of the jaw, quality of life and resource use.

Methodology

What is a clinical guideline?

Guidelines are recommendations for the care of individuals in specific clinical conditions or circumstances – from prevention and self-care through to primary and secondary care and onto more specialised services. NICE clinical guidelines are based on the best available evidence of clinical and cost effectiveness, and are produced to help healthcare professionals and patients make informed choices about appropriate healthcare. While guidelines assist the practice of healthcare professionals, they do not replace their knowledge and skills.

Who is the guideline intended for?

This guideline does not include recommendations covering every detail of the diagnosis and management of myeloma. Instead, this guideline has tried to focus on those areas of clinical practice (i) that are known to be controversial or uncertain; (ii) where there is identifiable practice variation; (iii) where there is a lack of high quality evidence; or (iv) where NICE guidelines are likely to have most impact. More detail on how this was achieved is presented later in the section on 'Developing clinical evidence based questions'.

This guideline is relevant to all healthcare professionals who come into contact with people with myeloma, as well as to the people with myeloma themselves and their carers. It is also expected that the guideline will be of value to those involved in clinical governance in both primary and secondary care to help ensure that arrangements are in place to deliver appropriate care to this group of people.

The remit of the guideline

Involvement of Stakeholders

Key to the development of all NICE guidelines are the relevant professional and patient/carer organisations that register as stakeholders. Details of this process can be found on the NICE website or in the 'NICE guidelines manual' (NICE 2014). In brief, their contribution involves commenting on the draft scope, submitting relevant evidence and commenting on the draft version of the guideline during the end consultation period. A full list of all stakeholder organisations who registered for the myeloma guideline can be found in Appendix F.

The guideline development process – who develops the guideline?

Overview

The development of this guideline was based upon methods outlined in the 'NICE guidelines manual' (NICE 2012). A team of health professionals, lay representatives and technical experts known as the Guideline Committee (GC) (Appendix F), with support from the NCC-C staff, undertook the development of this clinical guideline. The basic steps in the process of developing a guideline are listed and discussed below:

- using the remit, define the scope which sets the inclusion/exclusion criteria of the guideline
- forming the GC

- developing clinical questions
- identifying the health economic priorities
- developing the review protocols
- systematically searching for the evidence
- critically appraising the evidence
- incorporating health economic evidence
- distilling and synthesising the evidence and writing recommendations
- agreeing the recommendations
- structuring and writing the guideline
- consultation and validation

The scope

The scope was drafted by the GC Chair and Lead Clinician and staff at the NCC-C in accordance with processes established by NICE (NICE 2012). The purpose of the scope was to:

- set the boundaries of the development work and provide a clear framework to enable work to stay within the priorities agreed by NICE and the NCC-C
- inform professionals and the public about the expected content of the guideline
- provide an overview of the population and healthcare settings the guideline would include and exclude
- specify the key clinical issues that will be covered by the guideline
- inform the development of the clinical questions and search strategies

Before the guideline development process started, the draft scope was presented and discussed at a stakeholder workshop. The list of key clinical issues were discussed and revised before the formal consultation process. Further details of the discussion at the stakeholder workshop can be found on the NICE website (www.nice.org.uk).

The scope was subject to a four week stakeholder consultation in accordance with NICE processes. The full scope is shown in Appendix E. During the consultation period, the scope was posted on the NICE website. Comments were invited from registered stakeholder organisations and NICE staff. The NCC-C and NICE reviewed the scope in light of comments received, and the revised scope was reviewed and signed off by NICE and posted on the NICE website.

The Guideline Committee (GC)

The myeloma GC was recruited in line with the 'NICE guidelines manual' (NICE 2012). The first step was to appoint a Chair and a Lead Clinician. Advertisements were placed for both posts and shortlisted candidates were interviewed in person prior to being offered the role. The NCC-C Director, GC Chair and Lead Clinician identified a list of specialties that needed to be represented on the GC. Details of the adverts were sent to the main stakeholder organisations, cancer networks and patient organisations/charities (Appendix F). Individual GC members were selected for telephone interview by the NCC-C Director, GC Chair and Lead Clinician, based on their application forms. The guideline development process was supported by staff from the NCC-C, who undertook the clinical and health economics literature searches, reviewed and presented the evidence to the GC, managed the process and contributed to drafting the guideline. At the start of the guideline development process all GC members' interests were recorded on a standard declaration form that covered consultancies, fee-paid work, share-holdings, fellowships and support from the healthcare

industry. At all subsequent GC meetings, members declared new, arising conflicts of interest which were always recorded (see Appendix F).

Guideline Committee meetings

Thirteen GC meetings were held between 27-28 March 2014 and 5–6 November 2015. During each GC meeting (held over either 1 or 2 days) clinical questions and clinical and economic evidence were reviewed, assessed and recommendations formulated. At each meeting patient/carer and service-user concerns were routinely discussed.

NCC-C project managers divided the GC workload by allocating specific clinical questions, relevant to their area of clinical practice, to small sub-groups of the GC in order to simplify and speed up the guideline development process. These groups considered the evidence, as reviewed by the researcher, and synthesised it into draft recommendations before presenting it to the GC. These recommendations were then discussed and agreed by the GC as a whole. Each clinical question was led by a GC member with expert knowledge of the clinical area (usually one of the healthcare professionals). The GC subgroups often helped refine the clinical questions and the clinical definitions of treatments. They also assisted the NCC-C team in drafting the section of the guideline relevant to their specific topic.

Patient/carer members

Individuals with direct experience of myeloma services gave an important user focus to the GC and the guideline development process. The GC included three patient/carer members. They contributed as full GC members to writing the clinical questions, helping to ensure that the evidence addressed their views and preferences, highlighting sensitive issues and terminology relevant to the guideline and bringing service-user research to the attention of the GC.

Expert advisers

During the development of the guideline the GC identified the management of non spinal and spinal bone disease as topics that required additional expert input. Four experts were identified by the NCC-C and GC (Appendix F) and invited to advise the GC on drafting their recommendations for that clinical question.

Developing clinical evidence-based questions

Background

Clinical guidelines should be aimed at changing clinical practice and should avoid ending up as ‘evidence-based textbooks’ or making recommendations on topics where there is already agreed clinical practice. Therefore the key clinical issues listed in the scope were developed for areas that were known to be controversial or uncertain, where there was identifiable practice variation, or where NICE guidelines were likely to have most impact.

Method

From each of the key clinical issues identified in the scope, the GC formulated a clinical question. For clinical questions about interventions, the PICO framework was used. This structured approach divides each question into four components: P – the population (the population under study); I – the interventions (what is being done); C – the comparison (other main treatment options); O – the outcomes (the measures of how effective the interventions have been).

Review of Clinical Literature

Scoping search

An initial scoping search for published guidelines, systematic reviews, economic evaluations and ongoing research was carried out on the following databases or websites: NHS Evidence, NICE, Cochrane Databases of Systematic Reviews (CDSR), Health Technology Assessment Database (HTA), TRIP, SIGN, NHS Economic Evaluations Database (NHSEED), Health Economic Evaluations Database (HEED), Medline and Embase.

At the beginning of the development phase, initial scoping searches were carried out to identify any relevant guidelines (local, national or international) produced by other groups or institutions.

Developing the review protocol

For each clinical question, the information specialist and researcher (with input from other technical team and GC members) prepared a review protocol. This protocol explains how the review was to be carried out (Table 1) in order to develop a plan of how to review the evidence, limit the introduction of bias and for the purposes of reproducibility. All review protocols can be found in the evidence review (Appendix G).

Table 1: Components of the review protocol

Component	Description
Clinical question	The clinical question as agreed by the GC
Rationale	An explanation of why the clinical question is important. For example, is the topic contentious? Is there variation in practice across the UK?
Criteria for considering studies for the review	Using the PICO (population, intervention, comparison and outcome) framework. Including the study designs selected.
How the information will be searched	The sources to be searched and any limits that will be applied to the search strategies; for example, publication date, study design, language. Searches should not necessarily be restricted to RCTs.
The review strategy	The methods that will be used to review the evidence, outlining exceptions and subgroups. Indicate if meta-analysis will be used.

Searching for the evidence

In order to answer each question the NCC-C information specialist developed a search strategy to identify relevant published evidence for both clinical and cost effectiveness. Key words and terms for the search were agreed in collaboration with the GC. When required, the health economist searched for supplementary papers to inform detailed health economic work (see section on 'Incorporating Health Economic Evidence').

Search filters, such as those to identify systematic reviews (SRs) and randomised controlled trials (RCTs) were applied to the search strategies when necessary. No language restrictions were applied to the search; however, foreign language papers were not requested or reviewed (unless of particular importance to that question).

The following databases were included in the literature search:

- The Cochrane Library
- Medline and Premedline 1946 onwards
- Excerpta Medica (Embase) 1974 onwards
- Web of Science [specifically Science Citation Index Expanded (SCI-Expanded) 1900 onwards and Social Sciences Citation Index (SSCI) 1900 onwards]

Subject specific databases used for certain topics:

- Cumulative Index to Nursing and Allied Health Literature (CINAHL) 1937 onwards
- PsycINFO 1806 onwards
- Amed 1985 onwards

From this list the information specialist sifted and removed any irrelevant material based on the title or abstract before passing to the researcher. All the remaining articles were then stored in a Reference Manager electronic library.

In accordance with the 'NICE guidelines manual' (NICE 2012) searches were updated and re-run 6–8 weeks before the guideline was submitted to NICE for stakeholder consultation, thereby ensuring that the latest relevant published evidence was included in the database. Any evidence published after this date was not included. For the purposes of updating this guideline, 8th June 2015 should be considered the starting point for searching for new evidence.

Further details of the search strategies, including the methodological filters used, are provided in the evidence review (Appendix G).

Critical Appraisal and Evidence Grading

Following the literature search one researcher independently scanned the titles and abstracts of every article for each question, and full publications were obtained for any studies considered relevant or where there was insufficient information from the title and abstract to make a decision. When papers were obtained, the researcher applied inclusion/exclusion criteria to select appropriate studies, which were then critically appraised. If results from a study were published as more than one paper, the most recent or complete publication was used. For each question, data were extracted and recorded in evidence tables and an accompanying evidence summary prepared for the GC (see Appendix G). All evidence was considered carefully by the GC for accuracy and completeness.

GRADE (Grading of Recommendations, Assessment, Development and Evaluation)

For interventional questions, studies which matched the inclusion criteria were evaluated and presented using GRADE (NICE 2012; http://gradeworkinggroup.org/). Where possible this included meta-analysis and synthesis of data into a GRADE 'evidence profile'. The evidence profile shows, for each outcome, an overall assessment of both the quality of the evidence as a whole (very low, low, moderate or high) as well as an estimate of the size of effect. A narrative summary (evidence statement) was also prepared.

Each outcome was examined for the quality elements defined in Table 2 and subsequently graded using the quality levels listed in Table 3. The reasons for downgrading or upgrading specific outcomes were explained in footnotes.

Table 2: Descriptions of quality elements of GRADE

Quality element	Description
Limitations	Limitations in the study design and implementation may bias the estimates of the treatment effect. Major limitations in studies decrease the confidence in the estimate of the effect
Inconsistency	Inconsistency refers to unexplained heterogeneity of results
Indirectness	Indirectness refers to differences in study population, intervention, comparator or outcomes between the available evidence and the clinical question
Imprecision	Results are imprecise when studies include relatively few patients and few events and thus have wide confidence intervals around the estimate of the effect

Quality element	Description
Publication bias	Publication bias is a systematic underestimate or overestimate of the underlying beneficial or harmful effect due to the selective publication of studies

Table 3: Overall quality of outcome evidence in GRADE

Quality element	Description
High	Further research is very unlikely to change our confidence in the estimate of effect
Moderate	Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate
Low	Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate
Very low	Any estimate of effect is very uncertain

All procedures were fully compliant with NICE methodology as detailed in the 'NICE guidelines manual' (NICE 2012). In general, evidence was based on published data only. Study authors were contacted only to resolve any ambiguities, such as unclear presentation of data, or where clarification was needed in order to include or exclude a paper in the evidence review.

For non-interventional questions, for example questions regarding diagnostic test accuracy, a narrative summary of the quality of the evidence was provided. The quality of individual diagnostic accuracy studies was assessed using the QUADAS-2 tool (Whiting et al., 2011).

Incorporating health economics evidence

The aim of providing economic input into the development of the guideline was to inform the GC of potential economic issues relating to myeloma. Health economics is about improving the health of the population through the efficient use of resources. In addition to assessing clinical effectiveness, it is important to investigate whether health services are being used in a cost effective manner in order to maximise health gain from available resources.

Prioritising topics for economic analysis

After the clinical questions had been defined, and with the help of the health economist, the GC discussed and agreed which of the clinical questions were potential priorities for economic analysis. These economic priorities were chosen on the basis of the following criteria, in broad accordance with the NICE guidelines manual (NICE 2012):

- the overall importance of the recommendation, which may be a function of the number of patients affected and the potential impact on costs and health outcomes per patient
- the current extent of uncertainty over cost effectiveness, and the likelihood that economic analysis will reduce this uncertainty
- the feasibility of building an economic model

A review of the economic literature was conducted at scoping. Where published economic evaluation studies were identified that addressed the economic issues for a clinical question, these are presented alongside the clinical evidence.

For systematic searches of published economic evidence, the following databases were included:

- Medline
- Embase
- NHS Economic Evaluation Database (NHS EED)

- Health Technology Assessment (HTA)
- Health Economic Evaluations Database (HEED)

Methods for reviewing and appraising economic evidence

The aim of reviewing and appraising the existing economic literature is to identify relevant economic evaluations that compare both costs and health consequences of alternative interventions and that are applicable to NHS practice. Thus studies that only report costs, non-comparative studies of 'cost of illness' studies are generally excluded from the reviews (NICE 2012).

Economic studies identified through a systematic search of the literature are appraised using a methodology checklist designed for economic evaluations (NICE 2012). This checklist is not intended to judge the quality of a study per se, but to determine whether an existing economic evaluation is useful to inform the decision-making of the GC for a specific topic within the guideline. There are two parts of the appraisal process; the first step is to assess applicability (i.e. the relevance of the study to the specific guideline topic and the NICE reference case) (Table 4).

Table 4: Applicability criteria

Directly applicable	The study meets all applicability criteria, or fails to meet one or more applicability criteria but this is unlikely to change the conclusions about cost effectiveness
Partially applicable	The study fails to meet one or more applicability criteria, and this could change the conclusions about cost effectiveness
Not applicable	The study fails to meet one or more applicability criteria, and this is likely to change the conclusions about cost effectiveness. These studies are excluded from further consideration

In the second step, only those studies deemed directly or partially applicable are further assessed for limitations (i.e. the methodological quality, Table 5).

Table 5: Methodological quality

Minor limitations	Meets all quality criteria, or fails to meet one or more quality criteria but this is unlikely to change the conclusions about cost effectiveness
Potentially serious limitations	Fails to meet one or more quality criteria and this could change the conclusions about cost effectiveness
Very serious limitations	Fails to meet one or more quality criteria and this is highly likely to change the conclusions about cost effectiveness. Such studies should usually be excluded from further consideration

Where relevant, a summary of the main findings from the systematic search, review and appraisal of economic evidence is presented in an economic evidence profile alongside the clinical evidence.

If high-quality published economic evidence relevant to current NHS practice was identified through the search, the existing literature was reviewed and appraised as described above. However, it is often the case that published economic studies may not be directly relevant to the specific clinical question as defined in the guideline or may not be comprehensive or conclusive enough to inform UK practice. In such cases, for priority topics, consideration was given to undertaking a new economic analysis as part of this guideline.

Economic modelling

Once the need for a new economic analysis for high priority topics had been agreed by the GC, the health economist investigated the feasibility of developing an economic model. In the development of the analysis, the following general principles were adhered to:

- the GC subgroup was consulted during the construction and interpretation of the analysis
- the analysis was based on the best available clinical evidence from the systematic review
- assumptions were reported fully and transparently
- uncertainty was explored through sensitivity analysis
- costs were calculated from a health services perspective
- outcomes were reported in terms of quality-adjusted life years

Linking to NICE technology appraisals

There are several published technology appraisals (TAs) which are relevant to this guideline (TA129, 171, 228, 311 and 338- see www.nice.org.uk/TA/published). In line with NICE methodology, the recommendations from these TAs have either been cross-referenced (TA171, 311 and 338) or incorporated (TA228 and 129). (See Developing NICE guidelines: the manual).

Agreeing the recommendations

For each clinical question the GC were presented with a summary of the clinical evidence, and, where appropriate, economic evidence, derived from the studies reviewed and appraised. The GC derived their guideline recommendations from this information. The link between the evidence and the view of the GC in making each recommendation is made explicitly in the accompanying LETR statement (see below).

Wording of the recommendations

The wording used in the recommendations in this guideline denotes the certainty with which the recommendations were made. Some recommendations were made with more certainty than others. Recommendations are based on the trade-off between the benefits and harms of an intervention, whilst taking into account the quality of the underpinning evidence.

For all recommendations, it is expected that a discussion will take place with the patients about the risks and benefits of the interventions, and their values and preferences. This discussion should help the patient reach a fully informed decision. Terms used within this guideline are:

- 'Offer' – for the vast majority of patients, an intervention will do more good than harm
- 'Do not offer' – the intervention will not be of benefit for most patients
- 'Consider' – the benefit is less certain, and an intervention will do more good than harm for most patients. The choice of intervention, and whether or not to have the intervention at all, is more likely to depend on the patient's values and preferences than for an 'offer' recommendation, and so the healthcare professional should spend more time considering and discussing the options with the patient.

LETR (Linking evidence to recommendations) statements

As clinical guidelines were previously formatted, there was limited scope for expressing how and why a GC made a particular recommendation from the evidence of clinical and cost effectiveness. To make this process more transparent to the reader, NICE have introduced an explicit, easily understood and consistent way of expressing the reasons for making each

recommendation. This is known as the 'LETR statement' and will usually cover the following key points:

- the relative value placed on the outcomes considered
- the strength of evidence about benefits and harms for the intervention being considered
- the costs and cost effectiveness of an intervention
- the quality of the evidence (see GRADE)
- the degree of consensus within the GC
- other considerations – for example equalities issues

Where evidence was weak or lacking the GC agreed the final recommendations through informal consensus. Shortly before the consultation period five key research recommendations were selected by the GDG for implementation and the patient algorithms were agreed.

Consultation and validation of the guideline

The draft of the guideline was prepared by NCC-C staff in partnership with the GC Chair and Lead Clinician. This was then discussed and agreed with the GC and subsequently forwarded to NICE for consultation with stakeholders.

Registered stakeholders (Appendix F) had one opportunity to comment on the draft guideline which was posted on the NICE website between 19 August 2015 and 1 October 2015 in line with NICE methodology (NICE 2014).

Guideline implementation

NICE invited stakeholders to give their responses to the following questions during consultation of the guideline:

- Which areas will have the biggest impact on practice and be challenging to implement? Please say for whom and why.
- What would help users overcome any challenges? (For example, existing practical resources or national initiatives, or examples of good practice.)

NICE will use the feedback received as well as consultation with members of the committee, engagement with relevant key partners and relevant desk research, to write a chapter which aims to help users of the guideline to get started with implementation. It will highlight up to 3 areas for attention, describing the benefits, barriers and enablers as well as signposting to any relevant resources or examples of practice that may help.

The pre-publication process

An embargoed pre-publication version of the guideline was released to registered stakeholders who have signed a confidentiality form to allow them to see how their comments have contributed to the development of the guideline and to give them time to prepare for publication (NICE 2014).

The final document was then submitted to NICE for publication on their website. The other versions of the guideline (see below) were also discussed and approved by the GC and published at the same time.

Other versions of the guideline

This full version of the guideline is available to download free of charge from the NICE website (www.nice.org.uk) and the NCC-C website (www.wales.nhs.uk/nccc).

NICE also produces three other versions of the myeloma guideline which are available from the NICE website:

- the short version, containing all recommendations and the key research recommendations.
- NICE pathways, which is an online tool for health and social care professionals that brings together all related NICE guidance and associated products in a set of interactive topic-based diagrams.
- 'Information for the Public (IFP)', which summarises the recommendations in the guideline in everyday language for patients, their family and carers, and the wider public.

Updating the guideline

Literature searches were repeated for all of the clinical questions at the end of the guideline development process, allowing any relevant papers published before 8th June 2015to be considered. Future guideline updates will consider evidence published after this cut-off date.

A formal review of the need to update a guideline is usually undertaken by NICE after its publication. NICE will conduct a review to determine whether the evidence base has progressed significantly to alter the guideline recommendations and warrant an update.

Funding

The National Collaborating Centre for Cancer (NCC-C) was commissioned by NICE to develop this guideline.

Disclaimer

The GC assumes that healthcare professionals will use clinical judgement, knowledge and expertise when deciding whether it is appropriate to apply these guidelines. The recommendations cited here are a guide and may not be appropriate for use in all situations. The decision to adopt any of the recommendations cited here must be made by the practitioner in light of individual patient circumstances, the wishes of the patient and clinical expertise.

The NCC-C disclaims any responsibility for damages arising out of the use or non-use of these guidelines and the literature used in support of these guidelines.

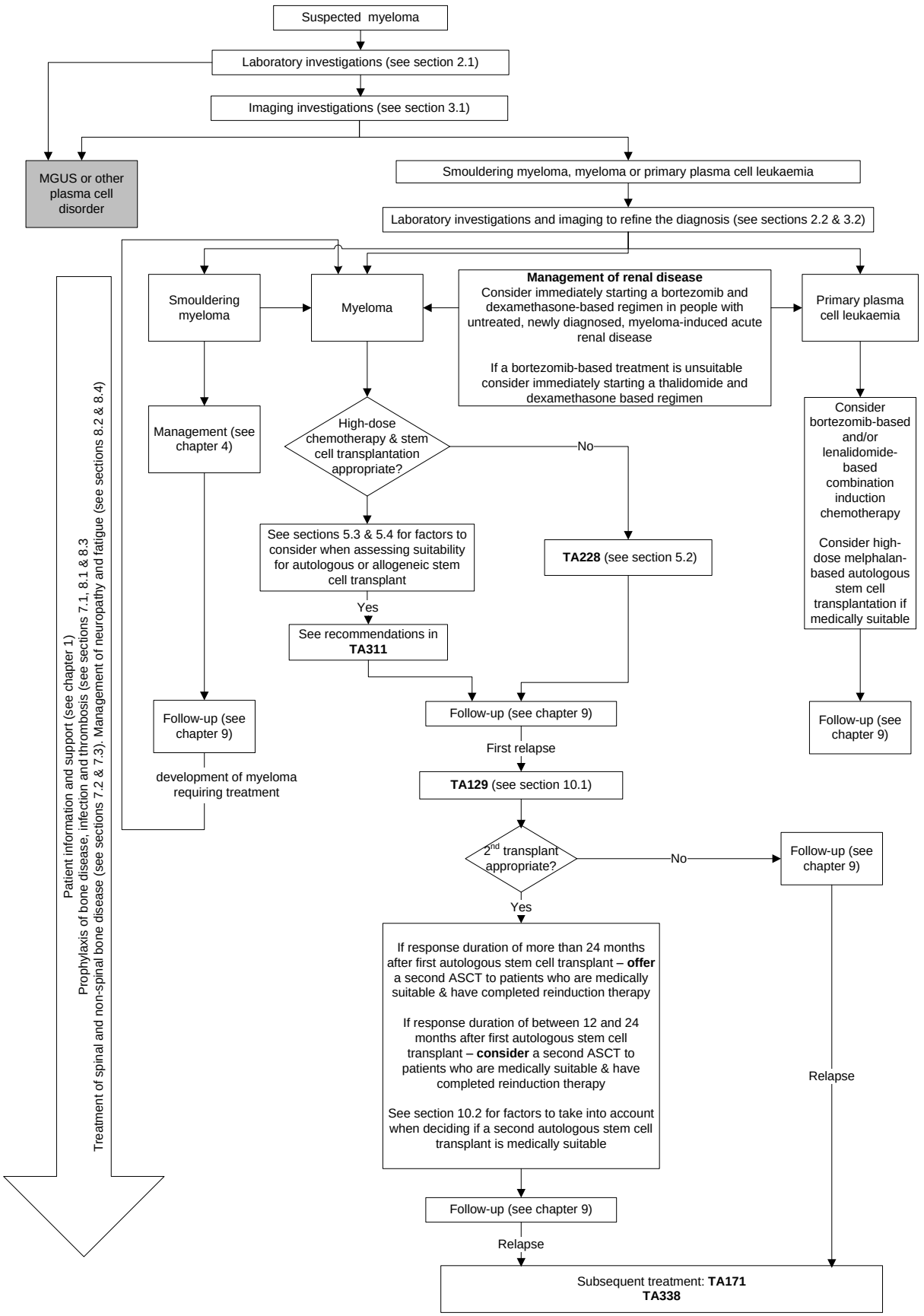
References

National Institute for Health and Clinical Excellence (2012) The guidelines manual. London: National Institute for Health and Clinical Excellence. Available from The guidelines manual

National Institute for Health and Clinical Excellence (2014) Developing NICE guidelines: the manual. London: National Institute for Health and Clinical Excellence. Available from Developing NICE guidelines: the manual

Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MMG, Sterne JAC, Bossuyt PMM, Group Q-2 (2011) QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Annals of Internal Medicine, 155: 529-536.

1 Algorithm



2
3

1 Communication and Support

Myeloma is a rare, complex cancer, and many people have not heard of it at the point of diagnosis. Myeloma is treatable but not curable and requires multiple lines of treatment usually involving chemotherapy with or without stem cell transplantation.

High quality, appropriate and clear individualised information, at different points in the patient pathway is essential, including a clear care plan, changed as necessary. Some patients, carers and their families may want to know all the information available, while others may wish to know little or nothing.

The specific information needs of myeloma patients will depend on the method of disease presentation and the organs which are most affected by their myeloma. Information needs and content will be greatly influenced by the patient's age, fitness, social circumstances, cultural background and by other parameters such as prognostic factors. It is important that whatever the patient's information needs, these are communicated clearly and carefully to alleviate the psychological impact of the prognosis. The sharing of information between secondary and primary care and the multiple community and palliative care teams is essential, particularly given that the majority of myeloma care is based in ambulatory and day unit settings.

There are many differences in the experiences of myeloma patients and their families in relation to the information and support received during diagnosis, treatment, follow-up and end of life care. Patients and carers often report either too little or too much information, leading to poor patient experience. Whilst it is important to understand the information needs at an individual level, it is also important that there is consensus across all centres on the minimum information given, by whom and when, to ensure that informed consent, and patient understanding, is achieved at each stage. Many aspects of patient's information and support needs are covered by NICE guidance on patient experience in adult NHS services. However myeloma patients have additional specific needs.

Clinical question: What are the specific information and support needs of patients with myeloma and their families and carers?

Clinical evidence (see also Appendix G)

Study quality

Evidence about the information and support needs of patients with myeloma and carers was identified from 14 studies (Boland et al 2014, Kelly & Dowling 2011, Lamers et al., 2013, Maher & De Vries, 2001, McGrath et al 2013, Molassiotis et al., 2011a, Molassiotis et al., 2011b, Oerlemans et al., 2012, Osborne et al, 2014, Rini et al., 2007, Spencer et al 2014, Tariman et al, 2014, Vlossak & Fitch 2008 and Myeloma UK survey 2014), which were either qualitative interview (n=9) or questionnaire studies (n=5). All 14 studies addressed the needs of patients whilst 3 studies also examined carer needs. The studies are limited by the small numbers of participants which were recruited from single cancer centres/hospitals. Also, people who participate in these questionnaire/interview studies may have information and support needs that are not representative of other myeloma patients/carers. Furthermore, recall bias may have been present in some studies where participants were asked to retrospectively recall the information and support that was provided.

Eight studies (Kelly & Dowling 2011, Lamers et al., 2013, McGrath et al 2013, Oerlemans et al., 2012, Rini et al., 2007, Spencer et al 2014, Tariman et al, 2014 and Vlossak & Fitch

2008) were conducted in countries other than the UK, so their relevance to current UK practice may be limited.

The evidence identified was all qualitative and assessed as being of moderate quality using the NICE qualitative study checklist.

Information and support needs of myeloma patients

The evidence suggests that the unmet information needs of myeloma patients are low, and patients are generally satisfied with the information they receive. The most common unmet information needs surrounded the need for patients to know more about their future prognosis and include the cause and course of disease as well as side effects and long-term effects of treatment. A common theme throughout the evidence was that patients are interested in experiential information (information from other myeloma patients' experiences). Many patients who had access to such information found it helpful and those who didn't have access to such information would have liked it. However there were some patients who found experiential information unhelpful or even harmful. Evidence from one study on palliative care demonstrated that information on palliative care was not easily available and most patients who were aware of palliative care gained their information from personal experiences they had in the past. There was a contrast between some participants wanting early discussions on palliative care and some only wanting information when needed.

With regards to support needs the evidence suggests that the majority of the unmet support needs of myeloma patients are emotional and psychosocial. In the identified studies many patients were anxious (8-27%) or depressed (5-25%) and many patients desired psychosocial interventions. The most common preferences were relaxation and counselling. Other common support needs include continuity of care, seeing the person in the patient, more time with healthcare professionals and support to manage ongoing symptoms such as fatigue, pain and mobility.

Information and support needs of carers

Evidence concerning carers determined that carers' information needs were in relation to understanding myeloma symptoms better and what is normal, financial advice and information around prognosis.

Whilst the most frequently reported unmet supportive care needs of the carers were the same as the patients, the partners had their own additional needs that were not reported by patients. Additional partner needs were mostly around the practical and informational aspects of the patients care: the need for help to manage ongoing side effects and/or complications experienced by patients as a result of their treatment, provision of up-to-date information, local health-care services that are available when the patient requires them, help in dealing with changes that myeloma has caused to the patient, emotional support to themselves, information to be provided in a way that they can understand.

Anxiety and depression were common in carers with anxiety being higher in partners than in patients.

Cost-effectiveness evidence

A literature review of published cost-effectiveness analyses did not identify any relevant papers for this topic. Whilst there were potential cost implications of making recommendations in this area, other questions in the guideline were agreed as higher priorities for economic evaluation. Consequently no further economic modelling was undertaken for this question.

1

Recommendations	**Provide information and support to people with myeloma or primary plasma cell leukaemia and their family members or carers (as appropriate) at diagnosis, at the beginning and end of each treatment, at disease progression and at transition to end of life care.** **Consider providing the following information in an individualised manner to people with myeloma and their family members or carers (as appropriate):** • **the disease process, relapse and remission cycle, and the person's overall prognosis** • **the treatment plan, including (if appropriate) the process and the potential benefits, risks and complications of stem cell transplantation** • **symptoms of myeloma and treatment-related side effects (including steroid-related side effects, infection and neuropathy)** • **lifestyle measures to optimise bone health and renal function** • **how to identify and report new symptoms (especially pain and spinal cord compression)** • **the role of supportive and palliative care** • **how to access peer support and patient support groups.** **Offer prompt psychological assessment and support to people with myeloma at diagnosis and (as appropriate) at the beginning and end of each treatment, at disease progression and at transition to end of life care.** **Refer people who are assessed as needing from further psychological support (see previous recommendation) to psychological services.** **Advise family members or carers (as appropriate) about available support services at diagnosis, at the beginning and end of each treatment, at disease progression and at transition to end of life care.** **For guidance on communication and patient-centred care see the NICE guideline on patient experience in adult NHS services.**
Relative value placed on the outcomes considered	The Guideline Committee considered the information and support needs reported by patients and carers to be the most important outcome of this review question. A number of potential themes were listed in the review question as examples but this was not an exhaustive list. Evidence was presented on any information and support needs reported by patients with myeloma and their carers that were identified in the literature. Themes from the review question that were not reported in the literature were education, pregnancy/fertility issues, advance care planning, online services and perceived problems with the number of specialists/sites involved in care.
Quality of the evidence	The evidence identified was all qualitative and assessed as being of moderate quality using the NICE qualitative study checklist. The qualitative studies were generally well conducted and provided rich data about patients' experiences and more limited data on carer

	experiences. However, the qualitative studies that were found were limited because the data produced might not generalise to other people or other settings (i.e. the findings are unique to the relatively few people included in the study). Apart from the qualitative nature of the studies some other limitations with the evidence were identified. Most notably, many of the studies had small sample sizes and the majority were not UK based studies. Also the studies were cross-sectional in their design and thus limited in that they only captured data about patient information and support needs at one point in time. It is likely that patient/carer opinions could have changed over time but this would not have been recorded. Some studies required patients to retrospectively report their experiences, and so may be subject to participant recall bias. Many of the information and support needs identified were general to cancer and not specific to myeloma. Due to these limitations the Guideline Committee used their expertise and clinical experience to identify the information gaps that were specific to patients with myeloma. This is why they made a consider reccommendation rather than an offer reccomendation to reflect the strength of the evidence
Trade off between clinical benefits and harms	The Guideline Committee considered that the potential benefits of timeline, individualised communication would include improved patient and carer information and support, leading to improved knowledge, better symptom control and improved quality of life. The Guideline Committee also agreed that improved communication could potentially lead to increased anxiety, resulting from either insufficient or excessive information. The Guideline Committee considered that the majority of patients were likely to benefit from the recommendations. To minimise the potential harm of information overload the Guideline Committee recommended that the information is provided in an individualised manner.
Trade-off between net health benefits and resource use	No health economic evidence was identified and no economic model was developed for this review question. The Guideline Committee considered that there was likely to be an increased workload for healthcare professionals from providing information and support. They also agreed that there may be potential additional costs resulting from the recommendation for onward referral to psychological services. However the Guideline Committee noted that the recommendations should result in better informed and supported patients. They considered that this was likely to reduce emergency admissions, resulting in a corresponding decrease in costs. The Guideline Committee considered that overall these recommendations may result in cost savings from improved efficiency.

References

Boland EG, Boland JW, Ezaydi Y, et al. (2014) Holistic needs assessment in advanced, intensively treated multiple myeloma patients. Support Care Cancer. 2014 Apr 15. [Epub ahead of print]

Kelly M & Dowling M (2011) Patients' lived experience of myeloma. Nursing Standard, 25: 38-44.

Lamers J, Hartmann M, Goldschmidt H et al. (2013) Psychosocial support in patients with multiple myeloma at time of diagnosis: who wants what? Psycho-Oncology, 22: 2313-2320.

Maher K & de VK (2011) An exploration of the lived experiences of individuals with relapsed multiple myeloma. European Journal of Cancer Care, 20: 267-275.

McGrath P (2013) End-of-life care in hematology: update from Australia. Journal Of Social Work In End-Of-Life & Palliative Care, 9: 96-110.

Molassiotis A, Wilson B, Blair S et al. (2011a) Unmet supportive care needs, psychological well-being and quality of life in patients living with multiple myeloma and their partners. Psycho-Oncology, 20: 88-97.

Molassiotis A, Wilson B, Blair S, et al. (2011b) Living with multiple myeloma: experiences of patients and their informal caregivers. Supportive Care in Cancer, 19: 101-111.

Myeloma UK survey. March 2014. Understanding patient experience of high-dose therapy and stem cell transplantation in myeloma.

Oerlemans S, Husson O, Mols F et al. (2012) Perceived information provision and satisfaction among lymphoma and multiple myeloma survivors--results from a Dutch population-based study. Annals of Hematology, 91: 1587-1595.

Osborne TR (2014). Understanding what matters most to people with multiple myeloma: a qualitative study of views on quality of life. BMC Cancer, 14, 496.

Rini C (2007) Peer mentoring and survivors' stories for cancer patients: Positive effects and some cautionary notes. Journal of Clinical Oncology, 25: 163-166.

Stephens M (2014). The work of living with a rare cancer: multiple myeloma. Journal of Advanced Nursing, 70, 2800-2809.

Tariman JD (2014). Patient, physician and contextual factors are influential in the treatment decision making of older adults newly diagnosed with symptomatic myeloma. Cancer Treatment Communications, 2, 34-47.

Vlossak D & Fitch MI (2008) Multiple myeloma: the patient's perspective. Canadian Oncology Nursing Journal, 18: 141-151

2 Laboratory investigations

2.1 Laboratory investigations for people with suspected myeloma

A diagnosis of myeloma may be suspected as a result of a wide range of clinical features and laboratory abnormalities, and in some cases incidental laboratory findings. The key question in diagnosis is to establish whether the individual has (symptomatic) myeloma requiring treatment, smouldering (asymptomatic) myeloma or the precursor condition monoclonal gammopathy of undetermined significance (MGUS), as defined by the International Myeloma Working Group (IMWG) (Rajkumar et al, 2014). The latter two conditions can remain stable for many years and may not ever progress to myeloma requiring treatment. MGUS has an approximately 1% per year risk of progression to myeloma whereas smouldering myeloma has an approximately 10% per year risk of progressing to myeloma during the first five years from diagnosis decreasing thereafter.

Unlike other haematological malignancies, the diagnosis of myeloma is not based on a single test such as a bone marrow or lymph node biopsy but on a combination of clinical features, laboratory tests and radiological findings (these are covered in chapter 3). The laboratory tests used to diagnose myeloma include the examination of bone marrow to show plasma cell infiltration, detection and quantification of monoclonal protein (M protein/M band/paraprotein) in the serum and/or urine, assessment for hypercalcaemia, renal impairment, anaemia, immunosuppression and hyperviscosity. A bone marrow biopsy is a potentially painful invasive test and therefore it is preferable to undertake diagnostic and prognostic tests simultaneously rather than repeat the procedure (see section 2.2 for prognostic tests).

There are now a number of different tests available for the detection and estimation of abnormal monoclonal proteins in serum and urine. There is currently variation in which tests or combinations of these tests are used and when they are done. This question looks at the optimal strategy for using the currently available tests.

Clinical question: What is the optimal laboratory testing strategy for suspected myeloma?

Clinical evidence (see also Appendix G)

Study Quality

The studies were at generally low risk of bias and there were few applicability concerns. There was an unclear risk of bias due to reference standard and flow/timing, due to poor reporting. Three studies had unclear applicability concerns due to patient selection (Park 2012, Cirit 2012, and Hutchison 2008) because they included only patients with renal failure. In other studies there were applicability concerns because patients were included on the basis of the index test results (e.g. Bergon 2010, Frebert 2011). In Katzmann (2005) although myeloma patients were the largest group their results were excluded from the analysis. For studies looking at discrimination of myeloma from MGUS, the reference standard consensus diagnostic criteria often included the index test itself.

Figure 1: Study quality assessment

Diagnostic accuracy of laboratory tests for suspected plasma cell disorders

Serum protein electrophoresis (SPE)

Evidence from 4 studies including 4888 patients (McTaggart et al 2013, Hill et al 2006, Piehler et al 2008 and Vermeersch et al 2008) suggests serum protein electrophoresis has sensitivity 85% [95%C.I. 75% – 92%] and specificity of 95% [95%C.I. 85% – 98%] for the diagnosis of plasma cell disorders.

Serum free light chain (sFLC) analysis

Evidence from of 4 studies including 4888 patients (McTaggart et al 2013, Hill et al 2006, Piehler et al 2008 and Vermeersch et al 2008) suggests serum free light chain ratio outside the normal range has sensitivity of 47% [33% – 60%] and specificity of 95% [85% – 99%] for the diagnosis of plasma cell disorders.

Combined SPE and sFLC

Evidence from 3 studies including 4054 patients (McTaggart et al 2013, Hill et al 2006, Piehler et al 2008) suggests that combining serum free light chain analysis with serum protein electrophoresis, improves sensitivity for the diagnosis of plasma cell disorders with a pooled estimate of 94% [72% – 99%]. In this strategy patients with a negative serum protein electrophoresis test would go on to have a serum free light chain test.

Other tests for plasma cell disorders

Three studies were identified which aimed to determine the most clinically effective diagnostic testing strategy for plasma cell disorders. In one UK study, 2,799 patients with suspected plasma cell dyscrasias were tested with serum protein electrophoresis with either urine protein electrophoresis (UPE) or serum free light chain analysis (McTaggart et al., 2013). The combination of sFLC and SPE had the greatest sensitivity (100% (95% CI 97 to 100), detecting all 124 patients with plasma cell disorders, and had specificity of 97% (95% CI 96 to 97). This was greater than the diagnostic accuracy of SPE and UPE, which had a sensitivity of 96% (95% CI 89 to 99) and a specificity of 95% (95% CI 93 to 97), although only this was based on fewer patients (n=579) and there is overlap in the confidence intervals for sensitivity and specificity of the two testing strategies.

One study reported the diagnostic accuracy of different testing strategies in 833 patients investigated for monoclonal gammopathy. SPE with follow-up immunofixation electrophoresis (IFE) plus sFLC had a sensitivity of 82% and a specificity of 97%. Serum IFE plus urine IFE had a sensitivity of 92% and a specificity of 100%. Neither of these testing strategies missed a case of myeloma (Vermeersch et al., 2008).

A further study only included patients with an existing plasma cell disorder (including 467 myeloma, 191 smouldering myeloma, 524 MGUS, 581 primary amyloidosis) (Katzmann et al., 2009). The combinations of SPE/IFE/sFLC and SPE/sFLC both detected 100% of the 467 patients with multiple myeloma.

Behdad et al (2014) reported that multiparameter flow cytometry had sensitivity 94% and specificity 68% for the diagnosis of plasma cell neoplasm versus not in a study of 361 patients with suspected plasma cell neoplasm.

Diagnostic accuracy of tests for the discrimination of myeloma versus MGUS

Serum protein electrophoresis – monoclonal protein

M-protein in serum ≥ 30 g/l is one of the International Myeloma Working Group (2003) consensus diagnostic criteria – so by definition it has 100% specificity for the diagnosis of myeloma versus MGUS in studies using those criteria. Some patients with myeloma have lower M-protein levels so this criterion alone has imperfect sensitivity for myeloma. Frebert et al (2011) in a study of 161 patients with myeloma or MGUS estimated the sensitivity for myeloma of this 30 g/L cutoff as only 41%.

In a study of 67 patients with monoclonal gammopathy, Wolff et al (2007) reported that the presence of a monoclonal band on serum protein electrophoresis had a sensitivity of 85% for intact immunoglobulin myeloma but only 40% for light chain myeloma.

Bone marrow plasma cell percentage

Similarly a clonal bone marrow plasma cell percentage ≥ 10% is one of the International Myeloma Working Group (2003) diagnostic criteria – so by definition it has 100% specificity for the diagnosis of myeloma versus MGUS in studies using those criteria. Some patients with myeloma have lower clonal bone marrow plasma cell percentages so this criterion alone has imperfect sensitivity for myeloma. In two studies including 229 patients with myeloma or MGUS (Milla et al 2001, Frebert et al 2011) with myeloma or MGUS, a ≥10% threshold had a sensitivity of 79% and a ≥30% threshold a sensitivity of 58% for myeloma.

Goyal et al (2014) reported that bone marrow aspirate was less sensitive than bone marrow trephine biopsy for myeloma, 74% versus 84% respectively, in a series of 31 patients with myeloma. In 5/31 patients however neither bone marrow aspirate or trephine biopsy showed plasmacytosis.

Cytomorphology

Evidence from one study (Milla et al 2001) including 68 patients with MGUS or myeloma suggests that a cytomorphologist's diagnosis has a sensitivity of 100% for myeloma with a specificity of 88%. In this study the use of a formal cytomorphologic atypia scoring system reduced the sensitivity for myeloma to 83%.

Serum free light chain analysis

Evidence about the use of serum free light chains for discrimination of myeloma from MGUS came from two studies (Wolff et al 2007 and Bergon et al 2005) including 484 patients. In Wolf et al (2007) free light chain quantification had a sensitivity of 76% and specificity of 75% for the discrimination of myeloma from MGUS when using a normal range for κ/λ ratio of 0.19 – 1.48. FLC testing had a sensitivity of 100% in the subgroup of five patients with light chain multiple myeloma.

Bergon et al (2005) explored the use of different thresholds for lower and higher bounds of the normal κ/λ ratio. Expanding the normal range for κ/λ ratio has the effect of increasing specificity but lowering sensitivity for the diagnosis of myeloma versus MGUS.

Flow cytometry

Two studies (Carulli et al, 2012 and Frebert et al, 2011), including 297 patients, evaluated multiparameter flow cytometry (MFC) for the discrimination of myeloma from MGUS. MFC measurement of the ratio of immunophenotypically abnormal to normal plasma cells had sensitivity of 74% to 98% and specificity 85% to 92% for myeloma.

Bacher et al (2010) compared the proportion of plasma cells identified using bone marrow cytomorphology with those found using MFC in 682 patients. This proportion was higher with bone marrow cytomorphology than with MFC: the median proportion of plasma cells was 8.5% versus 2% for cytomorphology and MFC respectively. However in 1.3% of cases MFC was able to detect monoclonal plasma cells when cytomorphology did not.

Cytogenetic abnormalities on FISH

Evidence from about cytogenetic abnormalities came from one study (Bacher et al, 2010) including 682 patients with myeloma or MGUS. Although cytogenetic abnormalities were more likely in myeloma than MGUS (87% versus 56% respectively, P<0.001) there was no cytogenetic abnormality unique to either diagnosis. FISH testing was more likely to be successful in patients with myeloma than in those with MGUS (90% versus 79% respectively) – test failures were related to insufficient amounts of plasma cells.

Diagnostic accuracy of tests for detection of myeloma in patients with renal failure

In one study of 82 patients with acute renal failure, seven were diagnosed with multiple myeloma using SPE, IFE and bone marrow biopsy. The FLC κ/λ ratio based on FLC measurement (using the published range of 0.26-1.65) had a sensitivity of 71% (95% CI 0.29 to 0.96) and a specificity of 96% (95% CI 89 to 99) for the diagnosis of multiple myeloma, with 3 false positives and 2 false negatives (Cirit et al., 2012). Another study of 471 patients with renal insufficiency reported that renal range FLC showed the highest sensitivity (92%) to differentiate multiple myeloma from non-multiple myeloma among four tests (conventional range FLC, SPE, UPE). Combined analysis with FLC and SPE improved the diagnostic accuracy to 98% sensitivity (Park et al., 2012). In a UK study, 142 patients with dialysis-dependant renal failure were assessed with SPE, IFE, and FLC (Hutchison et al., 2008). 41 patients had a clinical diagnosis of multiple myeloma, all of whom had abnormal serum FLC ratios. The modified renal reference FLC range (0.37-3.1) increased specificity from 93% to 99%, with no loss of sensitivity.

Cost-effectiveness evidence

A literature review of published cost-effectiveness analyses did not identify any relevant papers for this topic. Whilst there were potential cost implications of making recommendations in this area, other questions in the guideline were agreed as higher priorities for economic evaluation. Consequently no further economic modelling was undertaken for this question.

Recommendations	**Use serum protein electrophoresis and serum-free light-chain assay to confirm the presence of a paraprotein indicating possible myeloma or monoclonal gammopathy of undetermined significance (MGUS).** **Use serum immunofixation if serum protein electrophoresis is abnormal to confirm the presence of a paraprotein indicating possible myeloma or MGUS.** **Do not use serum protein electrophoresis, serum immunofixation, serum-free light-chain assay or urine electrophoresis (urine Bence–Jones protein assessment) alone to exclude a diagnosis of myeloma.** **When performing a bone marrow aspirate and trephine biopsy to confirm a diagnosis of myeloma, use morphology to determine plasma cell percentage and flow cytometry to determine plasma cell phenotype.**
Relative value placed on the outcomes considered	The Guideline Committee considered the outcomes of diagnostic accuracy, rate of confirmed diagnosis, delay in diagnosis, test-related adverse events and patient awareness of diagnosis to be the most relevant to identify the optimal laboratory testing strategy for suspected myeloma. No evidence was identified for delay in diagnosis, test-related adverse events and patient awareness of diagnosis. Diagnostic accuracy was reported in terms of both sensitivity and specificity but when making the recommendation the Guideline Committee prioritised sensitivity as they felt that any false positives would be dealt with by additional strategies, for example further blood tests.
Quality of the evidence	The evidence was assessed by QUADAS-2 as high quality. The Guideline Committee noted that different reference standards had been used both between and within studies. Also three studies had unclear applicability concerns due to patient selection because they included only patients with renal failure. Also, 1 study included a potential biased sample of patients who had the urine test. The Guideline Committee agreed to assume that the patient sample was representative of the population as a whole.
Trade off between clinical benefits and harms	The Guideline Committee noted, based on the evidence, that whilst urine testing for Bence-Jones protein is almost as effective in diagnosing plasma cell disorders as testing for serum free light chains, the evidence also showed that urine testing was only done in a fraction of the people being tested. This could have resulted in potential missed diagnoses if the serum free light chain test was not performed as an alternative. Hence the Guideline Committee did not include urine testing in their recommendations.

	The Guideline Committee agreed to recommended serum immunofixation if there was abnormal serum protein electrophoresis. This was based on the testing strategy in Vermeech et al., 2008 (only doing immunofixation if the electrophoresis was abnormal). The Guideline Committee recommended that each of the available tests should not be used alone. This was based on the evidence that individual testing had a low sensitivity compared to combinations of tests and as such using just one individual test could potentially miss myeloma patients. Evidence was presented to the Guideline Committee for the diagnostic accuracy of tests for detection of myeloma in patients with renal failure. However the Guideline Committee were unable to make any specific recommendation for these patients as they would need to include reference ranges, which vary depending on the laboratory processing the test. However, the Guideline Committee agreed that the existing recommendations were appropriate for patients with renal failure so the lack of a specific recommendation for these patients was not an issue. The Guideline Committee concluded that optimised laboratory investigations for suspected myeloma would result in a number of benefits including potential earlier diagnosis leading to a reduction in complications and uniformity of access to serum free light chain assay. The Guideline Committee also recognised that there may be an increase in referral of people who do not have myeloma (and associated anxiety) resulting from a false positive test. The Guideline Committee agreed that the benefits of diagnosing those people who actually have myeloma outweighed the potential harms.
Trade-off between net health benefits and resource use	The Guideline Committee noted that no relevant published economic evaluations had been identified and no additional economic analysis had been undertaken in this area. As a result of the recommendations made, the Guideline Committee agreed that there would be an increase in costs associated with the increased use of serum free light chain, but a decrease in costs as a result of the reduction in urine tests. The Guideline Committee also thought that as a consequence of earlier diagnosis there would be reduced costs of dealing with the complications of the disease (for example dialysis, bone disease etc.). Taking all this into account it was agreed that the net effect of the recommendations was likely to be neutral – with no significant increase or saving in costs.
Other	The Guideline Committee discussed current practice and that whilst there is a pathway already set up to look at serum free light chain assay, the volume of use is currently variable. They agreed that the recommendations were likely to results in an increase in the use of this pathway.

2.2 Laboratory investigations to provide prognostic information

Multiple myeloma is a heterogeneous disease with a wide range of clinical outcomes. Advances in treatments over the last decade have improved median overall survival in younger people with myeloma to around 7 to 10 years from diagnosis. However there remains a group of patients with significantly worse outlook, loosely defined as having high-risk myeloma. A wide range of techniques performed on the diagnostic bone marrow sample have been used to provide prognostic information in both newly diagnosed and relapsed myeloma cases. These can be broadly separated in to immunological techniques (immunophentotyping and immunohistochemistry) and genetic techniques (including cytogenetics, fluorescent in situ hybridisation, polymerase chain reaction techniques, sequencing and microarray technologies). Whilst a large range of molecular techniques have been developed in a research capacity, there remains minimal incorporation in to clinical practice at present.

Immunological techniques are used to both identify prognostic variables and to monitor response to treatment while most genetic techniques are concerned with prognostic information.

New drugs have been developed which appear to have the ability to treat myeloma with adverse genetic prognostic factors so evaluating these factors may become particularly important.

This question will review the current utility of tests available in specialist myeloma practice for prognostic information and management of patients.

Clinical question: Can investigations done at the diagnosis of myeloma, including trephine biopsy, immunophenotyping and cytogenetic and molecular genetic tests accurately predict treatment outcomes (for example, can they identify patients with a poor prognosis for whom an alternative treatment approach may be preferable)?

Clinical evidence (see also Appendix G)

Study quality

The included studies are high quality studies with a low risk of bias, although some studies do not include a multivariate model in the analysis to determine whether the assessed prognostic risk factor is independent of other risk factors. Treatment heterogeneity is an issue between as well as within studies.

Immunohistochemistry

Five studies were identified that investigated the prognostic value of immunohistochemistry. Each of the 5 studies investigated different markers. P53 expression and ki-67 antigen expression were found to be independent risk factors for OS (Chang et al., 2007 and Gastinee et al., 2007), whilst CD56, CD99 and cyclin D1 expression were not associated with patient survival (Chang et al., 2006; Shin et al., 2014; Tinguely et al., 2007).

Flow cytometry

Fourteen studies were identified that investigated the prognostic value of flow cytometry. All 14 studies found flow cytometry was able to identify myeloma patients with a poor prognosis. However not all studies could confirm their results in a multivariate model.

The identified studies all used flow cytometry to investigate a number of different markers. Five studies assessed the prognostic value of clonal circulating plasma cells and all 5 studies concluded that clonal circulating plasma cells were an independent risk factor for patient survival (Gonsalves et al., 2014; Nowakowski et al., 2005; Paiva et al., 2009a; 2009b; 2013).

CD antigens were investigated by flow cytometry in a number of studies. CD28+ (Mateo et al., 2008), CD81+ (Paiva et al., 2012a) and CD19+/CD117- (Caltagirone et al., 2014) were all found to be independent prognostic risk factors for survival in myeloma patients, whereas CD19 (Caltagirone et al., 2014; Mateo et al., 2008), CD45 (Caltagirone et al., 2014; Mateo et al., 2008), CD20 (Caltagirone et al., 2014; Mateo et al., 2008), CD56 (Caltagirone et al., 2014; Mateo et al., 2008) and CD33 (Mateo et al., 2008) were all reported to not be associated with clinical outcomes. CD117 was found to be prognostic in one study (Mateo et al., 2008) but not in another (Caltagirone et al., 2014).

DNA content/ hyperdiploidy was assessed in 3 studies. All 3 studies found that hyperdiploid patients had increased survival compared to non-hyperdiploid patients. But whether DNA content is an independent risk factor remains uncertain. One study reported that DNA content remained significant in a multivariate model (Paiva et al., 2012b), but another study reported that it lost significance (Mateos et al. 2011) whilst a third study did not include a multivariate model (Chng et al., 2006).

A high plasma cell proliferation index was reported to be associated with worse survival compared to a lower plasma cell proliferation index in 4 studies. The association remained significant after taking into account other risk factors in a multivariate model in one study (Paiva et al., 2012b). A multivariate model was not included in the other 3 studies (Minarik et al., 2005; 2010; 2011). The poor prognosis associated with a high proliferative index may be overcome by the use of novel agents (Minarik et al., 2010; Paiva et al., 2012b).

A low plasma cell apoptosis index was reported to be associated with worse survival compared to a higher plasma cell apoptosis index in 2 studies (Minarik et al., 2005; 2011). These studies did not include a multivariate model so it is uncertain whether the apoptosis index is an independent prognostic factor for patient survival in myeloma.

Serum free light chains

Eight studies were identified that investigated the prognostic value of serum free light chains (FLC). All 8 studies found serum FLC to be prognostic. Two studies reported that abnormal FLC was independently prognostic for a higher risk of progression from smouldering myeloma to active myeloma (Dispenzieri et al., 2008a; Larsen et al., 2013) and three studies reported that abnormal FLC was independently prognostic for myeloma patient survival (Kumar et al., 2010; Snozek et al., 2008; Van Rhee et al., 2007; Xu et al., 2013). Two further studies also reported serum FLC to be predictive for patient survival in myeloma, however multivariate analysis was not done and so it is unclear whether serum free chains were an independent prognostic factor in these studies (Dispenzieri et al., 2008b; Maltezas et al., 2013).

Heavy/light chain ratio

Three studies were identified that investigated the prognostic value of heavy/light chain ratio (Bradwell et al., 2013; Koulieris et al., 2012, Ludwig et al., 2013). All 3 studies found the heavy/light chain ratio to be independently prognostic for either OS or PFS.

FISH

Thirty four studies were identified that investigated the prognostic value of FISH. Thirty one studies examined genetic abnormalities in newly diagnosed myeloma patients and determined the prognostic impact of these genetic abnormalities on patient survival (PFS

and/or OS) and three studies examined genetic abnormalities in smouldering myeloma patients and determined the prognostic impact of these genetic abnormalities on time to progression to active myeloma.

The most common genetic abnormalities assessed were: t(11;14), t(4;14), t(14;16), del(17p), del(13q), del(1p), 1q gains, del(p53) and hyperdiploidy.

Patients with newly diagnosed myeloma):

t(11:14) was included in 13 studies (An et al., 2013, Avet-Loiseau et al., 2007, Avet-Loiseau et al., 2012, Avet-Loiseau et al., 2013a, Bang et al., 2006, Boyd et al., 2012, Caltagitone et al., 2014, Chang et al., 2005a, Chang et al., 2010, Gutierrez et al., 2007, Neben et al., 2010, Nemec et al., 2012 and Walker et al., 2010) but only 1 study found an association with patient survival. This association did not remain significant in the multivariate model.

t(4:14) was included in 16 studies (Avet-Loiseau et al., 2007, Avet-Loiseau et al., 2010, Avet-Loiseau et al., 2011, Avet-Loiseau et al., 2012, Avet-Loiseau et al., 2013a, Avet-Loiseau et al., 2013b, Boyd et al., 2012, Caltagitone et al., 2014, Chang et al., 2005a, Chang et al., 2010, Grzasko et al., 2013, Gutierrez et al., 2007, Moreau et al., 2007, Neben et al., 2010, Nemec et al., 2012 and Walker et al., 2010) and 12 of these reported an association between the genetic abnormality and patient survival. 9 of the 12 studies reported t(4;14) to be an independent prognostic factor after multivariate analysis whilst no multivariate analysis was undertaken in the other 3 studies. t(14:16) was included in 8 studies (Avet-Loiseau et al., 2011, Avet-Loiseau et al., 2012, Avet-Loiseau et al., 2013a, Boyd et al., 2012, Caltagitone et al., 2014, Gutierrez et al., 2007, Neben et al., 2010 and Walker et al., 2010) only 1 of which reported this genetic abnormality to be prognostic for patient survival.

Del(17p) was included in 12 studies (Avet-Loiseau et al., 2007, Avet-Loiseau et al., 2010, Avet-Loiseau et al., 2011, Avet-Loiseau et al., 2012, Avet-Loiseau et al., 2013a, Avet-Loiseau et al., 2013b, Boyd et al., 2012, Caltagitone et al., 2014, Grzasko et al., 2013, Neben et al., 2010, Nemec et al., 2012 and Walker et al., 2010) and 10 of these reported an association between the genetic abnormality and patient survival. 7 of the 10 studies reported del(17p) to be an independent prognostic factor after multivariate analysis whilst no multivariate analysis was undertaken in the other 3 studies. Del(p53) was included in 3 studies (Avet-Loiseau et al., 2007, Boyd et al., 2012 and Walker et al., 2010) but only 1 study found an association with patient survival. This association did not remain significant in the multivariate model.

Del(13q) was included in 14 studies (Avet-Loiseau et al., 2007, Avet-Loiseau et al., 2011, Avet-Loiseau et al., 2012, Avet-Loiseau et al., 2013a, Avet-Loiseau et al., 2013b, Bang et al., 2006, Boyd et al., 2012, Caltagitone et al., 2014, Chang et al., 2005a, Chang et al., 2010, Grzasko et al., 2013, Lai et al., 2012, Neben et al., 2010 and Nemec et al., 2012) and 9 of these reported an association between the genetic abnormality and patient survival. 4 of the 9 studies reported del(13q) to be an independent prognostic factor after multivariate analysis and 4 reported del(13q) to not be an independent prognostic factor whilst no multivariate analysis was undertaken in 1 study.

Del(1p) was included in 6 studies (Boyd et al., 2012, Caltagitone et al., 2014, Chang et al., 2010, Chng et al., 2010, Hebraud et al., 2014 and Walker et al., 2010) and 5 of these reported an association between the genetic abnormality and patient survival. 3 of the 5 studies reported del(1p) to be an independent prognostic factor after multivariate analysis whilst no multivariate analysis was undertaken in the other 2 studies.

Amp(1q) was included in 13 studies (An et al., 2014, Avet-Loiseau et al., 2012, Bang et al., 2006, Boyd et al., 2012, Caltagitone et al., 2014, Chang et al., 2010, Fonseca et al., 2006, Grzasko et al., 2013, Hanamura et al., 2006, Lai et al., 2012, Neben et al., 2010, Nemec et al., 2012 and Walker et al., 2010) and 9 of these reported an association between the genetic abnormality and patient survival. 5 of the 9 studies reported amp(1q) to be an

independent prognostic factor after multivariate analysis and 2 reported amp(1q) to not be an independent prognostic factor whilst no multivariate analysis was undertaken in 2 studies.

Hyperdiploidy was included in 5 studies (Chang et al., 2005a, Chang et al., 2005b, Chang et al., 2010, Gutierrez et al., 2007 and Lai et al., 2012) and 3 of these found an association with patient survival all of which remained significant in the multivariate model.

Patients with smouldering myeloma

t(11:14) was included in 3 studies (Lopez-Coral et al., 2012, Neben et al., 2013 and Rajkumar et al., 2013) but none of these found t(11;14) to be prognostic for progression to myeloma.

t(4:14) was included in 3 studies (Lopez-Coral et al., 2012, Neben et al., 2013 and Rajkumar et al., 2013) and 2 of these reported an association between the genetic abnormality and TTP. 1 study reported t(4;14) to be an independent prognostic factor after multivariate analysis whilst in the other study the result lost significance after multivariate analysis.

t(14:16) was included in 1 study (Lopez-Coral et al., 2012) but it was not found to be prognostic for progression to myeloma.

Del(17p) was included in 2 studies (Lopez-Coral et al., 2012 and Neben et al., 2013). One study reported an association between the genetic abnormality and TTP but the result lost significance after multivariate analysis.

Del(13q) was included in 3 studies (Lopez-Coral et al., 2012, Neben et al., 2013 and Rajkumar et al., 2013) but none of these found del(13q) to be prognostic for progression to myeloma.

Amp(1q) was included in 2 studies (Lopez-Coral et al., 2012 and Neben et al., 2013). One study reported an association between the genetic abnormality and TTP but the result lost significance after multivariate analysis.

Hyperdiploidy was included in 2 studies (Lopez-Coral et al., 2012 and Neben et al., 2013). One study reported an association between the genetic abnormality and TTP but the result lost significance after multivariate analysis.

No studies investigated the prognostic importance of del(1p) or del(p53) in smouldering myeloma.

A number of studies divided patients into high, standard or low risk groups based on the genetic abnormalities they carried (or lacked). It is difficult to compare across studies as different studies used different genetic abnormalities. However all studies reported that myeloma patients classed as high risk (with adverse genetic abnormalities) had a worse prognosis for survival compared to patients that were in the low risk group (without the established adverse genetic abnormalities) (Boyd et al., 2012; Chang et al., 2005a; Jacobus et al., 2011; Kapoor et al., 2010; Kumar et al., 2012; Lu et al., 2014; Mateos et al., 2011; Paiva et al., 2012c). Similarly, smouldering myeloma patients defined as high risk had a worse prognosis for progression to active myeloma (Neben et al., 2013; Rajkumar et al., 2013).

Cost-effectiveness evidence

A literature review of published cost-effectiveness analyses did not identify any relevant papers for this topic. Whilst there were potential cost implications of making recommendations in this area, other questions in the guideline were agreed as higher priorities for economic evaluation. Consequently no further economic modelling was undertaken for this question.

Recommendations	**Use the same sample for all diagnostic and prognostic tests on bone marrow, so people only have to have one bone marrow aspirate and trephine biopsy.** **When performing a bone marrow aspirate and trephine biopsy to provide prognostic information:** • **Perform fluorescence in-situ hybridisation (FISH) on CD138-selected bone marrow plasma cells to identify the adverse risk abnormalities t(4;14), t(14;16), 1q gain, del(1p) and del(17p)(TP53 deletion). Use these abnormalities alongside International Staging System (ISS) scores to identify people with high-risk myeloma.** • **Consider performing FISH on CD138-selected bone marrow plasma cells to identify the adverse risk abnormality t(14;20), and the standard risk abnormalities t(11;14) and hyperdiploidy.** • **Consider performing immunophenotyping of bone marrow to identify plasma cell phenotype, and to inform subsequent monitoring.** • **Consider performing immunohistochemistry (including Ki-67 staining and p53 expression) on the trephine biopsy to identify plasma cell phenotype, cell proliferation and p53 deletion, to provide further prognostic information.** **Perform serum-free light-chain assay and use serum-free light-chain ratio to assess prognosis.**
Relative value placed on the outcomes considered	The Guideline Committee considered the outcomes of response to treatment, adverse events, overall survival, progression-free survival and time to next treatment (for asymptomatic patients) to be the most relevant for identifying which investigations should be done at the diagnosis of myeloma to accurately predict treatment outcomes. No evidence was identified on adverse events. When drafting the recommendations the Guideline Committee considered overall survival and progression-free survival to be the most important.
Quality of the evidence	The evidence was assessed by QUADAS-2 as high quality with a low risk of bias. It was noted that some studies did not include a multivariate model in the analysis to determine whether the assessed prognostic risk factor was independent of other risk factors. The Guideline Committee therefore applied more weight to those studies that did include a multivariate model.
Trade off between clinical benefits and harms	The Guideline Committee agreed to recommend FISH testing to identify those adverse risk abnormalities which had been shown in multivariate analyses to be independent prognostic marker of high-risk disease. FISH testing has been validated in a large number of clinical trials and scientific studies as being the most practical and broadly applicable technique to identify acquired genetic abnormalities in myeloma. Patients with high-risk disease have worse outcomes with conventional treatments so identifying this group of patients enables other treatments to be considered. It also provides prognostic information for patients with myeloma, their family and carers. The Guideline Committee noted based on the evidence, that combinations of adverse risk abnormalities alongside the International Staging System score could be used to identify high-risk disease. Combinations were shown to be more effective than in isolation. However the evidence reported on many different combinations and not all of the results were consistent, therefore it was not possible for the Guideline Committee to specify a particular combination to define

	high risk myeloma. The Guideline Committee noted that the evidence had reported t(14;20) as a prognostic marker for high-risk disease but that there was less volume of evidence for this marker than for others. The Guideline Committee considered, based on their clinical experience and the available evidence, that t(14;20) was a prognostic marker of high-risk disease and therefore recommended the use of FISH to identify this marker should be considered. The Guideline Committee noted that the evidence had shown the standard risk abnormalities t(11;14) and hyperdiploidy were markers of not having high-risk disease. Because they indicate standard as opposed to high-risk disease, the group made a recommendation to consider the use of FISH to identify these markers as knowing this information can be helpful in discussing prognosis with patients. The Guideline Committee agreed that the evidence had shown plasma cell phenotype by flow cytometry, P53 expression and proliferation by Ki-67 staining by immunohistochemistry were prognostic markers for high-risk disease. They therefore recommended immunophenotyping of bone marrow and immunohistochemistry on the trephine biopsy to identify these markers. The Guideline Committee also considered, based on their clinical experience, that the initial plasma cell phenotype could be used to inform subsequent monitoring. Based on the evidence, the Guideline Committee also noted that serum-free light chain assay and serum-free light chain ratio were independent prognostic markers for high-risk disease. However the evidence reported on many different ratios and not all of the results were consistent, therefore it was not possible for the Guideline Committee to specify a particular ratio to use. The Guideline Committee noted that the evidence had reported that the Hevylite® assay results can be used to identify high-risk disease. However they considered that this was a relatively new technology and that it had not been independently validated in many studies. Consequently they recommended further validated research to determine the prognostic significance of results using this technique. The Guideline Committee agreed that provision of prognostic information would result in several benefits particularly the standardisation of the identification of high-risk disease (currently there is variation in whether or not this is assessed and what panels of tests are used to do this). This would enable better informed discussion with patients and potentially lead to improvements in treatment. Another benefit would be the avoidance of inappropriate treatments. The Guideline Committee noted that the potential harms could be the psychological effect on patients of being identified as having high-risk disease. However they balanced this potential harm against the ability to give different, more appropriate treatment and better support to the patient.
Trade-off between net health benefits and resource use	The Guideline Committee noted that no relevant published economic evaluations had been identified and no additional economic analysis had been undertaken in this area. As a result of the recommendations made, the Guideline Committee

	agreed there would be additional costs for undertaking the tests (as these are not all currently being done as standard) and time taken to give this information to patients. There would also likely be changes to the costs of providing treatment – some savings from not prescribing inappropriate treatments and some additional costs from prescribing different treatments (based on identified risks) although the direction of change in treatment costs was uncertain.
Other	The Guideline Committee noted that tests undertaken using the bone marrow aspirate and trephine biopsy samples can be used to both diagnose myeloma and to provide prognostic information. Tests such as immunophenotyping and CD138 selection for subsequent FISH analysis need to be undertaken on fresh bone marrow potentially before the diagnosis of myeloma has been confirmed. However the Guideline Committee agreed, based on their clinical experience, that a bone marrow aspirate and trephine biopsy is a potentially painful procedure and it was preferable to request diagnostic and prognostic tests simultaneously to avoid duplicate procedures and a negative patient experience as well as to minimise resource wastage.

Research recommendation	**A prospective randomised multi-centre trial of different treatment strategies should compare the prognostic value of the Hevylite® assay and ratio with other prognostic factors and tests, including the serum-free light-chain assay and fluorescence in situ hybridisation (FISH), in people with newly diagnosed myeloma who are starting treatment. Outcomes of interest are overall response, complete response, minimal residual disease, progression-free survival, overall survival and resource use.**
Why this is important	Hevylite® is a new assay which some studies have indicated is a useful prognostic tool. However, it is not clear how robustly it has been evaluated against other prognostic factors and tests, or whether it is an independent prognostic factor. The Hevylite® assay should be evaluated in an accredited centralised laboratory independent of links with the manufacturer.

References

An G, Xu Y, Shi L et al. (2013) t(11;14) multiple myeloma: a subtype associated with distinct immunological features, immunophenotypic characteristics but divergent outcome. Leukemia Research, 37: 1251-1257.

An G, Xu Y, Shi LH et al. (2014) Chromosome 1q21 gains confer inferior outcomes in multiple myeloma treated with bortezomib but copy number variation and percentage of plasma cells involved have no additional prognostic value. Haematologica, 99: 353-359.

Avet-Loiseau H, Attal M, Moreau P et al. (2007) Genetic abnormalities and survival in multiple myeloma: the experience of the Intergroupe Francophone du Myelome. Blood, 109: 3489-3495.

Avet-Loiseau H (2010) Bortezomib plus dexamethasone induction improves outcome of patients with t(4;14) myeloma but not outcome of patients with del(17p). Journal of Clinical Oncology, 28: 4630-4634.

Avet-Loiseau H, Malard F, Campion L, et al. (2011) Translocation t(14;16) and multiple myeloma: is it really an independent prognostic factor? Blood, 117: 2009-2011

Avet-Loiseau H, Attal M, Campion L et al. (2012) Long-term analysis of the IFM 99 trials for myeloma: cytogenetic abnormalities [t(4;14), del(17p), 1q gains] play a major role in defining long-term survival. Journal of Clinical Oncology, 30: 1949-1952.

Avet-Loiseau H, Durie BG, Cavo M, et al. (2013a) Combining fluorescent in situ hybridization data with ISS staging improves risk assessment in myeloma: an International Myeloma Working Group collaborative project. Leukemia, 27: 711-717.

Avet-Loiseau H, Hulin C, Campion L et al. (2013) Chromosomal Abnormalities Are Major Prognostic Factors in Elderly Patients With Multiple Myeloma: The Intergroupe Francophone du Myelome Experience. Journal of Clinical Oncology, 31: 2806-3809.

Bacher U et al. (2010) Correlation of cytomorphology, immunophenotyping, and interphase fluorescence in situ hybridization in 381 patients with monoclonal gammopathy of undetermined significance and 301 patients with plasma cell myeloma. Cancer Genetics and Cytogenetics 2010; 203: 169-175.

Bang SM (2006) Identification of 13q deletion, trisomy 1q, and IgH rearrangement as the most frequent chromosomal changes found in Korean patients with multiple myeloma. Cancer Genetics and Cytogenetics, 168: 124-132.

Behdad A (2014). Utility of nine-color, 11-parameter flow cytometry for detection of plasma cell neoplasms: a comparison with bone marrow morphologic findings and concurrent M-protein studies in serum and urine. American Journal of Clinical Pathology, 142, 398-410.

Bergon E, Miravalles E, Bergón E,et al. (2005) The predictive power of serum kappa/lambda ratios for discrimination between monoclonal gammopathy of undetermined significance and multiple myeloma. Clin Chem Lab Med; 43: 32-37.

Boyd KD, Ross FM, Chiecchio L et al. (2012) A novel prognostic model in myeloma based on co-segregating adverse FISH lesions and the ISS: analysis of patients treated in the MRC Myeloma IX trial. Leukemia, 26: 349-355.

Bradwell A, Harding S, Fourrier N, et al. (2013) Prognostic utility of intact immunoglobulin Ig'/Ig' ratios in multiple myeloma patients. Leukemia, 27: 202-207.

Caltagirone SR (2014) Chromosome 1 abnormalities in elderly patients with newly diagnosed multiple myeloma treated with novel therapies. Haematologica, 99: 1611-1617.

Carulli G Ottaviano V, Cannizzo E, et al. (2012) Multiparameter immunophenotyping by flow cytometry as a diagnostic tool in multiple myeloma and monoclonal gammopathy of undetermined significance. Clinica Terapeutica; 163: 387-392.

Chang H, Qi XY, Samiee S, et al. (2005a) Genetic risk identifies multiple myeloma patients who do not benefit from autologous stem cell transplantation. Bone Marrow Transplantation, 36: 793-796

Chang H, Qi C, Yi QL, et al. (2005b) p53 gene deletion detected by fluorescence in situ hybridization is an adverse prognostic factor for patients with multiple myeloma following autologous stem cell transplantation. Blood, 105: 358-360.

Chang H, Samiee S & Yi QL (2006) Prognostic relevance of CD56 expression in multiple myeloma: a study including 107 cases treated with high-dose melphalan-based chemotherapy and autologous stem cell transplant. Leukemia & lymphoma, 47: 43-47.

Chang H, Yeung J, Qi C, et al. (2007) Aberrant nuclear p53 protein expression detected by immunohistochemistry is associated with hemizygous P53 deletion and poor survival for multiple myeloma. British Journal of Haematology, 138: 324-329.

Chang H, Qi X, Jiang A, et al. (2010) 1p21 deletions are strongly associated with 1q21 gains and are an independent adverse prognostic factor for the outcome of high-dose chemotherapy in patients with multiple myeloma. Bone Marrow Transplantation, 45: 117-121.

Chng WJ (2006) Prognostic factors for hyperdiploid-myeloma: Effects of chromosome 13 deletions and IgH translocations. Leukemia, 20: 807-813.

Chng WJ, Gertz MA, Chung TH, et al. (2010) Correlation between array-comparative genomic hybridization-defined genomic gains and losses and survival: identification of 1p31-32 deletion as a prognostic factor in myeloma. Leukemia, 24: 833-842.

Cirit M Uzüm A, Ozen P, et al. (2012) The value of serum immunoglobulin free light chain assessment in patients with monoclonal gammopathies and acute renal failure. Turkish Journal of Haematology 2012; 29: 385-391.

Dispenzier, A, Kyle RA, Katzmann JA, et al. (2008a) Immunoglobulin free light chain ratio is an independent risk factor for progression of smoldering (asymptomatic) multiple myeloma. Blood, 111: 785-789.

Dispenzieri A, Zhang L, Katzmann JA, et al. (2008b) Appraisal of immunoglobulin free light chain as a marker of response. Blood, 111; 4908-4915.

Fonseca R, Van Wier SA, Chng WJ, et al. (2006) Prognostic value of chromosome 1q21 gain by fluorescent in situ hybridization and increase CKS1B expression in myeloma. Leukemia, 20: 2034-2040.

Frebet E Abraham J, Geneviève F, et al. (2011) A GEIL flow cytometry consensus proposal for quantification of plasma cells: application to differential diagnosis between MGUS and myeloma. Cytometry B Clin Cytom; 80: 176-185.

Gastinne T (2007) Plasma cell growth fraction using Ki-67 antigen expression identifies a subgroup of multiple myeloma patients displaying short survival within the ISS stage I. European Journal of Haematology, 79: 297-304.

Gonsalves WI, Rajkumar SV, Gupta V, et al. (2014) Quantification of clonal circulating plasma cells in newly diagnosed multiple myeloma: implications for redefining high-risk myeloma. Leukemia, 28: 2060-2065.

Goyal S, Singh UR, & Rusia U (2014) Comparative evaluation of bone marrow aspirate with trephine biopsy in hematological disorders and determination of optimum trephine length in lymphoma infiltration. Mediterranean Journal of Hematology & Infectious Diseases, 6, e2014002.

Grzasko N, Hus M, Pluta A, et al. (2013) Additional genetic abnormalities significantly worsen poor prognosis associated with 1q21 amplification in multiple myeloma patients. Hematological Oncology, 31: 41-48.

Gutierrez NC, Castellanos MV, Martin ML, et al. (2007) Prognostic and biological implications of genetic abnormalities in multiple myeloma undergoing autologous stem cell transplantation: t(4;14) is the most relevant adverse prognostic factor, whereas RB deletion as a unique abnormality is not associated with adverse prognosis. Leukemia, 21: 143-150.

Hanamura I, Stewart JP, Huang Y, et al. (2006) Frequent gain of chromosome band 1q21 in plasma-cell dyscrasias detected by fluorescence in situ hybridization: incidence increases from MGUS to relapsed myeloma and is related to prognosis and disease progression following tandem stem-cell transplantation. Blood, 108: 1724-1732.

Hebraud B, Leleu X, Lauwers-Cances V, et al. (2014) Deletion of the 1p32 region is a major independent prognostic factor in young patients with myeloma: the IFM experience on 1195 patients. Leukemia, 28: 675-679.

Hill PG, Forsyth JM, Rai B, et al. (2006) Serum free light chains: An alternative to the urine Bence Jones proteins screening test for monoclonal gammopathies. Clinical Chemistry; 52: 1743-1748.

Hutchison CA, Plant T, Drayson M, et al. (2008) Serum free light chain measurement aids the diagnosis of myeloma in patients with severe renal failure. BMC Nephrol; 9: 11

Jacobus SJ, Kumar S, Uno H, et al. (2011) Impact of high-risk classification by FISH: an eastern cooperative oncology group (ECOG) study E4A03. British.journal of haematology., 155: 340-348.

Kapoor P, Fonseca R, Rajkumar SV, et al. (2010) Evidence for cytogenetic and fluorescence in situ hybridization risk stratification of newly diagnosed multiple myeloma in the era of novel therapie. Mayo Clinic Proceedings, 85: 532-537.

Katzmann JA, Abraham RS, Dispenzieri A, et al. (2005) Diagnostic performance of quantitative and free light chain assays in clinical practice. Clinical Chemistry; 51: 878-881.

Katzmann JA, Kyle RA, Benson J, et al. (2009) Screening panels for detection of monoclonal gammopathies. Clin Chem; 55: 1517-1522.

Koulieris E, Panayiotidis P, Harding SJ, et al. (2012) Ratio of involved/uninvolved immunoglobulin quantification by Hevylite assay: clinical and prognostic impact in multiple myeloma. Experimental Hematology & Oncology, 1: 9.

Kumar S, Zhang L, Dispenzieri A, et al. (2010) Relationship between elevated immunoglobulin free light chain and the presence of IgH translocations in multiple myeloma. Leukemia, 24: 1498-1505.

Kumar S, Fonseca R, Ketterling RP, et al. (2012) Trisomies in multiple myeloma: impact on survival in patients with high-risk cytogenetics.[Erratum appears in Blood. 2014 Mar 6;123(10):1621]. Blood, 119: 2100-2105.

Lai YY, Huang XJ, Cai Z, et al. (2012) Prognostic power of abnormal cytogenetics for multiple myeloma: a multicenter study in China. Chinese Medical Journal, 125: 2663-2670.

Larsen JT, Kumar SK, Dispenzieri A, et al. (2013) Serum free light chain ratio as a biomarker for high-risk smoldering multiple myeloma. Leukemia, 27: 941-946.

Lopez-Corral L (2012) Genomic analysis of high-risk smoldering multiple myeloma. Haematologica, 97: 1439-1443.

Lu J, Lu J, Chen W, et al. (2014) Clinical features and treatment outcome in newly diagnosed Chinese patients with multiple myeloma: results of a multicenter analysis. Blood Cancer Journal, 4: e239.

Ludwig H, Milosavljevic D, Zojer N, et al. (2013) Immunoglobulin heavy/light chain ratios improve paraprotein detection and monitoring, identify residual disease and correlate with survival in multiple myeloma patients.[Erratum appears in Leukemia. 2013 Apr;27(4):996]. Leukemia, 27: 213-219.

Maltezas D, Dimopoulos MA, Katodritou I, et al. (2013) Re-evaluation of prognostic markers including staging, serum free light chains or their ratio and serum lactate dehydrogenase in multiple myeloma patients receiving novel agents. Hematological Oncology, 31: 356-362.

Mateo G, Montalban MA, Vidriales MB, et al. (2008) Prognostic value of immunophenotyping in multiple myeloma: a study by the PETHEMA/GEM cooperative study groups on patients uniformly treated with high-dose therapy. Journal of Clinical Oncology, 26: 2737-2744.

Mateos MV, Gutiérrez NC, Martín-Ramos ML, et al. (2011) Outcome according to cytogenetic abnormalities and DNA ploidy in myeloma patients receiving short induction with weekly bortezomib followed by maintenance. Blood, 118: 4547-4553.

McTaggart MP, Lindsay J, & Kearney EM (2013) Replacing urine protein electrophoresis with serum free light chain analysis as a first-line test for detecting plasma cell disorders offers

increased diagnostic accuracy and potential health benefit to patients. Am J Clin Pathol; 140: 890-897.

Milla F, Oriol A, Aguilar J, et al. (2001) Usefulness and reproducibility of cytomorphologic evaluations to differentiate myeloma from monoclonal gammopathies of unknown significance. Am J Clin Pathol; 115: 127-135.

Minarik J, Scudla V, Ordeltova M, et al. (2011) Combined measurement of plasma cell proliferative and apoptotic index in multiple myeloma defines patients with good and poor prognosis. Leukemia Research, 35: 44-48.

Minarik J (2005) Evaluation of plasma cell propidium-iodide and annexin-V indices: their relation to prognosis in multiple myeloma. Biomedical papers of the Medical Faculty of the University Palacky, Olomouc, Czechoslovakia, 149: 271-274.

Minarik J, Scudla V, Bacovsky J, et al. (2010) Thalidomide and bortezomib overcome the prognostic significance of proliferative index in multiple myeloma. Neoplasma, 57: 8-14.

Moreau P (2007) Heterogeneity of t(4;14) in multiple myeloma. Long-term follow-up of 100 cases treated with tandem transplantation in IFM99 trials. Leukemia, 21: 2020-2024.

Neben K, Jauch A, Bertsch U, et al. (2010) Combining information regarding chromosomal aberrations t(4;14) and del(17p13) with the International Staging System classification allows stratification of myeloma patients undergoing autologous stem cell transplantation. Haematologica, 95: 1150-1157.

Neben K, Jauch A, Hielscher T, et al. (2013) Progression in smoldering myeloma is independently determined by the chromosomal abnormalities del(17p), t(4;14), gain 1q, hyperdiploidy, and tumor load. Journal of Clinical Oncology, 31: 4325-4332.

Nemec P, Zemanova Z, Kuglik P, et al. (2012) Complex karyotype and translocation t(4;14) define patients with high-risk newly diagnosed multiple myeloma: Results of CMG2002 trial. Leukemia & lymphoma, 53: 920-927.

Nowakowski GS, Witzig TE, Dingli D, et al. (2005) Circulating plasma cells detected by flow cytometry as a predictor of survival in 302 patients with newly diagnosed multiple myeloma. Blood, 106: 2276-2279.

Paiva B, Vidriales MB, Perez JJ, et al. (2009a) Multiparameter flow cytometry quantification of bone marrow plasma cells at diagnosis provides more prognostic information than morphological assessment in myeloma patients. Haematologica, 94: 1599-1602.

Paiva B, Vidriales MB, Mateo G, et al. (2009b) The persistence of immunophenotypically normal residual bone marrow plasma cells at diagnosis identifies a good prognostic subgroup of symptomatic multiple myeloma patients. Blood, 114: 4369-4372.

Paiva B, Gutierrez NC, Chen X, et al. (2012a) Clinical significance of CD81 expression by clonal plasma cells in high-risk smoldering and symptomatic multiple myeloma patients. Leukemia, 26: 1862-1869.

Paiva B, Vídriales MB, Montalbán MÁ, et al. (2012b) Multiparameter flow cytometry evaluation of plasma cell DNA content and proliferation in 595 transplant-eligible patients with myeloma included in the Spanish GEM2000 and GEM2005<65y trials. American.journal of pathology, 181: 1870-1878.

Paiva B, Gutiérrez NC, Rosiñol L, et al. (2012c) High-risk cytogenetics and persistent minimal residual disease by multiparameter flow cytometry predict unsustained complete response after autologous stem cell transplantation in multiple myeloma. Blood, 119: 687-691.

Paiva B, Vidriales MB, Rosinol L, et al. (2013) A multiparameter flow cytometry immunophenotypic algorithm for the identification of newly diagnosed symptomatic myeloma with an MGUS-like signature and long-term disease control.[Erratum appears in Leukemia. 2013 Oct;27(10):2112]. Leukemia, 27: 2056-2061.

Park JW, Kim YK, Bae EH, et al. (2012) Combined analysis using extended renal reference range of serum free light chain ratio and serum protein electrophoresis improves the diagnostic accuracy of multiple myeloma in renal insufficiency. Clinical Biochemistry; 45: 740-744.

Rajkumar SV, Gupta V, Fonseca R, et al. (2013) Impact of primary molecular cytogenetic abnormalities and risk of progression in smoldering multiple myeloma. Leukemia, 27: 1738-1744.

Piehler,AP, Gulbrandsen N, Kierulf P, et al. (2008) Quantitation of Serum Free Light Chains in Combination with Protein Electrophoresis and Clinical Information for Diagnosing Multiple Myeloma in a General Hospital Population. Clinical Chemistry; 54: 1823-1830.

Rajkumar S, Dimopoulos MA, Palumbo A, et al. (2014) International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma. Lancet Oncology 15(12): e538-48

Shin SJ, Lee H, Jung G, et al. (2014) Expression of CD99 in Multiple Myeloma: A Clinicopathologic and Immunohistochemical Study of 170 Cases. The Korean Journal of Pathology, 48: 209-216.

Snozek CL, Katzmann JA, Kyle RA, et al. (2008) Prognostic value of the serum free light chain ratio in newly diagnosed myeloma: proposed incorporation into the international staging system. Leukemia, 22: 1933-1937.

Tinguely M, Jenni B, Reineke T, et al. (2007) Chromosomal translocations t(4;14), t(11;14) and proliferation rate stratify patients with mature plasma cell myelomas into groups with different survival probabilities: a molecular epidemiologic study on tissue microarrays. American Journal of Surgical Pathology, 31: 690-696.

van RF, Bolejack V, Hollmig K, 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: 827-832.

Vermeersch P, Van Hoovels L, Delforge M, et al. (2008) Diagnostic performance of serum free light chain measurement in patients suspected of a monoclonal B-cell disorder. Br J Haematol; 143: 496-502.

Walker BA, Leone PE, Chiecchio L, et al. (2010) A compendium of myeloma-associated chromosomal copy number abnormalities and their prognostic value. Blood, 116: e56-e65

Wolff F, Thiry C & Willems D (2007) Assessment of the analytical performance and the sensitivity of serum free light chains immunoassay in patients with monoclonal gammopathy. Clinical Biochemistry; 40: 351-354.

Xu Y, Sui W, Deng S, et al. (2013) Further stratification of patients with multiple myeloma by International Staging System in combination with ratio of serum free to light chains. Leukemia & lymphoma, 54: 123-132.

3 Imaging investigations

3.1 Imaging for people with suspected myeloma

The diagnosis of myeloma is based on a combination of clinical features, laboratory tests (these are covered in chapter 2) and radiological findings.

Skeletal survey (a series of plain X-rays) has traditionally been the primary imaging investigation used to diagnose myeloma bone disease in the UK. This imaging test is widely available, low cost and has relatively low radiation exposure. However, there is variation in whether the long bones are included in a skeletal survey and it is known that this form of imaging is less sensitive than newer techniques.

Other imaging techniques include whole body computed tomography (WB-CT), magnetic resonance imaging (MRI) and positron emission tomography CT (PET-CT). These are more sensitive and specific than the skeletal survey and may identify soft tissue lesions not seen on skeletal examination. However, they are more costly and may increase radiation burden (especially PET CT and WB-CT).

This section covers people in secondary care with a known plasma cell disorder on laboratory investigations which is suspected to be myeloma. It does not cover investigation of MGUS.

Clinical question: What is the optimal imaging strategy for patients with suspected myeloma?

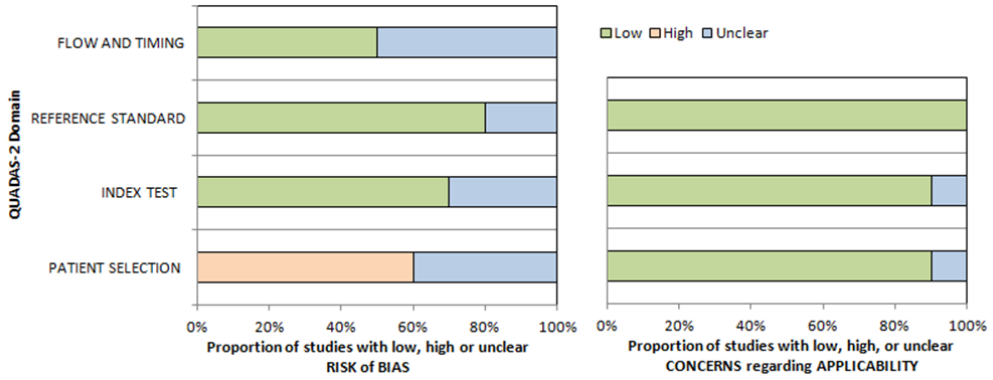
Clinical evidence (see also Appendix G)

Study quality

The QUADAS-2 assessment tool was used to evaluate risk of bias in the studies. Generally there was a low risk of bias across the studies and the studies were found to be applicable to the review question. For some of the studies the risk of bias is unclear due to under-reporting in some studies of the timing of the index and reference tests and whether they were interpreted blind to each other's results.

There was most uncertainty in the patient selection methods: many studies did not report this. Some studies were considered to have a high risk of bias in the patient selection category as the population did not include controls i.e. patients without myeloma.

Figure 2: Risk of bias and applicability across studies



Diagnostic accuracy

12 studies were identified and included in the evidence review. 10 studies used biopsy as the reference standard whilst 2 studies used x-ray. All 12 studies reported sensitivity for myeloma. Only 6 reported specificity (due to a lack of people without myeloma in the other 6 studies). The data can be seen in tables 6 and 7. Some studies reported high sensitivity with MRI and TC99MIBI bone scan, however there was considerable heterogeneity in sensitivity and specificity estimates. This could be related to the differences in techniques and diagnostic criteria used in the individual studies.

Patient acceptability, Radiation exposure

We did not find evidence for these outcomes.

1 **Table 6: diagnostic accuracy of various imaging methods compared to the reference standard biopsy**

Index tests	Study	Myeloma prevalence	TP	FN	FP	TN	sensitivity	specificity	PPV	NPV
MRI	Whole body (WB) MRI (Cascini et al., 2013)	100%	22	0	NR	NR	100%	-	-	-
	WB MRI (Erten et al., 2007)	100%	11	2	NR	NR	85%	-	-	-
	WB MRI - focal lesions (Kloth, 2014)	75%	259	150	33	105	63%	76%	87%	41%
	WB MRI – any bone marrow infiltration (Kloth, 2014)	75%	251	158	53	85	61%	62%	83%	35%
	Spinal MRI STIR (Myslivecek et al., 2008)	79%	38	3	0	11	93%	100%	100 %	79%
	Spinal MRI T1 w.i. (Myslivecek et al., 2008)	79%	38	3	6	5	93%	45%	86%	63%
	Spinal MRI SI - b1000 image (Dutoit, 2014)	41%	55	9	45	46	86%	51%	55%	84%
	Spinal MRI ADC1000 value (Dutoit, 2014)	41%	48	16	61	30	75%	33%	44%	65%
FDG PET/CT	Cascini et al., 2013	100%	18	4	NR	NR	82%	-	-	-
	Sager et al., 2011	100%	29	3	NR	NR	90%	-	-	-
x-ray bone survey	Sohn et al., 2002	100%	14	8	NR	NR	64%	-	-	-
	Alper et al., 2003	100%	18	2	NR	NR	90%	-	-	-
	Alexandrakis et al, 2001	100%	26	2	NR	NR	93%	-	-	-
TC99MIBI bone scan	Myslivecek et al., 2008	79%	39	2	0	11	95%	100%	100 %	85%
	Svaldi et al., 2001	66%	58	0	2	28	100%	93%	97%	100 %
	Alexandrakis et al, 2001	100%	22	6	NR	NR	79%	-	-	-
	Alper et al., 2003	100%	20	0	NR	NR	100%	-	-	-
	Erten et al., 2007	100%	17	1	NR	NR	94%	-	-	-

Index tests	Study	Myeloma prevalence	TP	FN	FP	TN	sensitivity	specificity	PPV	NPV
TC99MDP bone scan	Sohn et al., 2002	100%	11	11	NR	NR	50%	-	-	-
	Alexandrakis et al, 2001	100%	15	13	NR	NR	54%	-	-	-
	Alper et al., 2003	100%	15	5	NR	NR	75%	-	-	-
Bone marrow immunoscintigraphy (BMIS) using technetium-99m-labelled AGA	Sohn et al., 2002	100%	18	4	NR	NR	82%	-	-	-

Table 7: diagnostic accuracy of various imaging methods compared to the reference standard x-ray

Index tests	study	Myeloma prevalence	TP	FN	FP	TN	sensitivity	specificity	PPV	NPV
TC99MIBI	Catalano et al., 1999	100%	7	3	3	10	70%	77%	70%	77%
FDG-PET CT	Zamagni et al., 2007	100%	12	4	21	9	75%	30%	36%	69%

TP: true positive, FN: false negative, FP: false positive, TN: true negative, PPV: positive predictive value, NPV: negative predictive value, NR: not reported

Cost-effectiveness evidence (see also Appendix A)

Uncertainty remains around whether performing cross-sectional imaging for diagnosis is cost effective compared to skeletal survey and if so which imaging modalities are optimal. The aim of the economic analysis was to assess the cost effectiveness of skeletal survey compared to whole body CT (WB-CT), MRI spine with plain radiograph of the long bones (MRI spine), whole body MRI (WB-MRI) and PET-CT for diagnosis in secondary care patients with a plasma cell disorder suspected to be myeloma.

Economic evidence statement

A systematic literature review was performed to assess the current economic literature for this topic. The review identified 463 possibly relevant economic papers relating to myeloma. Of these, no papers were deemed relevant for this topic and therefore no papers were included in the review of existing economic evidence.

De novo economic analysis

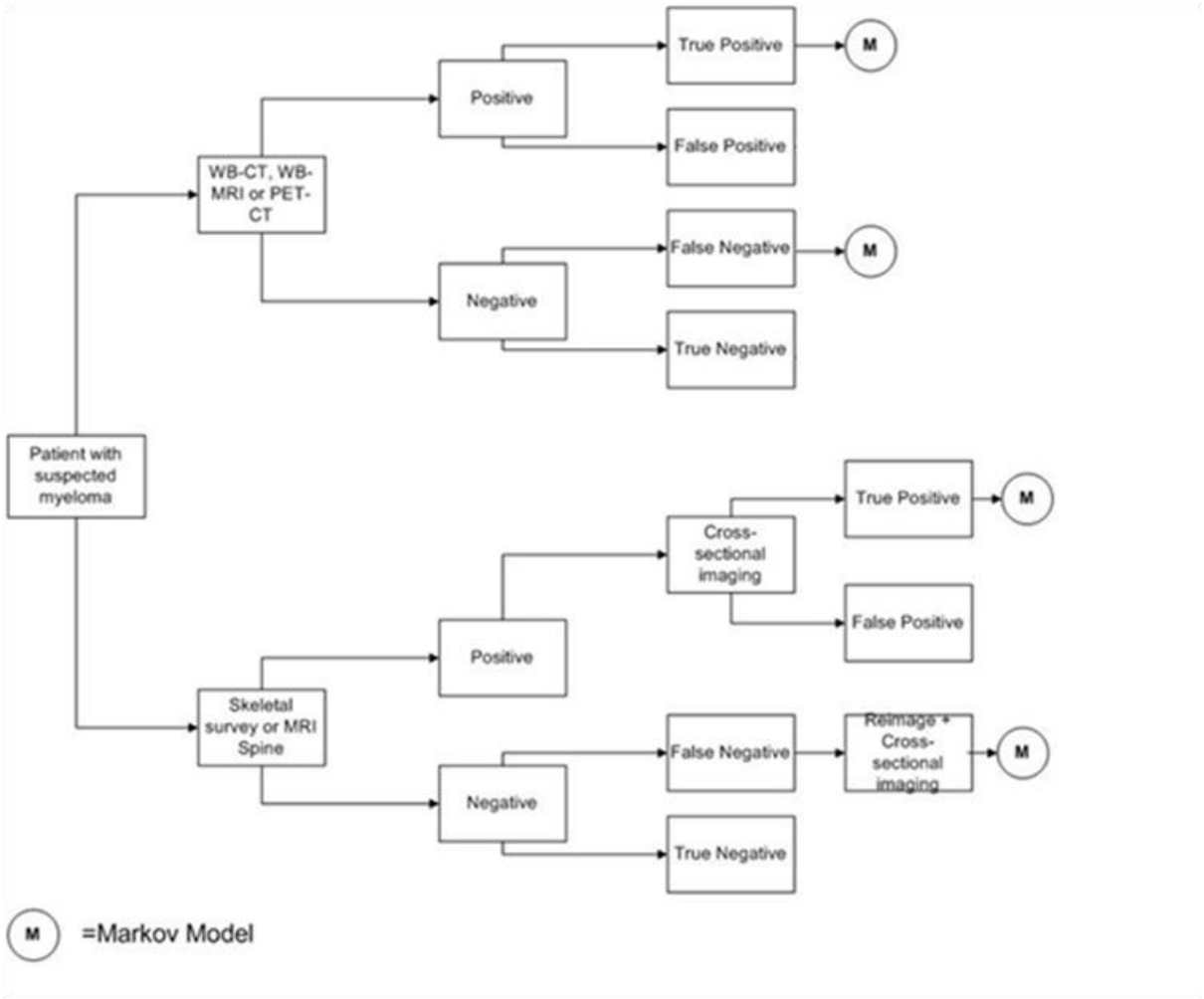
A decision tree model was created to estimate the proportion of people correctly diagnosed using each competing imaging strategy. Following the decision tree portion of the model all people with myeloma were assigned to a simple Markov model with two states, alive and dead, All people with correct non-myeloma diagnoses (true negatives) and incorrect myeloma diagnoses (false positives) were not followed up after the decision tree component of the model.

Model structure

The economic model considered five potential diagnostic imaging interventions skeletal survey, WB-CT, MRI spine (with plain radiograph of the long bones), WB-MRI and PET-CT.

Patients receive either a positive or negative result based upon the diagnostic accuracy of the imaging modality. Patients in the skeletal survey or MRI spine arm of the model receive cross-sectional imaging to guide treatment decisions, assumed to be WB-CT in the base case, if initial imaging is positive for myeloma. People in the WB-CT, WB-MRI or PET-CT arms are assumed to have received sufficient imaging to be able to make these treatment decisions. The structure of the model is shown in Figure 3.

Figure 3: Structure of the economic model



Following the initial decision tree portion of the model people with myeloma enter one of three Markov models based on their time from first symptoms upon entering the model and whether they received a correct diagnosis or not. Discussion of the alternate Markov models is presented later.

The model assumes that people with myeloma who receive a false negative result from imaging will receive diagnostic reimaging six months later which would always correctly diagnose myeloma. People falsely diagnosed with myeloma were assumed to receive a correct diagnosis at cross-sectional imaging or at a reimaging six months later before receiving treatment.

Prevalence

A systematic review identified no studies reporting on the prevalence of myeloma amongst people receiving diagnostic imaging for plasma cell disorders suspected to be myeloma. Prevalence figures therefore had to be estimated from other myeloma populations. For the base case prevalence figures were taken from 1684 patients at the US Mayo clinic in 2006 with a M protein in the serum or urine. Two other similar studies were identified (Table 8). Both these alternate prevalence estimates were used as a one-way sensitivity analysis.

Table 8: Summary of estimated prevalence of myeloma, MGUS and other disorders in imaged population

	Myeloma	MGUS	Other

	Myeloma	MGUS	Other
USA (Kyle & Rajkumar, 2007)	19.8%	54.7%	25.5%
Sweden (Bird et al., 2009)	19%	72%	4%
Italy (Malacrida et al. 1987)	25.6%	69.6%	4.8%

Diagnostic accuracy

Eight studies considering interventions included as part of the economic modelling were identified by the accompanying clinical evidence review (Alexandrakis et al, 2001; Sohn et al., 2002; Alper et al., 2003; Zamagni et al., 2007; Myslivecek et al., 2008; Sager et al., 2011; Cascini et al., 2013; Erten et al., 2007). Two other studies were identified which only covered interventions outside of those considered in the economic analysis (Catalano et al., 1999; Svaldi et al., 2001). Evidence was found on all included interventions apart from WB-CT.

The evidence assessed using QUADAS-2 was of moderate to low quality. The risk of bias in the included studies was generally low or unclear due to underreporting with regards to the timing and blinding of the index and reference tests. There was great uncertainty in the patient selection methods as many studies did not report this and it was unclear whether a consecutive or random sample of patients had been recruited and if inappropriate exclusions had been avoided. It was therefore impossible to tell how representative they are of the patient group that would be seen in NHS practise. Some studies were highlighted as having a high risk of bias in the patient selection category as the population did not include a control group (i.e. negative reference standard results) meaning that the specificity of the test could not be determined. The reviewer highlighted to the Guideline Committee (GC) that only 4 of the 10 included studies included negative cases (not myeloma) to determine specificity. This again implies that the patient populations are not generalisable given that the majority of cases seen during diagnosis would be negative for myeloma. It also makes it unclear whether studies which reported a high sensitivity did so at the expense of a lower specificity and consequently more false positives.

Given these issues with the included evidence, especially the issues of applicability, patient selection and lack of negative cases in the patient cohorts the GC found it difficult to give weight to the values reported in the evidence review and to estimate values for inclusion in the economic model. This was particularly true around values of specificity. The base case values used in the economic model (Table 9), whilst based on the evidence review where possible are intended to be illustrative and not an estimate or ranking of the diagnostic accuracy of the different imaging modalities. From the GC's clinical experience and supported somewhat by low quality evidence MRI was assumed to be the most diagnostically accurate, in terms of both sensitivity and specificity.

Table 9: Sensitivity and specificity used in the base case economic model

	Sensitivity	Specificity	Source (Sensitivity,Specificity)
Skeletal Survey	64%	50%	Sohn, Illustrative value
WB-CT	94%	94%	Illustrative value, Illustrative value
MRI Spine	93%	95%	Myslivecek, Illustrative value
WB-MRI	100%	95%	Cascini, Illustrative value
PET-CT	90%	35%	Sager, Zamagni

Given the lack of or weak evidence around the diagnostic accuracy estimates they were extensively explored during sensitivity analysis. Sensitivity and specificity were also assigned a wide, uninformative uniform distribution, between 50% and 100% during probabilistic sensitivity analysis. This was deemed appropriate to cover all potential sensitivity and specificity values based on the GCs clinical experience.

Population demographics

The age and sex of our modelled cohort was likely to be similar to that reported in the Kariyawasan et al (2007) study used for the clinical inputs in the model described in detail below (Kariyawasan et al., 2007). The sex and age of the cohort had no influence on the outcome of the economic evaluation and are intended to be illustrative of the likely population cohort in the UK.

Clinical inputs

A systematic review identified two studies which linked time to diagnosis to survival and myeloma related complications.(Friese et al., 2009; Kariyawasan et al., 2007) Kariyawasan et al (2007) investigated the relationship between time to diagnosis of myeloma and number of complications and survival. The study concluded that time to diagnosis predicted both the number of complications and disease free survival of people with myeloma.

Friese et al was a US retrospective study of 5483 patients diagnosed with myeloma. The study concluded that time to diagnosis did not predict outcomes in patients with myeloma (OR 0.9, CI 0.8-1.1)

In the base case economic model outcomes from Kariyawasan et al (2007) were used. Even though the study size was significantly smaller than Friese et al (2009) the patient group and treatment pathway was more likely to be reflective of that in the NHS. A sensitivity analysis was performed that assumed there was no improvement in patient outcomes from prompter diagnosis (a de-facto cost-minimisation analysis) to investigate the robustness of the results to these differing conclusions.

Kariyawasan et al (2007) grouped time from first symptoms to diagnosis into three groups 0-3 months, 3-6 months and greater than 6 months. The outcomes from these groupings were used to inform the Markov models. The proportion of patients in each group is reported in Table 10.

The model assumed that the time between first symptom and diagnosis in people with a correct diagnosis would be identical to the time between first symptom and presentation (i.e. diagnosis would be made shortly after presentation in secondary care) and would enter the Markov model corresponding to this. People with a false non-myeloma diagnosis were assumed to have an additional 6 months until diagnosis and would always enter the 'greater than 6 months' Markov model. Table 10 shows the Markov model entered by time from first symptom at presentation in secondary care and the results of diagnostic imaging.

Table 10: Time from first symptom to presentation and Markov pathway following diagnostic result.

Time from first symptoms to presentation	Proportion of model cohort on entry	Markov entered True Positive Result	Markov entered False Negative result
0-3 months	21.6%	0-3 months	>6 months
3-6 Months	21.6%	3-6 Months	>6 months
>6 months	56.9%	>6 months	>6 months

Complications

Complication rates used in the model were identical to those reported in Kariyawasan et al., (2007). Kariyawasan et al., (2007) reported five different types of complications: infection, neurological, renal disease, bone disease and anaemia (Table 11).

Table 11: Proportion of complications at presentation for time from first symptoms to diagnosis

	Time from first symptoms to diagnosis		
Complication	0-3 Months	3-6 Months	>6 Months
Infection	7.1%	4.8%	11.6%
Neurological	10.7%	4.8%	7.0%
Renal Disease	21.4%	19.1%	55.8%
Bone Disease	25.0%	38.1%	62.8%
Anaemia	7.1%	38.1%	93.0%

Health-related quality of life

Health related quality of life values used in the model were taken from one study (Proskorovsky et al., 2014) of 154 patients with a current diagnosis of myeloma presenting for routine care at five UK and six German sites using the EQ-5D and Uk population preference weights.

Proskorovsky et al reported their utility values for four groups: Asymptomatic, Mildly Symptomatic, Moderately Symptomatic and Severely Symptomatic based on the number and the severity of the symptoms reported. Summary of the utility values and the definition for each grouping is presented in Table 12. For the economic model people with no complications reported had a utility value equal to that of the asymptomatic group and those presenting with any complication had a utility value equal to that of the moderately symptomatic group.

Table 12: Symptom level and reported utility value

Symptom Level	Definition	Utility Value
Asymptomatic	Patient had no symptoms/AEs	0.923
Mildly Symptomatic	≥1 mild symptom/AE no moderate or severe symptoms/AE	0.806
Moderately Symptomatic	≥1 moderate symptom/AE no severe symptoms/AE	0.675
Severely Symptomatic	≥1 severe symptom/AE	0.501

AE: Adverse Event

Overall survival

Survival for the three Markov models was based on an annual probability of survival based on the person's Durie-Salmon stage of myeloma at the time of diagnosis, disaggregated again by time from first symptoms to diagnosis, reported in Kariyawasan et al., (2007).

The annual probability of survival for each stage was based on one retrospective study of the clinical and laboratory data of 10,750 previously untreated people with myeloma from 17 institutions including centres in Europe and North America between 1981 and 2002 (Greipp et al., 2005). Median survival and annual probabilities of survival are reported in Table 13.

Table 13: Estimates of median and annual probability of survival used in the economic model

Durie-Salmon Stage	Median Survival (Months)	Annual Probability Survival
1a	69	88.6%
1b	22	68.5%
2a	58	86.6%
2b	34	78.3%
3a	45	83.1%

Durie-Salmon Stage	Median Survival (Months)	Annual Probability Survival
3b	24	70.7%

The economic model considered a time horizon of 10 years. This was considered sufficient to cover all significant differences in costs and quality of life between the different imaging modalities.

Costs

Imaging costs

The costs of the differing imaging modalities were taken from NHS reference costs apart from skeletal survey for which reference costs were not reported (Table 14). Skeletal survey costs were taken from internal recharge costs used in one UK myeloma centre (King's College Hospital, personal communication, April 4, 2015).

Table 14: Imaging costs used in the base case analysis and probabilistic sensitivity analysis (PSA)

	Reference Cost	Source	PSA
Skeletal Survey	£108.82	Internal recharge (Personal correspondence)	Triangular(£54,163)
CT(Whole body)	£147.17	NHS Reference Costs	Gamma(α=8.2, β=17.9)
MRI (spine)	£199.01	NHS Reference Costs+50% cost Skeletal Survey	Gamma(α=17.8, β=11.2)
MRI (Whole Body)	£203.06	NHS Reference Costs	Gamma(α=10.0, β=20.2)
PET-CT	£651.96	NHS Reference Costs	Gamma(α=7.0, β=92.7)

Appointment costs

All appointment costs were taken from NHS reference costs (Table 15).

Complication costs

With a lack of evidence around costs associated with these complications it was assumed that they would result in one additional consultant led appointment costed as 'consultant led non-admitted follow-up appointment' (Table 15).

Table 15: Other costs used in the base case analysis and probabilistic sensitivity analysis

	Cost	Source	PSA
First Appointment	£212.83	NHS Reference Costs	Gamma(α=5.3, β=40.2)
Subsequent Appointments	£156.41	NHS Reference Costs	Gamma(α=5.8, β=27.1)
Complication Costs	£156.41	NHS Reference Costs	Gamma(α=5.8, β=27.1)

All costs in the economic model were already at 2014 prices, the latest for which inflation figures were available. Therefore it was unnecessary to inflate any costs. All costs and QALYs were discounted at 3.5% per annum as recommended by the NICE Guidelines Manual (National Institute for Health and Care Excellence, 2014).

Sensitivity analysis

For the base case analyses a range of deterministic threshold and probabilistic sensitivity analyses were carried out.

Incremental net monetary benefit

All results are presented as incremental net monetary benefit (INMB) assuming a willingness to pay per QALY of £20,000.

Results

Deterministic base case results

Table 16 shows the base case results for the different imaging modalities. WB-CT, MRI spine and WB-MRI are cost effective when compared to skeletal survey alone with them all being cost saving and health improving. WB-MRI showed the largest rise in incremental QALYs although this was directly as a result of it being illustratively assigned the highest sensitivity. WB-CT had the highest INMB although it was only marginally higher than that of WB-MRI. PET-CT was the only intervention to report a negative NMB. These results were consistent when the other two estimates of prevalence were used.

Table 16: Deterministic base case results for a willingness to pay of £20,000 per QALY

	Incremental Cost	Incremental QALYs	INMB
Skeletal Survey	Reference	Reference	Reference
WB-CT	-£ 142.40	0.0119	£ 379.49
MRI spine	-£ 33.39	0.0115	£ 262.57
WB-MRI	-£ 92.06	0.0142	£ 376.56
PET-CT	£ 792.85	0.0103	-£ 587.37

Probabilistic base case results

The probabilistic base case results are shown in Table 17. In our probabilistic results the strategy of WB-CT ends up being the least costly option followed by skeletal survey, WB-MRI and MRI spine. This is also a de facto cost minimisation and shows WB-CT as a preferred option when we assume promptness of diagnosis has no impact on health outcomes as reported in Friese et al. (2009).

Table 17: Probabilistic base case results

	Incremental Cost
Skeletal Survey	Reference
WB-CT	-£ 10.96
MRI spine	£ 93.84
WB-MRI	£ 60.61
PET-CT	£ 639.84

Sensitivity/specificity

A sensitivity analysis was performed assuming an arbitrary specificity (80%) for all imaging interventions. All other values were identical to the base case. Two other sensitivity analyses were performed, one assuming 80% sensitivity across all imaging modalities and one assuming both 80% sensitivity and specificity. When 80% specificity is assumed in all interventions then the ranking of interventions remains the same with WB-CT remaining the preferred option. Similar results are seen for 80% sensitivity with WB-CT remaining the preferred option. When 80% sensitivity and specificity is assumed for all interventions there

will be no difference in QALYs between interventions so the preferred option will also always be the least costly (WB-CT). The conclusions were identical when both 60% and 100% specificities were used. Whilst these results were explicitly arbitrary they provided a starting point for threshold sensitivity analysis presented below (Table 18).

Assuming a starting point of perfect diagnostic accuracy if skeletal survey had sensitivity less than 4 percentage points or a specificity less than 12 percentage points of that of WB-CT then WB-CT became both cost saving and health improving. For the same to be true of MRI these values needed to be 5 percentage points and 15 percentage point respectively. Higher values were needed for the same to be true of MRI spine although it never became the preferred option over either WB-CT or WB-MRI. PET-CT was never preferred to skeletal survey, or any of the other intervention, for any values of diagnostic accuracy. These conclusions were not sensitive to the starting point in terms of sensitivity and specificity with the results being consistent for all starting values of sensitivity and specificity.

Table 18: Incremental net monetary benefit around different Sensitivity and Specificity assumptions

	Sensitivity			Specificity			Both=		
	80%	60%	100%	80%	60%	100%	80%	60%	100%
Skeletal Survey	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref	Ref
WB-CT	£122	£114	£129	£272	£272	£272	£14	£7	£22
MRI Spine	£16	£12	£19	£153	£153	£153	-£94	-£97	-£90
WB-MRI	£61	£50	£73	£260	£251	£269	-£55	-£76	-£34
PET-CT	-£830	-£877	-£783	-£369	-£450	-£288	-£611	-£740	-£483

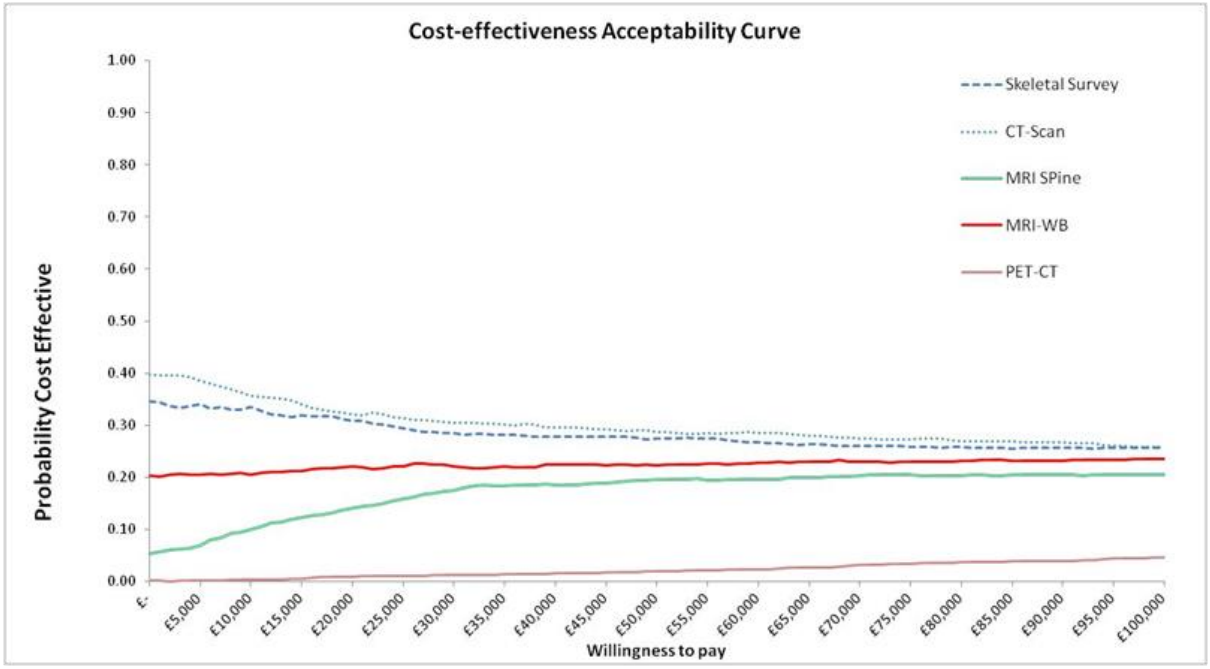
Utility values and survival

The use of alternate methods of estimating the utility values did not change the conclusions of the analysis. The same was true for survival with improvements in survival increasing the INMB of all cross sectional imaging approaches in the base-case.

Probabilistic sensitivity analysis

It can be seen from the cost effective acceptability curve (Figure 4) that WB-CT remains the most likely preferred option for all willingness to pay values up to £100,000 followed closely by skeletal survey. At £20,000 per QALY there is a 32% probability that WB-CT is the preferred option closely followed by skeletal survey (31%). MRI spine and PET-CT only become the most probable preferred option for values significantly exceeding the NICE threshold of £20,000.

Figure 4: Cost effectiveness acceptability curve



As the model considers three cross-sectional imaging approaches compared to two alternate options it is possible that when the distributions of the PSA allow for favourable parameter estimates for a cross-sectional approach to be preferred it will be distributed across a greater number of potential options. This potentially underestimates the probability of the cost effectiveness of these cross-sectional approaches compared to the comparator of skeletal survey. Therefore, the CEAC analysis was re-run to just compare WB-CT (Figure 5) and WB-MRI (Figure 6) to skeletal survey.

Under these direct comparisons both WB-CT and WB-MRI are the preferred option when compared to skeletal survey both for a willingness to pay of £20,000 and £0 (where the least costly option is preferred).

Figure 5: Cost effectiveness acceptability curve: whole body CT versus skeletal survey

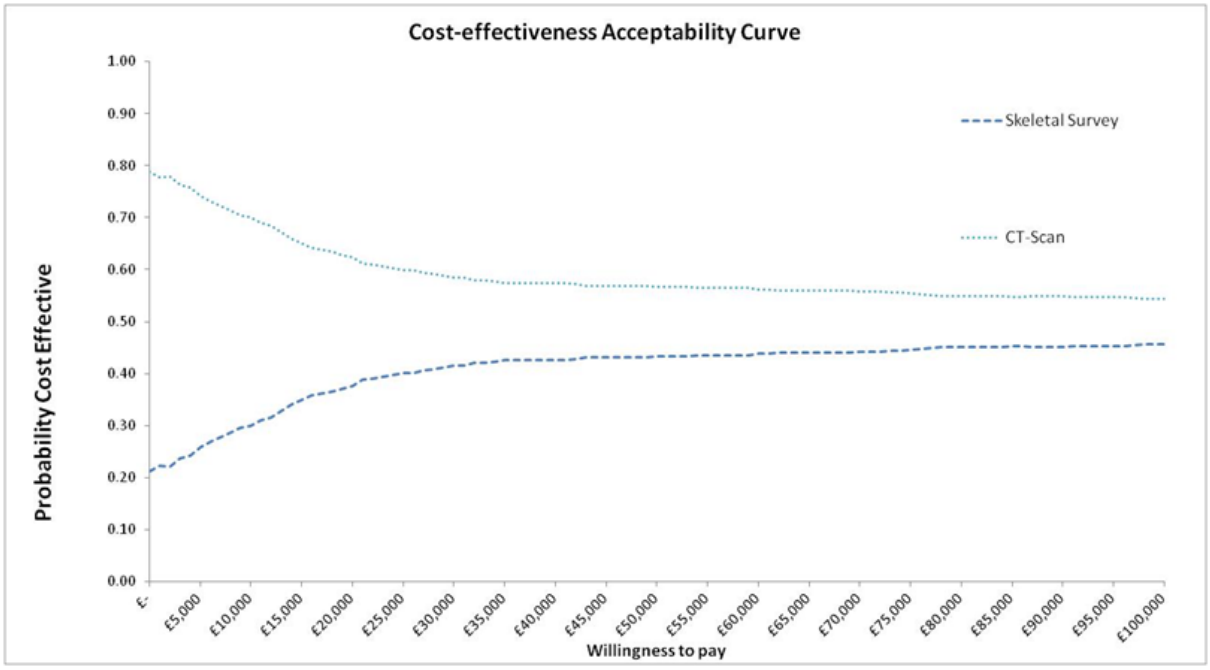
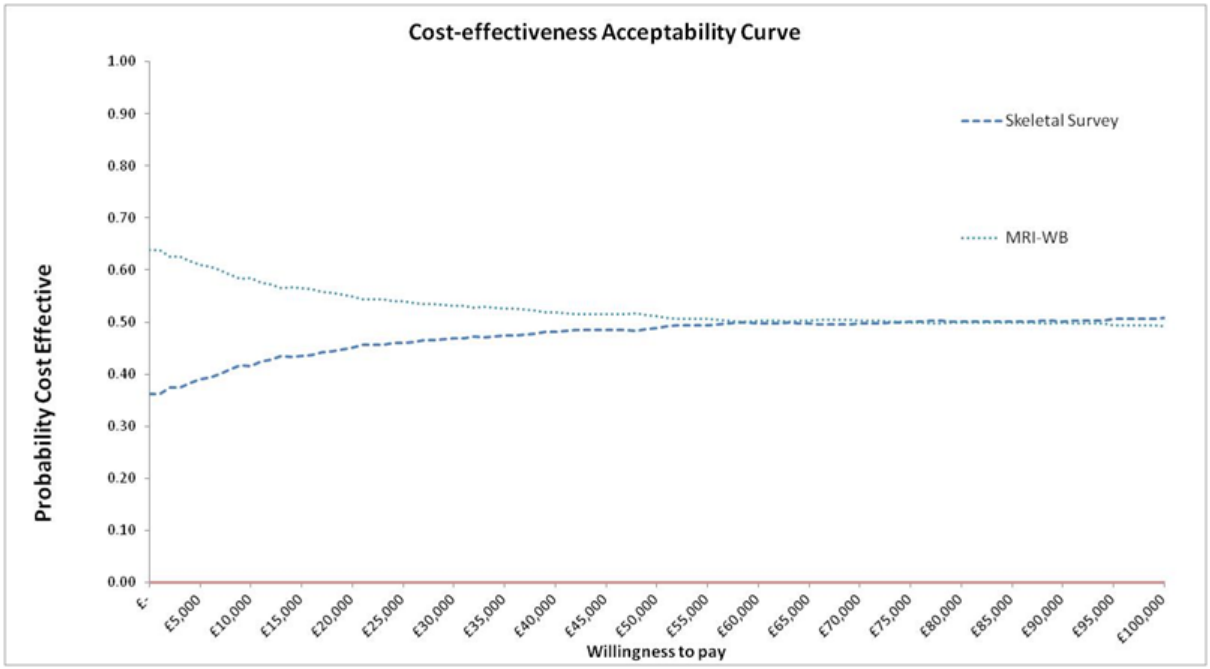


Figure 6: Cost effectiveness acceptability curve: whole body –MRI versus skeletal survey



Conclusion

Even under the very conservative assumption that there is no difference in diagnostic accuracy between the different imaging modalities there is a strong case that an approach of using cross-sectional imaging at the time of diagnosis is a cost effective strategy for diagnosis in patients with a plasma cell disorder suspected to be myeloma. The main driver of this result appears to be the avoidance of the need for further cross-sectional imaging, to guide treatment decisions, following a positive result on skeletal survey. Even under these conservative assumptions this approach could be both cost saving and health improving

even with the use of WB-CT or WB-MRI being the preferred option in greater than 50% of cases even when the health provider's willingness to pay per QALY is zero. The case becomes stronger when the cross-sectional imaging starts to have a higher diagnostic accuracy than skeletal survey with the illustrative base case values again suggesting an approach using either WB-CT or WB-MRI could be cost saving and health improving. It is unclear which is the most cost-effective between WB-CT and WB-MRI given a paucity of good evidence around diagnostic accuracy and the decision sensitivity to differences in diagnostic accuracy between both. Whilst it was the GCs opinion that MRI was the most sensitive of the considered imaging modalities it was difficult to quantify by exactly how much, if at all, without higher quality evidence. It is unclear which is the most cost-effective between WB-CT and WB-MRI given a paucity of good evidence around diagnostic accuracy and the decision sensitivity to differences in diagnostic accuracy between both. Whilst it was the GCs opinion that MRI was the most sensitive of the considered imaging modalities it was difficult to quantify by exactly how much, if at all, without higher quality evidence.

Recommendations	**Offer imaging to all people with a plasma cell disorder suspected to be myeloma.** **Consider whole-body MRI as first-line imaging.** **Consider whole-body low-dose CT as first-line imaging if whole-body MRI is unsuitable or the person declines it.** **Only consider skeletal survey as first-line imaging if whole-body MRI and whole-body low-dose CT are unsuitable or the person declines them.** **Do not use isotope bone scans to identify myeloma-related bone disease in people with a plasma cell disorder suspected to be myeloma.**
Relative value placed on the outcomes considered	The Guideline Committee considered the outcomes of diagnostic accuracy, radiation exposure and patient acceptability to be the most relevant in determining the optimal imaging strategy for patients with suspected myeloma. Lesion detection rate had been included as an outcome in the clinical question but the reviewer noted that this question concerned the use of imaging at diagnosis, i.e., whether a patient has lesions or not. The number of lesions is not important for diagnosis, although may provide information for management and follow-up. The Guideline Committee agreed that lesion detection rate was not relevant for this question and as such did not include evidence on this outcome. Evidence was reported on diagnostic accuracy, but no evidence was identified for radiation exposure and patient acceptability. When drafting the recommendations the Guideline Committee considered diagnostic accuracy to be the most important although this evidence was limited.
Quality of the evidence	The quality of the evidence for the outcome of diagnostic accuracy was assessed using QUADAS-2 as moderate to low quality. Generally there was a low risk of bias across the studies and the studies were found to be applicable to the review question. It was noted that for some of the studies the risk of bias was unclear as there was under-reporting in some studies with regards to the timing of the index and reference tests and whether they were conducted blind to each other.

	The greatest uncertainty was in the patient selection methods as many studies did not report this and so it was unclear whether a consecutive or random sample of patients had been recruited and if inappropriate exclusions had been avoided. Some studies were highlighted as having a high risk of bias in the patient selection category as the population did not include a control group, meaning that the specificity of the test could not be determined. The Guideline Committee noted that only 4 of the 10 included studies included negative cases (not myeloma) to determine specificity. So although many of the diagnostic imaging tests report a high sensitivity the implications of using a highly sensitive test are unclear due to a lack of information about test specificity. Using a more sensitive test could increase the number of false positives. Other limitations of the included studies were that they were all single centre studies with small sample sizes.
Trade off between clinical benefits and harms	The Guideline Committee recommended that whole body MRI be used as first line imaging as the evidence indicated that this method was the most sensitive. This was also in line with the Guideline Committee's clinical experience. There was no clinical evidence about the diagnostic accuracy ofWB-CT but, based on their clinical experience, the Guideline Committee agreed that it was an effective option where MRI was not suitable. The Guideline Committee highlighted a number of situations in which MRI would be unsuitable and soWB-CT should be considered. These include patients who could not tolerate MRI (claustrophobic), patients with pacemakers and patients who would find it difficult/painful lying on their back for a long period of time. Skeletal survey was reported to be less accurate when compared to WB-CT and whole body MRI. However the Guideline Committee still thought skeletal survey could be useful if both MRI and CT were not suitable. The Guideline Committee made a recommendation not to use isotope bone scans as the clinical evidence reported it to be of low sensitivity/accuracy. The Guideline Committee concluded that the benefits of the improved imaging for diagnosis would be earlier and more accurate diagnosis, with minimal need for second line imaging. The Guideline Committee agreed that a potential harm of the use of WB-CT was radiation exposure. However they minimised this by recommending that MRI should be considered first and CT only if MRI was not suitable.
Trade-off between net health benefits and resource use	The Guideline Committee noted that no relevant economic evaluations were identified for this topic. As this topic was considered a high economic priority a de novo economic analysis was performed. A decision tree and Markov model was used to assess the cost effectiveness of WB-CT, MRI spine, whole body MRI, and PET-CT compared to skeletal survey in patients with a plasma disorder suspected to be myeloma. The model assumed that by performing cross-sectional imaging (WB-CT, whole body MRI and PET-CT) at

	the time of diagnosis it would avoid this imaging needing to be done at the treatment planning stage. During the base case analysis WB-CT was found to be the preferred imaging modality, having the highest net monetary benefit, being both cost saving and health improving when compared to skeletal survey. Both whole body MRI and MRI spine were also cost saving and health improving, with whole body MRI having the largest QALY gain, although they reported lower NMBs than WB-CT. However, the difference in NMB was less than £3 in the case of whole body MRI and WB-CT and their ranking is dependent on the assumptions used in the model. The results were sensitive to the prevalence of myeloma with both MRI and CT no longer being cost effective if the prevalence was below 10% and 7% below our base case estimate of 19.6% respectively. PET-CT was never the preferred imaging option even under very favourable assumptions and is unlikely to be cost effective for the diagnosis of myeloma. The results were not sensitive to changes in other parameters in the economic model. Importantly results were not sensitive to diagnostic accuracy - a parameter the Guideline Committee had difficulty estimating given the paucity of evidence in the clinical evidence review. Probabilistic results in which all imaging modalities were given equal and wide distributions around both sensitivity and specificity (including a very conservative assumption for cross-sectional imaging) showed WB-CT was still cost saving. The probabilistic sensitivity analysis (again using the conservative assumptions around diagnostic accuracy) suggested that whole body MRI and CT were cost saving in the majority of iterations. Given that the Guideline Committee's clinical experience and weak evidence from the clinical evidence review suggested that both whole body MRI and CT were more diagnostically accurate than skeletal survey there is a good probability that these interventions are both cost saving and health improving. MRI spine and PET-CT were always the less preferred option to all other interventions when the willingness to pay per QALY was below £100,000. The model did not consider radiation exposure given that adverse effects were likely to be limited due to the life expectancy of patients with myeloma. Whilst WB-CT was the preferred option to whole body MRI in most iterations the difference in probability was small when under the conservative equal distributions of diagnostic accuracy. When MRI is considered a few percentage points more accurate than WB-CT, as in the base case, MRI becomes the preferred option. Given the adverse events from the increased radiation burden of WB-CT compared to MRI, which were not included in the analysis, and the Guideline Committee's clinical opinion that MRI is more accurate than WB-CT there is a reasonable probability that whole body MRI is the most cost effective intervention for the diagnosis of myeloma in this patient group.

3.2 Imaging for people with newly diagnosed myeloma

Once myeloma has been diagnosed, it is important to establish whether the patient has smouldering myeloma or myeloma. This is achieved by a combination of laboratory testing and imaging.

The main imaging techniques used are skeletal survey, CT, MRI and PET-CT. These vary in their sensitivity and specificity for detecting myeloma bone disease, bone marrow infiltration

and extra medullary disease. There is also variation in anatomical coverage, radiation exposure, suitability and practicality for each test. In addition there is uncertainty on which modality to use for certain sites, for example skull, ribs.

Currently there is limited access to multi-parametric MRI, WB CT and PET-CT in some areas.

Patients with non-secretory myeloma are much harder to assess as there is no laboratory marker to use. Imaging is particularly valuable in these patients.

Clinical question: What is the most effective imaging to guide treatment decisions in patients with newly diagnosed myeloma?

Clinical evidence (see also Appendix G)

Study quality

A modified version of the QUADAS-2 assessment tool was used to evaluate the risk of bias and applicability concerns in the included studies. It was clear a priori that it would not be likely that any studies included a reference standard, so it was therefore decided not to make this a part of the inclusion criteria, although this strategy naturally means that none of the index/comparator test results were verified. Consequently, it is not possible to know, based on the present evidence, which of the index/comparison tests is better when the results differ between the tests, nor indeed if the results are correct even when they do not differ between the included tests.

In a number of the included studies, it was unclear whether the patient selection was consecutive (Baur-Melnyk et al., 2008; Bäuerle et al, 2009; Fonti et al., 2009; Lin et al., 2014; Mahnken et al., 2002; Spinnato et al., 2012) and in one study it was clear that it was not (Wolf et al., 2014; high risk) whereas in the remainder patient selection was consecutive and therefore considered at low risk of bias (Kröpil et al, 2008; Nanni et al., 2006; Princewill et al., 2013, Razek et al., 2013).

The majority of the studies employed blinded assessment of the index and comparator tests, that is, the results were blinded, at least, to those of the other imaging tests, and were therefore considered at low risk whereas the remaining four studies did not employ blinded reading of the index and comparator test results and, consequently, these studies were rated at high risk of bias (Baur-Melnyk et al., 2008; Kröpil et al, 2008; Mahnken et al., 2002; Nanni et al., 2006).

The time interval between the index and comparator tests was acceptable in all but two of the included studies where it was unclear (Kröpil et al, 2008; Wolf et al., 2014).

Generally the studies were found to be applicable to the review question in terms of the index/comparator tests employed and, for the most part, the populations. However, the applicability of the populations of four studies was unclear (Lin et al., 2014; Mahnken et al., 2002; Princewill et al, 2013; and Wolf et al., 2014) as these populations seemed to either be subject to excessive exclusions (for the present purposes: Lin et al., 2014), consist of a narrow range of patients (i.e., all stage III who may or may not have been treated, Mahnken et al., 2002) or be a mix of patients only some of whom are applicable to the current question (Princewell et al., 2013; Wolf et al., 2014).

The small sample sizes of all the included studies should also be noted as a clear limitation.

Imaging results

11 studies were identified and included in the evidence review. None of the studies employed a reference standard to verify the imaging results.

CT identified more lesions than radiography (3 studies [Kröpil et al., 2008; Princewill et al., 2013; Razek et al., 2013], N = 108; low quality) and was also associated with a higher radiation exposure than radiography (2 studies [Kröpil et al., 2008; Princewill et al., 2013], N = 80; low quality);

MRI identified more lesions than radiography (1 study [Wolf et al., 2014], N = 119; low quality);

MRI and CT each identified more lesions than radiography (1 study, N = 18 [Mahnken et al., 2002]; low quality);

PET-CT identified more lesions than radiography and an equivalent number of lesions to MRI in half of the included patients with more and less lesions detected, respectively, in the other two quarters of patients, compared to MRI (1 study [Nanni et al., 2006], N = 28; low quality);

MRI identified more regions affected by myeloma than CT (1 study [Baur-Melnyk et al., 2008], N = 41; low quality);

WB-MRI identified more extensive disease than axial skeleton MRI (1 study [Bäuerle et al., 2009], N = 73; low quality)

MRI identified a different pattern of disease than PET-CT (3 studies [Fonti et al., 2008; Lin et al., 2014; Spinnato et al., 2012], N = 239; low quality)

Risk of second primary cancers, patient acceptability, and prognostic accuracy for progression-free survival and overall survival:

We did not find evidence for these outcomes.

Cost-effectiveness evidence

A literature review of published cost-effectiveness analyses did not identify any relevant papers for this topic. Whilst there were potential cost implications of making recommendations in this area, other questions in the guideline were agreed as higher priorities for economic evaluation. Consequently no further economic modelling was undertaken for this question.

Recommendations	**For people with newly diagnosed myeloma or smouldering myeloma who have not had whole-body imaging with 1 of the following, consider whole-body imaging to assess for myeloma-related bone disease and extra-medullary plasmacytomas with one of:** • **MRI** • **CT** • **fluorodeoxyglucose positron emission tomography CT (FDG PET-CT).** **For guidance on imaging for people with suspected spinal cord compression, see the NICE guideline on metastatic spinal cord compression.** **Consider baseline whole-body imaging with MRI or FDG PET-CT for people who have non-secretory myeloma or suspected or confirmed soft tissue plasmacytomas and have not already had 1 of these tests.**
Relative value placed on the outcomes considered	The Guideline Committee considered the outcome of diagnostic yield to be the most important outcome in guiding treatment decisions because the number and site of the lesions detected

	would affect management of patients with newly diagnosed myeloma. No evidence was found for patient acceptability and prognostic accuracy for progression-free survival and overall survival.
Quality of the evidence	The quality of the evidence was assessed by QUADAS-II and was of low quality for all outcomes. The Guideline Committee noted three major issues with the evidence: (1) There was no reference standard, (2) for most of the test comparisons, the evidence consisted of only one study, and (3) the sample sizes were generally very small. As a result of these limitations and the low quality evidence the Guideline Committee used their clinical experience alongside the evidence and made 'consider' recommendations rather than 'offer'. Given the lack of evidence to determine which imaging modality is the best test for patients with newly diagnosed myeloma, the Guideline Committee made a recommendation for further research, comparing the effectiveness of MRI, FDG-PET/CT and CT.
Trade off between clinical benefits and harms	The Guideline Committee noted that all patients with newly diagnosed myeloma should receive prompt imaging because this is a major determinant of management strategy. The Guideline Committee also noted that in addition to skeletal lesions, whole body imaging picks up lesions outside the spine, in the soft tissue and in all areas with marrow. The Guideline Committee therefore decided to recommend whole-body imaging in all patients with newly diagnosed or smouldering myeloma to assess for skeletal disease and extra-medullary plasmacytomas with MRI, FDG-PET/CT or CT. This information can be used to make treatment decisions and also a as a baseline to assess future response. The Guideline Committee noted that the evidence did not allow them to distinguish between the relative merits of MRI, FDG-PET/CT or CT, but that the evidence did show that they were all superior to conventional radiography. They therefore decided, based on their clinical experience, to recommend MRI, FDG-PET/CT and CT. The Guideline Committee noted that non-secretory disease cannot be adequately assessed by CT. Based on their clinical experience, they therefore also recommended baseline whole body FDG-PET/CT or MRI in patients with suspected or confirmed non-secretory disease. These modalities were also recommended to assess soft tissue plasmacytomas to act as a baseline and for follow-up. The Guideline Committee concluded that effective imaging would result in a number of benefits, including better assessment of bone disease, people getting the right treatment sooner, prevention of deterioration in symptoms, and the incidental identification of other serious, but hitherto unknown lesions that can then receive earlier treatment, all of which the Guideline Committee expected to lead to improved outcomes. The Guideline Committee agreed that potential harms would be: radiation exposure from CT and FDG-PET/CT; issues with patient

	acceptability (e.g. claustrophobia) or contra-indications for the tests recommended (e.g. pacemakers); over-diagnosis where the recommended tests may identify lesions, requiring investigation, which turn out to be of no significance. Overall the Guideline Committee agreed that the benefits of the recommendations in terms of improved patient outcomes, outweighed the potential harms.
Trade-off between net health benefits and resource use	The Guideline Committee noted that no relevant published economic evaluations had been identified and no additional economic analysis had been undertaken in this area. There may be some cost savings resulting from using cross sectional imaging as a 'one stop shop' for diagnosis and to guide treatment decisions. The recommendations are likely to result in more cross sectional imaging, which will increase costs. This increase would be offset against a decrease in the number of skeletal surveys being performed. The Guideline Committee considered that cross sectional imaging is already being done following skeletal surveys.
Other considerations	The Guideline Committee noted that the recommendations will result in a significant change in practice. Currently, skeletal surveys are routinely used as the main imaging modality in myeloma, whereas implementation of these recommendations requires the use of other imaging modalities. Although all the recommended imaging modalities are available within the NHS, there may be extra pressure on the services to provide this imaging and some capacity issues.

Research recommendation	**Observational studies should be carried out comparing the effectiveness of whole-body MRI, fluorodeoxyglucose positron emission tomography CT (FDG PET-CT) and whole-body low-dose CT in detecting lesions that may determine the start of treatment for people with newly diagnosed myeloma. Outcomes of interest are lesion detection, sensitivity and specificity for myeloma-related bone disease, patient acceptability, incremental upstaging, radiation exposure, risk of second primary cancer, the impact of additional information on predicting progression-free survival, overall survival and skeletal-related events.**
Why this is important	Newer imaging techniques are replacing skeletal surveys for assessing myeloma-related bone disease in people with newly diagnosed myeloma. However, the most effective technique is not known.

References

Alexandrakis MG, Kyriakou DS, Passam F, et al. (2001). Value of Tc-99m sestamibi scintigraphy in the detection of bone lesions in multiple myeloma: comparison with Tc-99m methylene diphosphonate. Annals of Hematology, 80, 349-353.

Alper E, Gurel M, Evrensel T, et al. (2003). 99mTc-MIBI scintigraphy in untreated stage III multiple myeloma: comparison with X-ray skeletal survey and bone scintigraphy. Nuclear Medicine Communications, 24, 537-542.

Bauerle T, Hillengass J, Fechtner K, et al. (2009) Multiple myeloma and monoclonal gammopathy of undetermined significance: importance of whole-body versus spinal MR imaging. Radiology, 252: 477-485.

Baur-Melnyk A, Buhmann S, Becker C, et al. (2008) Whole-body MRI versus whole-body MDCT for staging of multiple myeloma. AJR, American Journal of Roentgenology. 190: 1097-1104.

Bird J, Behrens J, Westin J, et al. (2009). UK Myeloma Forum (UKMF) and Nordic Myeloma Study Group (NMSG): guidelines for the investigation of newly detected MGÇÉproteins and the management of monoclonal gammopathy of undetermined significance (MGUS). British journal of haematology, 147, 22-42.

Cascini GL, Falcone C, Console D, et al. (2013). Whole-body MRI and PET/CT in multiple myeloma patients during staging and after treatment: personal experience in a longitudinal study. Radiologia Medica, 118, 930-948.

Catalano L, Pace L, Califano C, et al. (1999). Detection of focal myeloma lesions by technetium-99m-sestaMIBI scintigraphy. Haematologica, 84, 119-124.

Dutoit JC, Vanderkerken MA & Anthonissen (2014). The diagnostic value of SE MRI and DWI of the spine in patients with monoclonal gammopathy of undetermined significance, smouldering myeloma and multiple myeloma. European Radiology, 24, 2754-2765

Erten N, Saka B, Berberoglu K, et al. (2007). Technetium-99m 2-methoxy-isobutyl-isonitrile uptake scintigraphy in detection of the bone marrow infiltration in multiple myeloma: correlation with MRI and other prognostic factors. Annals of Hematology, 86, 805-813.

Fonti R, Salvatore B, Quarantelli M, et al. (2008) 18F-FDG PET/CT, 99mTc-MIBI, and MRI in evaluation of patients with multiple myeloma. Journal of Nuclear Medicine, 49: 195-200.

Friese CR, Abel GA, Magazu LS, et al. (2009). Diagnostic delay and complications for older adults with multiple myeloma. Leukemia & lymphoma, 50, 392-400.

Greipp PR, San Miguel J, Durie BG, et al. (2005). International staging system for multiple myeloma. Journal of Clinical Oncology, 23, 3412-3420.

Kariyawasan CC, Hughes DA, Jayatillake MM et al. (2007). Multiple myeloma: causes and consequences of delay in diagnosis. QJM, 100, 635-640.

Kloth JK (2014). Appearance of monoclonal plasma cell diseases in whole-body magnetic resonance imaging and correlation with parameters of disease activity. International Journal of Cancer, 135, 2380-2386.

Kropil P, Fenk R, Fritz LB, et al. (2008) Comparison of whole-body 64-slice multidetector computed tomography and conventional radiography in staging of multiple myeloma. European Radiology, 18: 51-58.

Kyle RA & Rajkumar SV (2007). Monoclonal gammopathy of undetermined significance and smouldering multiple myeloma: emphasis on risk factors for progression. British journal of haematology, 139, 730-743.

Lin C, Ho CL, Ng SH, et al. (2014) C-11-Acetate as a new biomarker for PET/CT in patients with multiple myeloma: initial staging and postinduction response assessment. European Journal of Nuclear Medicine and Molecular Imaging, 41: 41-49.

Mahnken AH, Wildberger JE, Gehbauer G, et al. (2002) Multidetector CT of the spine in multiple myeloma: comparison with MR imaging and radiography. AJR, American Journal of Roentgenology. 178: 1429-1436.

Malacrida VITT, De Francesco D, Banfi G, et al. (1987). Laboratory investigation of monoclonal gammopathy during 10 years of screening in a general hospital. Journal of clinical pathology, 40, 793-797.

Myslivecek M, Nekula J, Bacovsky J, et al. (2008). Multiple myeloma: predictive value of Tc-99m MIBI scintigraphy and MRI in its diagnosis and therapy. Nuclear Medicine Review, 11, 12-16.

Nanni C, Zamagni E, Farsad M, et al. (2006) Role of 18F-FDG PET/CT in the assessment of bone involvement in newly diagnosed multiple myeloma: preliminary results. European journal of nuclear medicine and molecular imaging, 33: 525-531.

National Institute for Health and Care Excellence (2014). The guidelines manual. London: NICE: 2014. National Institute for Clinical Excellence, 13, 4-2007.

Princewill K, Kyere S, Awan O et al. (2013) Multiple myeloma lesion detection with whole body CT versus radiographic skeletal survey. Cancer Investigation, 31: 206-211.

Proskorovsky I, Lewis P, Williams CD, et al. (2014). Mapping EORTC QLQ-C30 and QLQ-MY20 to EQ-5D in patients with multiple myeloma. Health and quality of life outcomes, 12, 35.

Razek AA, Ezzat A, Azmy E et al. (2013) Role of whole-body 64-slice multidetector computed tomography in treatment planning for multiple myeloma. Radiologia Medica, 118: 799-805.

Sager S, Ergul N, Ciftci H, et al. (2011). The value of FDG PET/CT in the initial staging and bone marrow involvement of patients with multiple myeloma. Skeletal Radiology, 40, 843-847.

Sohn SK, Ahn BC, Lee SW, et al. (2002). Bone marrow immunoscintigraphy using technetium-99m anti-granulocyte antibody in multiple myeloma. European Journal of Nuclear Medicine & Molecular Imaging, 29, 591-596.

Spinnato P, Bazzocchi A, Brioli A, et al. (2012) Contrast enhanced MRI and 8F-FDG PET-CT in the assessment of multiple myeloma: a comparison of results in different phases of the disease. European Journal of Radiology, 81: 4013-4018.

Svaldi M, Tappa C, Gebert U, et al. (2001). Technetium-99m-sestamibi scintigraphy: an alternative approach for diagnosis and follow-up of active myeloma lesions after high-dose chemotherapy and autologous stem cell transplantation. Annals of Hematology, 80, 393-397.

Wolf MB, Murray F, Kilk K, et al. (2014) Sensitivity of whole-body CT and MRI versus projection radiography in the detection of osteolyses in patients with monoclonal plasma cell disease. European Journal of Radiology, 83: 1222-1230.

Zamagni E, Nanni C, Patriarca F, et al. (2007). A prospective comparison of 18F-fluorodeoxyglucose positron emission tomography-computed tomography, magnetic resonance imaging and whole-body planar radiographs in the assessment of bone disease in newly diagnosed multiple myeloma. Haematologica, 92, 50-55.

4 Smouldering myeloma

Around 10-15% of people diagnosed with myeloma will have no evidence of myeloma related organ or tissue injury or a myeloma defining event. These people have what is termed smouldering (asymptomatic) myeloma. It is known that most of these patients will develop myeloma requiring treatment at some time in the future.

Historically it has been thought that patients with smouldering myeloma do not require specific treatments as this has not improved the long term consequences of the myeloma. More recently, with the introduction of newer more effective and better tolerated therapies, it is suggested that some patients with smouldering myeloma may benefit from earlier treatment. It is also suggested that the availability of more sensitive ways of assessing myeloma may identify specific groups of patients with smouldering myeloma who may benefit from earlier treatment with either the same chemotherapy treatments used to treat myeloma patients or specific treatments for asymptomatic myeloma.

Clinical question: What are the most effective primary management strategies (including observation) for patients with smouldering (asymptomatic) myeloma?

Clinical evidence (see also Appendix G)

See Tables 19-21.

Overall survival

Low quality evidence from five randomised trials (Mateos et al, 2013; Witzig et al, 2013; Hjorth et al, 1993; Riccardi et al, 2000; D'Arena et al 2011) including 552 patients with asymptomatic myeloma suggests uncertainty about the effect of immediate treatment on overall survival, when compared to treatment deferred until progression (HR 1.00; 95% C.I. 0.71 to 1.40; where HR < 1 favours immediate treatment).

Low quality evidence came from two trials which used immediate treatment with thalidomide plus zoledronate (Witzig et al, 2013) or lenalidomide plus dexamethasone (Mateos et al 2013). Pooling these IMiD trials suggests uncertainty about whether immediate treatment improves overall survival (HR 0.61; 95% C.I. 0.30 to 1.24; where HR < 1 favours immediate treatment), although Mateos et al (2013) did report a significant overall survival benefit with immediate treatment with lenalidomide plus dexamethasone (HR 0.31; 95% C.I. 0.10 to 0.94; where HR < 1 favours immediate treatment).

Progression to symptomatic disease

Low quality evidence from two randomised trials including 187 patients with asymptomatic myeloma (Mateos et al 2013; Witzig et al, 2013) suggests that immediate treatment with an IMiD regimen delays the progression to symptomatic disease (HR 0.36; 95% C.I. 0.23 to 0.55; where HR < 1 favours immediate treatment). In Mateos et al (2013) three year symptomatic progression free survival was around 78% in patients who received immediate treatment compared to 30% in those with deferred treatment.

Low quality evidence from two randomised trials including 340 patients with asymptomatic myeloma (Musto et al 2008; D'Arena et al, 2011) suggests uncertainty about the effect of treatment with bisphosphonates on progression to symptomatic disease when compared to observation alone (HR 0.94; 95% C.I. 0.72 to 1.23; where HR < 1 favours immediate treatment).

Disease progression (including biological progression)

Witzig et al (2013) defined disease progression as increased M-protein level 25% above the lowest level or new bone lesion or plasmacytoma. Using this definition of progression, low quality evidence suggests immediate treatment with lenalidomide plus zoledronate was more effective than treatment with zoledronate alone (HR 0.51; 95% C.I. 0.28 to 0.91).

Skeletal related events

Low quality evidence from two randomised trials including 274 patients with asymptomatic myeloma (D'Arena et al 2011; Musto et al 2008) suggests that immediate treatment with bisphosphonates reduces the risk of skeletal related events compared to observation alone (RR 0.61; 95% C.I. 0.45 to 0.81; where RR<1 favours bisphosphonate treatment). These figures suggest that an additional skeletal related event could be avoided for every ten patients treated with bisphosphonates instead of observation alone.

Low quality evidence from two RCTS (Hjorth et al 1993; Riccardi et al, 2000) including 188 patients with asymptomatic myeloma suggests uncertainty over whether immediate treatment melphalan and prednisone lowers the risk of vertebral compression when compared to deferred treatment (RR 0.19; 95% C.I. 0.02 to 1.60; where RR <1 favours immediate treatment). In these studies no vertebral compression occurred in the immediate treatment whereas 4% of patients in the deferred treatment group experienced vertebral compression.

Treatment related adverse events

Low quality evidence from two randomised trials including 187 patients (Mateos et al 2013; Witzig et al, 2013) suggests uncertainty about whether immediate IMiD treatment is associated with an increased rate of grade 3-4 adverse events (RR 1.70; 95% C.I. 0.60 to 5.06; where RR>1 favours deferred treatment).

Low quality evidence from three randomised trials including 288 patients (Mateos et al, 2013; Hjorth et al, 1993; Riccardi et al 2000) suggests that immediate treatment is associated with an increased risk of a second primary cancer when compared to deferred treatment (RR 4.49; 95% C.I. 1.15 to 17.49; where RR>1 favours deferred treatment).

Osteonecrosis of the jaw occurred in 1.3% of those treated with bisphosphonates (D'Arena et al 2011; Musto et al 2008; Witzig et al, 2013).

Outcomes not reported

HRQOL, patient acceptability, renal failure and disease related mortality were not reported in the trials.

Table 19: GRADE profile: What are the most effective primary management strategies (including observation) for patients with asymptomatic myeloma (immediate IMiD treatment versus deferred treatment)?

Quality assessment							No of patients		Effect		Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Immediate IMiD treatment	Deferred treatment	Relative (95% CI)	Absolute	
Overall survival (event is death from any cause)											
2[1]	randomised trials	serious[4]	no serious inconsistency	no serious indirectness	serious[3]	none	13/92 (14.1%)	22/95 (23.2%)	HR 0.61 (0.3 to 1.24)	-	LOW
Time to disease progression (event is progression to symptomatic disease)											
2[1]	randomised trials	serious[4]	no serious inconsistency	no serious indirectness	serious[3]	none	39/92 (42.4%)	72/95 (75.8%)	HR 0.31 (0.2 to 0.48)	-	LOW
Grade 3 or 4 adverse effects											
2[1]	randomised trials	serious[4]	no serious inconsistency	no serious indirectness	serious[3]	none	24/92 (26.1%)	15/95 (15.8%)	RR 1.74 (0.6 to 5.06)	117 more per 1000 (from 63 fewer to 641 more)	LOW

1 Mateos 2013; Witzig 2013; 3 Low number of events; 4 Allocation concealment and sequence generation unclear; no blinding

Table 20: GRADE profile: What are the most effective primary management strategies (including observation) for patients with asymptomatic myeloma (immediate mephalan+prednisone treatment versus deferred treatment)?

Quality assessment	No of patients	Effect	Quality

No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Immediate mephalan + prednisone treatment	Deferred treatment	Relative (95% CI)	Absolute	
Overall survival (event is death from any cause)											
2[1]	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[3]	none	58/97 (59.8%)	47/91 (51.6%)	HR 1.39 (0.78 to 2.47)	-	LOW
Time to disease progression (event is progression to symptomatic disease)											
1[4]	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[3]	none	5/72 (6.9%)	34/66 (51.5%)	HR 0.11 (0.05 to 0.24)	-	LOW
Acute leukaemia											
2[1]	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[3]	none	4/97 (4.1%)	1/93 (1.1%)	RR 3.01 (0.47 to 19.43)	22 more per 1000 (from 6 fewer to 198 more)	LOW
Secondary primary cancer											
2[1]	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[3]	none	6/82 (7.3%)	1/87 (1.1%)	RR 4.20 (0.71 to 23.57)	41 more per 1000 (from 2 fewer to 291 more)	LOW
Vertebral compression											
2[1]	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[3]	none	0/97 (0%)	4/91 (4.4%)	RR 0.19	41 more per	LOW

Quality assessment							No of patients		Effect		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Immediate mephalan + prednisone treatment	Deferred treatment	Relative (95% CI)	Absolute	Quality
									(0.02 to 1.60)	1000 (from 2 fewer to 291 more)	

1 Riccardi 2000; Hjorth 1993; 2 Allocation concealment and sequence generation unclear; no blinding; 3 Low number of events; 4 Riccardi 2000

Table 21: GRADE profile: What are the most effective primary management strategies (including observation) for patients with asymptomatic myeloma (immediate bisphosphonate treatment versus deferred treatment)?

Quality assessment							No of patients		Effect		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Immediate bisphosph-onate treatment	Deferred treatment	Relative (95% CI)	Absolute	Quality
Overall survival (event is death from any cause)											
2	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2,4]	none	0/89 (0%)	0/88 (0%)	Not estimable	-	LOW
Time to disease progression (event is progression to symptomatic disease)											
2[3]	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[4]	none	90/170 (52.9%)	90/170 (52.9%)	HR 0.94 (0.72 to 1.23)	-	LOW
Skeletal events											
2[3]	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[4]	none	24/126 (19%)	38/127 (29.9%)	RR 0.64 (0.41 to 0.99)	108 fewer per 1000 (from 3 fewer	LOW

Quality assessment							No of patients		Effect		Quality
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Immediate bisphosph-onate treatment**	**Deferred treatment**	**Relative (95% CI)**	**Absolute**	
										to 177 fewer)	
Osteonecrosis of the jaw											
2[3]	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[4]	none	2/170 (1.2%)	0/170 (0%)	RR 5.06 (0.25 to 103.83)	12 more per 1000 with bisphospho-nates	LOW

1 *1 Not intention-to-treat analysis in D'Arena (2011); no blinding in Musto (2008) or D'Arena (2011); 2 Number of deaths not reported; 3 Musto 2008, D'Arena 2011; 4 Low*
2 *number of events*

3

4

5

1 Cost-effectiveness evidence

2 A literature review of published cost-effectiveness analyses did not identify any relevant
3 papers for this topic. Whilst there were potential cost implications of making
4 recommendations in this area, other questions in the guideline were agreed as higher
5 priorities for economic evaluation. Consequently no further economic modelling was
6 undertaken for this question.

7

Recommendations	No recommendations were made for clinical practice
Relative value placed on the outcomes considered	The Guideline Committee considered disease-related mortality, overall survival, progression-free survival, progression to symptomatic myeloma, prevention of renal failure, health-related quality of life, patient acceptability, adverse events and skeletal related events to be the most relevant outcomes when determining the most effective primary management strategy for patients with asymptomatic myeloma. No evidence was found for health-related quality of life, patient acceptability, renal failure or disease-related mortality. The Guideline Committee considered the outcomes of overall survival and adverse events the most important when agreeing recommendations
Quality of the evidence	The evidence was assessed by GRADE methodology and appropriate NICE checklists as low quality for all reported outcomes. The Guideline Committee noted that allocation concealment was unclear in some of the studies and there was no blinding. Also the studies all had small sample sizes. In addition the Guideline Committee were aware that the International Myeloma Working Group has recently changed its' definitions of smouldering myeloma, such that some patients previously considered to have smouldering myeloma would now be considered to have myeloma. In addition the criteria used by the International Myeloma Working Group to define this population differed from the criteria used to define high risk smouldering myeloma in the evidence reviewed. The patient populations included in our evidence base included people with high risk disease who would no longer be termed smouldering disease under the current IMWG definition. The Guideline Committee were therefore unclear as to which patients might actually benefit from receiving treatment. Given these limitations with the evidence the Guideline Committee agreed not to make any recommendations for clinical practice. Instead they agreed to recommend further research into the most effective treatment strategy for people with smouldering myeloma (as defined by the International Myeloma Working Group 2014 classification). However the Guideline Committee acknowledged that in order for this research to be possible, additional research would be needed to identify the most effective way to risk stratify people with smouldering myeloma. They therefore also recommended this.
Trade off between clinical benefits and harms	As no recommendations were made, these issues were not discussed.
Trade-off between net	

health benefits and resource use	

Research recommendation	**A randomized multi-centre prospective trial should be carried out for patients with newly diagnosed smouldering myeloma (as defined by the International Myeloma Working Group 2014 classification) to:** • **identify which combinations of FISH, molecular technologies, bone marrow plasma cell percentage, whole-body imaging, immunophenotype, serum-free light-chain levels or ratio, Hevylite®, paraprotein levels, immunoparesis, and International Staging System (ISS) are most effective at risk stratification for people with smouldering myeloma.** • **compare fixed duration treatment (with or without bone-directed therapy), continuous treatment (with or without bone-directed therapy) and no treatment (with or without bone-directed therapy)** **Outcomes of interest are time to biochemical and/or clinical progression, overall survival, adverse events, quality of life and resource use.**
Why this is important	Changes to the International Myeloma Working Group definitions of smouldering myeloma and myeloma have affected the risk stratification process for smouldering myeloma. It is unclear if the previous risk stratification approach remains valid. It is also unclear if earlier treatment will be of benefit to people with smouldering myeloma.

References

D'Arena G, Gobbi PG, Broglia C, et al. (2011). Pamidronate versus observation in asymptomatic myeloma: final results with long-term follow-up of a randomized study. Leukemia & lymphoma, 52, 771-775.

Hjorth M, Hellquist L, Holmberg E, et al. (1993). Initial versus deferred melphalan-prednisone therapy for asymptomatic multiple myeloma stage I--a randomized study. Myeloma Group of Western Sweden. European.journal of haematology., 50, 95-102.

Mateos MV, Hernandez MT, Giraldo P, et al. (2013). Lenalidomide plus dexamethasone for high-risk smoldering multiple myeloma. New England Journal of Medicine, 369, 438-447.

Musto P, Petrucci MT, Bringhen S, et al. (2008). A multicenter, randomized clinical trial comparing zoledronic acid versus observation in patients with asymptomatic myeloma.[Erratum appears in Cancer. 2008 Nov 15;113(10):2835]. Cancer, 113, 1588-1595.

Musto P, Petrucci MT, Bringhen S, et al. (2007). Final Analysis of a Multicenter, Randomised Study Comparing Zoledronate vs Observation in Patients with Asymptomatic Myeloma. Blood, 110, 164A.

Riccardi A, Mora O, Tinelli C, et al. (2000). Long-term survival of stage I multiple myeloma given chemotherapy just after diagnosis or at progression of the disease: a multicentre randomized study. British Journal of Cancer, 82, 1254-1260.

Witzig TE, Laumann KM, Lacy MQ, et al. (2013). A phase III randomized trial of thalidomide plus zoledronic acid versus zoledronic acid alone in patients with asymptomatic multiple myeloma. Leukemia, 27, 220-225.

5 Service organisation

The myeloma journey is complex, with many complications and side effects and it is not uncommon for a patient to be seen by many different specialists. Some of the services required by patients with myeloma include radiological imaging, radiotherapy, renal and orthopaedic support, services for management of spinal disease, clinical trials, haemato-oncology pharmacy, chemotherapy trained nurses, transplant services, and supportive and palliative care teams. These services are normally co-ordinated by a haematologist with a special interest in myeloma working as a member of the appropriate MDT.

There is variation across the UK in terms of which specialist services are provided and also where these services are provided (locally or regionally). This may mean that patients have to travel long distances to access specialised treatments which can be problematic and inconvenient. The optimal configuration of local and regional services to provide best care for patients with myeloma is uncertain.

Clinical question: What is the optimal configuration of local and regional haematology services for management of myeloma (including access to specialised radiological imaging, radiotherapy services, the management of renal disease, spinal disease and bone disease, clinical trials and supportive & palliative care)?

Clinical evidence (see also Appendix G)

No studies were identified in the literature that examined the configuration of local and regional haematology services for management of myeloma.

Cost effectiveness evidence

A literature review of published cost-effectiveness analyses did not identify any relevant papers for this topic. Whilst there were potential cost implications of making recommendations in this area, other questions in the guideline were agreed as higher priorities for economic evaluation. Consequently no further economic modelling was undertaken for this question.

Recommendations	**For guidance on the facilities needed to provide intensive inpatient chemotherapy and transplants for adults aged 18 and over with myeloma, and the structure and function of multidisciplinary teams (MDTs), see the NICE cancer service guidance on improving outcomes in haematological cancers.**
	
	For guidance on service organisation for people younger than 18, see the NICE cancer service guidance on improving outcomes in children and young people with cancer.
	
	Each hospital treating myeloma in people aged 18 and over who are not receiving intensive inpatient chemotherapy or a transplant should provide local access to:
	• **an MDT specialising in myeloma**
	• **supportive and palliative care, including:**
	  ○ **psychological support services**
	  ○ **a 24-hour acute oncology and/or haematology helpline**
	  ○ **physiotherapy**
	  ○ **occupational therapy**
	  ○ **dietetics**
	  ○ **medical social services**

	○ **critical care** • **clinical trials via the myeloma MDT** • **dental services.** **Each hospital treating myeloma in people aged 18 and over should provide regional access through its network to:** • **facilities for intensive inpatient chemotherapy or transplantation** • **renal support** • **spinal disease management** • **specialised pain management** • **therapeutic apheresis** • **radiotherapy** • **restorative dentistry and oral surgery** • **clinical trials, in particular early phase trials.**
Relative value placed on the outcomes considered	The Guideline Committee considered patient-reported outcomes, travel times, health-related quality of life, overall survival and progression free survival to be the most relevant outcomes to define the optimal configuration of local and regional haematology services for management of myeloma. No clinical evidence was identified for any of the outcomes.
Quality of the evidence	No evidence was identified that examined the configuration of local and regional haematological services for the management of myeloma. Therefore the Guideline Committee relied on clinical and patient experience and consensus to make recommendations.
Trade off between clinical benefits and harms	The Guideline Committee concluded that optimising service organisation would ensure equitable access to the full range of services needed to improve patient care, patient experience, clinical outcomes and safety for patients with myeloma. Also the recommendation to provide regional access to more specialist services through a network should help to prevent a fragmented service leading to better patient care. The Guideline Committee noted that a harm could be the potential for de-skilling at local hospitals where services had been recommended at the regional level. However the Guideline Committee agreed that this was unlikely to happen and therefore the benefits outweighed the potential harm.
Trade-off between net health benefits and resource use	The Guideline Committee noted that no relevant published economic evaluations had been identified and no additional economic analysis had been undertaken in this area. The Guideline Committee noted that most local hospitals treating patients with myeloma are already providing the services that have been recommended. However there was the potential for some additional cost for those local hospitals that do not currently provide the full range of services. The Guideline Committee agreed that improvements in patient care would outweigh these potential additional costs.
Other	The Guideline Committee recognised that patients receiving intensive in-patient chemotherapy or transplant are covered by the NICE guidance on Improving Outcomes in Haematological Cancers and therefore cross referenced this guidance. However the Guideline Committee noted that services for patients not

	receiving intensive in-patient chemotherapy or transplant were not covered by this guidance and therefore separate recommendations were needed for this patient group. The Guideline Committee noted that the provision of services for people with myeloma was currently variable and needed to be standardised. When making their recommendations, the Guideline Committee noted that some services would need to be provided at the local hospital level (because these services would be needed frequently by patients and would involve non-complex management) and others would be best provided at the regional level (because they are more specialist/complex and affect fewer patients). Based on their clinical and patient experience, the Guideline Committee recommended the configuration of services that would help to improve quality of care and improve outcomes. They noted that similar service configurations in other disease areas have resulted in improvements and considered this was likely to be the case for patients with myeloma. The Guideline Committee stressed that the services they have highlighted exist already but there would need to be a change in practice to define the pathways that myeloma patients should follow.

6 Managing newly diagnosed myeloma

6.1 First-line treatment

NICE has developed a suite of technology appraisal guidance on myeloma. It has not been possible to develop recommendations on primary disease treatment, salvage therapy for relapsed myeloma and consolidation/maintenance therapy after primary management in this guideline due to published technology appraisals or those in development.

There is no significant new evidence that would lead to a change in the existing recommendations in the published appraisals, and following consultation with relevant stakeholders, it was decided that these appraisals should be moved to the static list, thus preserving the funding direction associated with any positive recommendations. It is therefore possible for these recommendations to be incorporated into any future clinical guideline, but they cannot be updated and replaced at this time.

Recommendations in this guideline will complement the existing technology appraisals.

For more information on the relationship between the technology appraisal and clinical guidelines programmes please see Updating technology appraisals in the context of clinical guidelines.

For guidance on the use of bortezomib for induction therapy, see Bortezomib for induction therapy in multiple myeloma before high-dose chemotherapy and autologous stem cell transplantation (NICE technology appraisal guidance 311).

Recommendations	**Thalidomide in combination with an alkylating agent and a corticosteroid is recommended as an option for the first-line treatment of multiple myeloma in people for whom high-dose chemotherapy with stem cell transplantation is considered inappropriate. [This recommendation is from Bortezomib and thalidomide for the first-line treatment of multiple myeloma (NICE technology appraisal guidance 228).]** **Bortezomib in combination with an alkylating agent and a corticosteroid is recommended as an option for the first-line treatment of multiple myeloma if:** • **high-dose chemotherapy with stem cell transplantation is considered inappropriate and** • **the person is unable to tolerate or has contraindications to thalidomide [This recommendation is from Bortezomib and thalidomide for the first-line treatment of multiple myeloma (NICE technology appraisal guidance 228).]**
	These recommendations are from Bortezomib and thalidomide for the first-line treatment of multiple myeloma (NICE technology appraisal guidance 228). They were formulated by the technology appraisal and not by the guideline developers. They have been incorporated into this guideline in line with NICE procedures for developing clinical guidelines, and the evidence to support these recommendations can be found at www.nice.org.uk/TA228.

6.1.1 First autologous stem cell transplantation

Autologous stem cell transplantation (ASCT) is a medical procedure in which blood-forming stem cells are removed from the patient prior to intense chemotherapy and one or two days later given back to the same patient. The chemotherapy is aimed at killing myeloma cells but

also affects normal blood-forming cells that are needed to fight infections, transport oxygen and control bleeding. By giving the patient back his or her own blood-forming cells, the recovery from the chemotherapy is notably faster and more predictable compared to allogeneic transplantation.

ASCT has become the first line standard of care in those myeloma patients deemed biologically fit enough for this option mainly because of the low transplant-related mortality and prolongation of event-free survival, resulting in improved quality of life. Many factors must be considered to determine whether a patient is a candidate for ASCT including how the myeloma responded to prior treatment, patient age and general physical condition, and myeloma related renal failure and need for dialysis.

In addition, recent studies have shown that ASCT may be possible in patients who have reduced kidney function or kidney failure, with proper precautions and chemotherapy dose modification.

At the moment there is no clear consensus on what makes a patient a suitable candidate for ASCT and different centres use different criteria.

Clinical question: Which patients with newly diagnosed myeloma should be considered for autologous stem cell transplantation?

Clinical evidence (see also Appendix G)

See Table 22.

Age

Overall survival

Low quality evidence, from an individual patient meta-analysis (Levy et al, 2005) of three randomised trials (Attal et al, 1996; Fermand et al, 1998 and Fermand et al 1999; N=575), suggests that the effectiveness of high dose therapy with autologous stem cell transplant (HDT-ASCT) compared to standard dose treatment (SDT) is similar in younger and older age groups. There was no significant interaction between age (< 60 years versus 60 to 65 years) and the relative effectiveness of HDT-ASCT and SDT (P=0.96). For patients aged 60 to 65 years the hazard ratio for all cause mortality for HDT-ASCT versus SDT was 0.91 (95% C.I. 0.63 to 1.31; where HR < 1 favours HDT-ASCT), for patients younger than 60 years the hazard ratio was 0.90 (95% C.I. 0.72 to 1.12; where HR < 1 favours HDT-ASCT).

Seven randomised trials (low quality evidence) looked at age as a prognostic factor for overall survival but only two of these trials found age (Bladé et al., 1996 and Sonneveld et al., 2007) to be an independent prognostic factor. In Bladé et al (1996) the 56 to 70 year old age group were at higher risk of all cause mortality compared to those younger than 56 years: HR 1.87 [95%C.I. 1.12 to 3.19]. In Sonneveld et al (2007), each additional year in age was associated with an increased risk of overall mortality: HR 1.04 [95%C.I. 1.02 to 1.07].

Progression free survival

Moderate quality evidence from nine randomized trials including 2474 patients, suggests progression free survival is better with HDT-ASCT, regardless of the age entry criteria used in the trial. For HDT-ASCT versus SDT, the HR for disease progression was 0.78 (95%C.I. 0.71 to 0.86; where HR <1 favours HDT-SCT). In only one of the nine trials was progression free survival significantly worse with autologous stem cell transplant (Facon et al, 2007), this was a trial in older patients (aged 65 to 75 years) comparing reduced intensity autologous stem cell transplantation with melphalan, prednisolone and thalidomide.

TWiSTT

Moderate quality evidence from two randomized trials (Fermand et al 1998, 2005) including 375 patients suggests that TWiSTT is 6.93 months longer (95%C.I. 1.61 to 12.26 months longer) with HDT-ASCT than with standard dose chemotherapy, regardless of the age entry criteria used in the trial.

Treatment related mortality

Low quality evidence from six randomized trials including 1588 patients suggests that the risk of treatment related mortality is higher with HDT-ASCT than with standard dose therapy, RR 2.00 [95%C.I. 1.25 to 3.19] where RR <1.0 favours HDT-ASCT. When grouping the trials by their age entry criteria, the highest relative risks of treatment related mortality were seen in trials that included patients aged 70 years or less, however the absolute risk of treatment related mortality with HDT-ASCT in this subgroup was around 4% - lower than the 8% to 10% seen in trials restricted to under 65s or under 55s respectively.

Treatment related morbidity

In patients randomized to receive transplantation (Attal et al, 1996; low quality evidence) the completion of allocated treatment was related to age, with older patients less likely to undergo transplantation. 12 of 67 patients (18%) aged 60 or less did not undergo transplantation compared to 14 of 33 patients (42%) aged 60-65 years (P=0.01).

Fragility/weakness

Overall survival

Moderate quality evidence suggested a difference in the effectiveness of HDT-ASCT versus standard dose therapy (SDT) according to the trials' performance status (PS) entry criteria (test for subgroup differences, P=0.01). For trials restricted to patients with WHO PS 0 to 2 there was uncertainty about the relative effectiveness of HDT-ASCT and SDT in terms of overall survival (HR = 1.06; 95% C.I. 0.92 to 1.23; HR <1 favours HDT-ASCT). For trials that did not state any PS entry criteria, overall survival was significantly better with HDT-ASCT than SDT (HR = 0.80; 95% C.I. 0.68 to 0.95; HR <1 favours HDT-ASCT). It was unclear, however, what the actual performance status was of the patients in trials not specifying performance status entry criteria.

Disease progression

Moderate quality evidence from nine randomized trials including 2474 patients, suggests a difference in the relative effectiveness of HDT-ASCT and SDT in terms of disease progression according to the performance status entry criteria used in the trial (test for subgroup differences, P<0.0001). For trials restricted to patients with WHO PS 0 to 2 there was uncertainty about the relative effectiveness of HDT-ASCT and SDT in terms of disease progression (HR = 0.93; 95% C.I. 0.82 to 1.05; HR <1 favours HDT-ASCT). For trials that did not state any PS entry criteria, progression free survival was significantly better with HDT-ASCT than SDT (HR = 0.63; 95% C.I. 0.55 to 0.72; HR <1 favours HDT-ASCT). It was unclear, however, what the actual performance status was of the patients in trials not specifying performance status entry criteria.

In only one of these nine trials was progression free survival significantly worse with autologous stem cell transplant (Facon et al, 2007), this was a trial in older patients (aged 65 to 75 years) comparing reduced intensity autologous stem cell transplantation with melphalan, prednisolone and thalidomide. The inclusion of this trial in the WHO PS 0-2 subgroup accounts for the subgroup differences.

Comorbidities (charlson score, ACE-27, FACT-BMT)

No evidence was identified about the influence of comorbidities on the relative effectiveness of high dose therapy or conventional dose therapy.

Renal impairment

Overall survival

Moderate quality evidence, from an individual patient meta-analysis (Levy et al, 2005) of three randomised trials (Attal et al, 1996; Fermand et al, 1998 and Fermand et al 1999; N=575), suggests that the effectiveness of high dose therapy with autologous stem cell transplant (HDT) compared to standard dose treatment (SDT) is similar in high and low creatinine groups. There was no significant interaction between creatinine level (< 120 µmol/L versus ≥ 120 µmol/L) and the relative effectiveness of high dose therapy with autologous stem cell transplant (HDT) and conventional treatment (P=0.72). For patients with creatinine level < 120 µmol/L the hazard ratio for all cause mortality for HDT versus conventional treatment was 0.86 (95% C.I. 0.69 to 1.08; where HR < 1 favours HDT), for patients creatinine level ≥ 120 µmol/L the hazard ratio was 0.94 (95% C.I. 0.65 to 1.12; where HR < 1 favours HDT).

Three randomised trials (low quality evidence) looked at creatinine as a prognostic factor for overall survival and in two of these trials (Barlogie et al 2006 and Child et al 2003) creatinine level was an independent prognostic factor for overall survival.

Disease progression

Two trials (Barlogie et al 2006 and Child et al 2003) looked at creatinine level as a prognostic factor for disease progression and in one of these trials (Child et al 2003) it was an independent prognostic factor for disease progression.

Genetic abnormalities

One trial (Barlogie et al, 2006) considered deletion of chromosome 13 on FISH as a prognostic factor. FISH del(13) was an independent prognostic factor for both overall survival and disease progression free survival. Compared with others, patients with FISH del(13) had an increased risk of all cause mortality (HR 1.96; 95%C.I. 1.30 to 2.94) and of disease progression (HR 1.48; 95%CI 1.03 to 2.12). No evidence was presented of the relative effectiveness of HDT-ASCT versus SDT within the subgroup of patients with FISH del(13).

Response depth

In Child (2003) the depth of response was associated with overall survival in the HDT-ASCT group – for minimal response median survival was 25.6 months (95% CI 7.0 to 31.3 months), for partial response median survival was 39.8 months (95% CI 33.8 to 61.4 months) and for complete response median survival was 88.6 months (lower limit of 95% CI 61.4 months).

Table 22: GRADE profile: Which patients with newly diagnosed myeloma should be considered for autologous stem cell transplantation (high dose therapy with autologous stem cell transplant versus standard dose therapy)?

Quality assessment							No of patients		Effect		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	High dose therapy with AutoSCT	Standard Chemo-therapy	Relative (95% CI)	Absolute	Quality
Death from any cause (age < 60 years) (follow-up median 8.67 years)											
3[1]	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	no serious imprecision	none	154/212 (72.6%)	161/215 (74.9%)	HR 0.896 (0.717 to 1.121)	-	MODERATE
Death from any cause (age 60 to 65 years) (follow-up median 8.67 years)											
3[1]	randomised trials	serious[2]	serious[3]	no serious indirectness	no serious imprecision	none	57/73 (78.1%)	63/75 (84%)	HR 0.906 (0.626 to 1.311)	-	LOW
Death from any cause (performance status not specified) (follow-up median 3.1 to 10 years)											
5[4]	randomised trials	no serious risk of bias	serious[5]	serious[6]	no serious imprecision	none	261/533 (49%)	300/528 (56.8%)	HR 0.80 (0.68 to 0.95)	-	LOW
Death from any cause (performance status 0 to 2) (follow-up median 4.7 to 7.7 years)											
4[7]	randomised trials	serious[5]	no serious inconsistency	no serious indirectness	no serious imprecision	none	374/623 (60%)	353/611 (57.8%)	HR 0.94 (0.84 to 1.05)	-	MODERATE
Death from any cause (creatinine < 120 µmol/L) (follow-up median 8.67 years)											
3[1]	randomised	serious[8]	no serious	no serious	no serious	none	154/217	167/226	HR	-	MODERATE

Quality assessment							No of patients		Effect		Quality
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**High dose therapy with AutoSCT**	**Standard Chemo-therapy**	**Relative (95% CI)**	**Absolute**	
	trials		inconsistency	indirectness	imprecision		(71%)	(73.9%)	0.864 (0.693 to 1.077)		
Death from any cause (creatinine ≥ 120 µmol/L) (follow-up median 8.67 years)											
3[1]	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[3]	none	57/68 (83.8%)	57/64 (89.1%)	HR 0.935 (0.645 to 1.355)	-	
Progression free survival (follow-up median 3.1 to 10 years)											
9[9]	randomised trials	serious[5]	no serious inconsistency	no serious indirectness	no serious imprecision	none	?/1223	?/1194	HR 0.78 (0.71 to 0.86)	-	MODERATE
TWiSTT (follow-up median 4.8 to 10 years; Better indicated by higher values)											
2[10]	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	no serious imprecision	none	185	190	-	MD 6.93 months longer (1.61 to 12.26 longer)	MODERATE
Treatment related mortality (follow-up median 3.1 to 10 years)											
6[11]	randomised trials	serious[5]	no serious inconsistency	no serious indirectness	serious[3]	none	50/796 (6.3%)	25/792 (3.2%)	RR 2.00 (1.25 to 3.19)	32 more per 1000 (from 8 more to	LOW

Quality assessment							No of patients		Effect		
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**High dose therapy with AutoSCT**	**Standard Chemo-therapy**	**Relative (95% CI)**	**Absolute**	**Quality**
										69 more)	
Health related quality of life - not reported											
0	-	-	-	-	-	none	-	-	-	-	
Treatment related morbidity - not reported											
0	-	-	-	-	-	none	-	-	-	-	
Patient acceptability - not reported											
0	-	-	-	-	-	none	-	-	-	-	

1 Attal (1996), Fermand (1998), Fermand (2005) - IPD meta analysis by Levy (2005); 2 Unclear random sequence generation and blinding in all studies; 3 Low number of events; 4 Attal (1996), Child (2003), Fermand (1998), Fermand (2005) and Palumbo (2004); 5 Unclear random sequence generation and blinding in most studies; 6 Only Child (2003) reported the actual performance status of included patients; 7 Barlogie (2006), Blade (2005), Facon (2007) and Sonneveld (2007); 8 No explanation was provided; 9 Attal (1996), Barlogie (2006), Blade (2005), Child (2003), Facon (2007), Fermand (1998), Fermand (2005), Palumbo (2004) and Sonneveld (2007); 10 Fermand (1998), Fermand (2005); 11 Attal (1996), Barlogie (2006), Fermand (1998), Fermand (2005), Palumbo (2004) and Sonneveld (2007)

Cost effectiveness evidence

The following databases were searched for economic evidence relevant to the clinical question: MEDLINE, EMBASE, COCHRANE, NHS EED. Studies conducted from any OECD countries were considered (Guidelines Manual 2014).

463 possibly relevant papers were identified. Of these, 11 full papers relating to this topic were obtained for appraisal. Three papers were not relevant to the clinical question, one only considered costs and four did not report quality of life based outcomes. Therefore three studies (Gulbrandsen et al 2001, Van Agthoven et al 2004, Corso et al 2013) were included in the current review of published economic evidence for this topic (Table 23).

Gulbrandsen et al (2001) considered the cost effectiveness of high dose chemotherapy in addition to autologous stem cell transplant versus high dose chemotherapy alone in patients under 60 years of age with newly diagnosed, symptomatic myeloma. The study reported the results in terms of cost per Quality Adjusted Life Year (QALY) gained and considered a Norwegian societal perspective. Gulbrandsen et al found the transplant strategy to be both more costly and more effective estimating a cost per QALY of $27,000. This ranged from $6,800 to $40,000 per QALY during sensitivity analysis. Gulbrandsen et al (2001) had limited exploration of uncertainty around the parameters and results and did not present a probabilistic sensitivity analysis.

Corso et al (2013) considered the cost effectiveness of high dose chemotherapy with autologous stem cell support versus high dose chemotherapy in previously untreated myeloma patients. The study reported results in terms of cost per QALY gained and considered an Italian health payer perspective. The transplant strategy was found to be both more expensive and more effective leading to a cost per QALY of €44,454. There was a lack of transparency in the Corso study around their elicitation of key parameters (in particular utility weights) and the distributions used for parameters in their probabilistic sensitivity analysis. Deterministic sensitivity analyses were not presented.

Van Agthoven et al (2004) considered the cost effectiveness of intensive chemotherapy with stem cell transplant versus intensive chemotherapy alone in patients ≤65 years of age with previously untreated stage II or III A/B myeloma. The study found the transplant strategy to be both more costly and less effective. Van Agthoven presented limited exploration of uncertainty around their estimate making it difficult to consider the robustness of these conclusions. The study was therefore deemed to have potentially serious limitations.

Given the methodological issues discussed above all studies were considered to have potentially serious limitations. All studies were considered only partially applicable to the decision problem. This is because all studies took a perspective other than a NHS+PSS one. Discounting of costs and health outcomes was also inconsistent, with that recommended by NICE. Only one study (Van Agthoven et al, 2004) elicited changes in ‘Health Related Quality of Life’ from a representative sample of the general public.

Despite all three studies considering similar interventions and comparators it is difficult to meaningfully compare results given the differing range of perspectives and time horizons considered taken. All studies though reported significantly higher costs for the transplant strategy than for the non-transplant strategy. The incremental QALYs between the transplant and non-transplant strategies differed widely across all studies ranging from -0.14 to 1.73 QALYs.

1 **Table 23: Modified GRADE profile: included economic studies**

Study	Population	Comparators	Costs	Effects	Incr costs	Incr effects	ICER	Uncertainty	Applicability	Limitations
Corso et al (2013) Italy	Previously untreated myeloma patients	High dose chemotherapy-Intermittent course of melphalan and prednisone.	€23,825	3.02	Reference			Probabilistic sensitivity analyses: The transplant strategy had an 80% chance of being cost effective at a WTP of €60,000 and a 90% probability at a WTP of €75,000.	Partially Applicable	Methods for identifying key inputs (utilities etc) not adequately described. Lack of transparency around probabilistic sensitivity analysis.
		High dose melphalan + autologous stem cell support	€102,373	4.75	€78548	1.73 QALYs	€44,454 per QALY			
	Comments:									
Gulbrandsen et al (2001) Norway	Patients under 60 years of age with newly diagnosed, symptomatic myeloma.	High dose chemotherapy-Intermittent course of melphalan and prednisone.	$9,500	Not reported in disaggregated form	Reference			One-way sensitivity analyses A range of sensitivity analyses were conducted analyses with the cost per QALY ranging from $6,800 (survival pf transplant group increased by 0.5 years) to $40,000 per QALY (survival pf transplant group decreased by 0.5 years)	Partially applicable	Probabilistic sensitivity analysis not performed. Health outcomes not discounted
		High dose melphalan + autologous stem cell support	$34,000	Not reported in disaggregated form	$24,500	1.2 QALYs	$27,000 per QALY			
	Comments:									

Study	Population	Comparators	Costs	Effects	Incr costs	Incr effects	ICER	Uncertainty	Applicability	Limitations
van Agthoven et al (2004) Netherlands	Patients with previously untreated stage II or Stage III A/B myeloma ≤65 years of age.	Intensive chemotherapy with melphalan	€67,563	2.46QALYs	Reference			Probabilistic sensitivity analyses: Varying the hospital days by±34% varied total costs by ±11%	Partially Applicable	Limited exploration of uncertainty around model parameters.
		Intensive chemotherapy with melphalan followed by myeloablative therapy with autologous stem-cell rescue.	€80,630	2.32QALYs	€13,067	-0.14	Transplant strategy dominated			
	Comments:									

1

2

3

Recommendations	**Consider using frailty and performance status measures that include comorbidities to assess the suitability of people with myeloma for first autologous stem cell transplant.** **Do not use age or the level of renal impairment alone to assess the suitability of people with myeloma for first autologous stem cell transplant.**
Relative value placed on the outcomes considered	The Guideline Committee considered the outcomes of overall survival, progression-free survival, health-related quality of live, treatment related mortality and morbidity, patient/carer/family acceptability, later effects and TWIST to be the most relevant in determining whether autologous stem cell transplant was effective in specific subgroups of patients. Of these, evidence was identified for all outcomes except health-related quality of life, later effects and patient/carer/family acceptability. When drafting the recommendations the Guideline Committee considered overall survival and progression-free survival to be the most important.
Quality of the evidence	The quality of the evidence was assessed using GRADE methodology and appropriate NICE Checklists. Using these methods it was determined that the quality of the evidence was low to moderate. As a result the Guideline Committee made 'consider' recommendations rather than 'offer'. The Guideline Committee noted that the included papers were relatively old (1996 – 2007) with out-dated comparative treatments and so the Guideline Committee felt it was unclear which patient groups would benefit from autologous stem cell transplant in comparison to newer (and more effective) treatments. Further concerns with the evidence reported by the Guideline Committee were an inconsistency between the studies and also the majority of the studies have self-selected a group of patients who are suitable for autologous stem cell transplant.
Trade off between clinical benefits and harms	The Guideline Committee discussed age and it was felt that in current practice age was being used to determine a person's suitability for autologous stem cell transplant. The Guideline Committee agreed that this was inappropriate as there may be older patients who are fitter than some younger frail patients with co morbidities. Furthermore the evidence from a meta-analysis of 3 RCTs suggested that the effectiveness of autologous stem cell transplant compared to standard dose treatment was similar in younger and older age groups. Therefore the Guideline Committee recommended that age alone is not used to determine someone's suitability for autologous stem cell transplant. Historically renal impairment has been regarded as a relative contraindication to autologous stem cell transplant. However the Guideline Committee noted, based on the evidence, that autologous stem cell transplant is well tolerated and effective in people with myeloma who have renal impairment. They therefore recommended that level of renal impairment should not be used to determine someone's suitability for autologous stem cell transplant. The Guideline Committee made a recommendation to consider frailty and performance status related to comorbidities when proposing autologous stem cell transplantation, as it was agreed these are more appropriate measures of a person's suitability for transplant. The

	evidence also supported this. However fitness and frailty scores are not validated. Therefore the Guideline Committee also made a research recommendation for validating these scores. The Guideline Committee concluded that the benefits would be autologous stem cell transplant being offered to all suitable patients. The Guideline Committee acknowledged that as there is no robust measure of frailty, autologous stem cell transplant could potentially be performed in someone who was too frail and there may be an increase in morbidity. However, it was agreed that the biological status of the patient will be the deciding factor in whether they are suitable for transplant.
Trade-off between net health benefits and resource use	The economic evidence review identified three previous economic evaluations for this topic. All three studies compared autologous stem cell transplant to high dose chemotherapy. No studies were identified which took a NHS and PSS perspective. All studies were considered to have potentially serious limitations with their methodology. Despite similarities in their interventions and comparators it was difficult to meaningfully compare results, given the differing range of perspectives and time horizons considered. All studies though reported significantly higher costs for the transplant strategy than for the non-transplant strategy. The studies were inconsistent in their conclusions about the cost effectiveness of autologous stem cell transplant. All three studies used high dose chemotherapy with melphalan and prednisone in their comparator arm, an intervention which the Guideline Committee agreed is no longer standard care within the NHS, with more costly and effective chemotherapy now used. This limits the applicability of any conclusions from these studies. Given these serious applicability concerns the Guideline Committee were unable to draw any conclusions from the economic evidence and felt unable to use it in informing their recommendations. The Guideline Committee thought the recommendations may result in a small increase in the number of patients transplanted and so an increase in costs associated with this. However the Guideline Committee discussed that this may be offset by delayed relapse therapy. The Guideline Committee also discussed that some patients may become dialysis independent and this would result in a huge cost saving.
Other considerations	The Guideline Committee felt that there may a small change in practice as a result of the recommendations with some older patients or those with renal impairment having autologous stem cell transplant. There would be implications for renal units to provide support.

1

Research recommendation	**Observational studies should be carried out of the use and validation of existing fitness and frailty scoring systems in patients with myeloma being considered for autologous stem cell transplant. Outcomes of interest are transplant related mortality, length of stay, progression free survival, overall survival and quality of life.**
Why is this important?	Whilst there are various existing scoring systems for assessing generic fitness and frailty, none of these have been validated in patients undergoing ASCT. A validated system would allow a more rational and

Research recommendation	**Observational studies should be carried out of the use and validation of existing fitness and frailty scoring systems in patients with myeloma being considered for autologous stem cell transplant. Outcomes of interest are transplant related mortality, length of stay, progression free survival, overall survival and quality of life.**
	uniform approach to selecting patients who will benefit from this procedure.

6.1.21 Allogeneic stem cell transplantation

Allogeneic stem cell transplantation (AlloSCT) is a complex procedure involving administration of high-dose cytotoxic therapy (chemotherapy with or without radiotherapy) followed by transplant of peripheral blood or bone marrow stem cells (and rarely cord blood) from a sibling or unrelated donor. The risks are significantly higher than ASCT (which is more commonly performed in myeloma patients) and include a long-term tendency to infection and graft versus host disease (GvHD). These toxicities can significantly compromise both short term and long term quality of life and amount to a treatment related mortality risk of over 10-30% depending on the type of transplant and the status of the donor. However, GvHD is closely associated with a beneficial graft-versus-myeloma effect. AlloSCT has the potential of very-long term disease control but relapse occurs in a substantial proportion of patients.

Outcomes of AlloSCT have improved with the use of reduced intensity transplant (often combined with an ASCT). Despite this, a decision to proceed with AlloSCT is increasingly challenging with the advent of new therapies in myeloma, which, although not curative, may offer prolonged periods of disease control, and have significantly extended the life expectancy in patients with myeloma. Thus, amongst the modern treatment of myeloma, the optimum selection of patients for AlloSCT is unclear.

Clinical question: Which patients with myeloma should be considered for allogeneic stem cell transplantation?

Clinical evidence (see also Appendix G)

See Tables 24-29.

Patients with newly diagnosed myeloma

Very low to low quality evidence suggests that outcomes are better (OS and PFS or EFS are longer) following treatment with a tandem approach of autologous-allogeneic stem cell transplant compared to treatment with a tandem autologous-autologous stem cell transplant in newly diagnosed myeloma patients in the following subgroups: patients with del13 (Björkstrand et al., 2011; Gahrton et al., 2013), ISS stage 3 patients (Lokhorst et al., 2012) and chemosensitive patients (Rosinol et al., 2008). Allogeneic transplant was also found to be superior to any other treatment in patients with beta-2-microglobulin (B2M) greater than 3 (Lokhorst et al., 2012).

There was also evidence to the contrary from 2 studies which reported that outcomes were better with tandem autologous stem cell transplant compared to allogeneic transplant in newly diagnosed high risk myeloma patients (Garban et al., 2006; Krishnan et al., 2011). In addition, one study reported no difference in outcomes for the two treatment strategies in high risk patients (Bruno et al., 2007).

Conflicting results between the different studies are unlikely to be due to a true difference in the effect of allogeneic transplant in specific subgroups of patients but more than likely can be explained by differences between studies such as different patient selections, different conditioning regimens, and different GvHD prophylaxis regimen. Variation in the length of

follow-up employed in the different studies may also account for the differences in results. The studies of high risk myeloma patients all report better results (longer OS and PFS or EFS) with tandem autologous transplant compared to autologous-allogeneic transplant whereas studies of other population subgroups report better outcomes with autologous-allogeneic transplant. But these studies of high risk patients have shorter follow-up times (24-45 months) compared to the other studies (62-96 months).

No evidence was identified for the outcomes treatment related morbidity, health related quality of life, adverse events, patient/carer/family acceptability and PROMs.

Patients with relapsed myeloma

Low quality evidence from a retrospective analysis suggests that outcomes are worse following treatment with allogeneic stem cell transplant compared to a second autologous stem cell transplant in relapsed patients with Durie-Salmon stage III myeloma. Allotransplant was associated with a higher risk of relapse and treatment failure compared to autologous transplantation (Freytes et al., 2014). Evidence from the same study suggests that there is little difference in outcomes between related and unrelated donor allogeneic transplantation. The 3-year OS of patients who underwent transplant from related donors was 19% compared to 21% in patients whose donors were unrelated. Furthermore the TRM was also similar irrespective of donor type (Freytes et al., 2014).

Moderate quality evidence from studies of allogeneic transplant that reported predictive factors (high quality prognostic factor studies but downgraded as comparative studies are better for answering the review question) suggest that in relapsed myeloma patients undergoing allogeneic transplant B2 microglobulin < 3.3mg/L is predictive of lower NRM and longer PFS and OS (Efebera et al., 2010), a longer interval between auto and relapse is predictive of poorer OS (Patriarca et al., 2012), an interval of more than 1 year between the first and the salvage transplant is predictive of longer OS (Qazilbash et al., 2006), previous auto STC is predictive of lower NRM and longer PFS and OS (Efebera et al., 2010), refractory disease is predictive of worse OS and PFS (Shimoni et al., 2010), disease duration of >5 years is predictive of worse PFS (Shimoni et al., 2010) and SCT from female donor to male recipient is predictive of worse OS and PFS (Shimoni et al., 2010).

No evidence was identified for the outcomes treatment related morbidity, health related quality of life, adverse events, patient/carer/family acceptability and PROMs.

Table 24: GRADE profile: Which patients with myeloma should be considered for allogeneic stem cell transplantation (allo versus second auto in patients with newly diagnosed myeloma del13)?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Allo	second auto	Relative (95% CI)	Absolute	Quality
PFS at 96 months											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	Serious[1]	none	29	63	-	PFS at 96 months was 16% greater in the allo group compared to those in the second auto group	VERY LOW
OS at 96 months											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	Serious[1]	none	29	63	-	OS at 96 months was 16% greater in the allo group compared to those in the second auto group	VERY LOW

1 imprecision due to small sample size

Table 25: GRADE profile: Which patients with myeloma should be considered for allogeneic stem cell transplantation (allo versus second auto in patients with newly diagnosed myeloma who have high risk disease)?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Allo	second auto	Relative (95% CI)	Absolute	Quality
EFS											
2	observational studies	no serious	no serious inconsistency	no serious indirectness	no serious imprecision	none	123	265	-	One study: HR 0.52 (95%CI: 0.22-1.21).	LOW

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Allo	second auto	Relative (95% CI)	Absolute	Quality
		limitations								Second study: mean EFS was 3 months longer in patients in the second auto group compared to those in the allo group.	
OS											
2	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none	123	265	-	One study: HR 0.34 (95%CI: 0.10-1.18). Second study: mean OS was 12 months longer in patients in the second auto group compared to those in the allo group.	LOW
3 yr PFS											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	Serious[1]	none	29	31	-	3 yr PFS was 3% greater in patients in the second auto group compared to those in the allo group.	VERY LOW
3 yr OS											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	Serious[1]	none	29	31	-	3 yr OS was 3% greater in patients in the second auto group compared to those in the allo	VERY LOW

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Allo	second auto	Relative (95% CI)	Absolute	Quality
										group.	
3 yr TRM											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	Serious[1]	none	29	31	-	3 yr TRM was 7% lower in patients in the second auto group compared to those in the allo group.	VERY LOW
relapse/progression at 3 yrs											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	Serious[1]	none	29	31	-	Relapse/progression at 3yrs was 4% greater in patients in the second auto group compared to those in the allo group.	VERY LOW

1 *1 imprecision due to small sample size*

2 **Table 26: GRADE profile: Which patients with myeloma should be considered for allogeneic stem cell transplantation (allo versus**
3 **second auto in patients with newly diagnosed myeloma who have ISS stage III)?**

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Allo	second auto	Relative (95% CI)	Absolute	Quality
5yr PFS											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	Serious[1]	none	17	17	-	5 yr PFS was 28% greater in patients in the allo group	VERY LOW

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Allo	second auto	Relative (95% CI)	Absolute	Quality
										compared to those in the second auto group.	
5yr OS											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	Serious[1]	none	17	17	-	5 yr OS was 23% greater in patients in the allo group compared to those in the second auto group.	VERY LOW

1 *1 imprecision due to small sample size*

2 **Table 27: GRADE profile: Which patients with myeloma should be considered for allogeneic stem cell transplantation (allo versus**
3 **other treatment in patients with newly diagnosed myeloma who have β2M greater than 3mg/L)?**

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Allo	other treatment	Relative (95% CI)	Absolute	Quality
5yr PFS											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	Serious[1]	none	46	47	-	5 yr PFS was 20% greater in patients in the allo group compared to those in the second auto group.	VERY LOW
5yr OS											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	Serious[1]	none	46	47	-	5 yr OS was 17% greater in patients in the allo group	VERY LOW

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Allo	other treatment	Relative (95% CI)	Absolute	Quality
										compared to those in the second auto group.	

1 *1 imprecision due to small sample size*

2 **Table 28: GRADE profile: Which patients with myeloma should be considered for allogeneic stem cell transplantation (allo versus**
3 **second auto in patients with newly diagnosed myeloma who are chemosensitive)?**

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Allo	second auto	Relative (95% CI)	Absolute	Quality
CR rate											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	Serious[1]	none	25	85	-	CR was 29% greater in patients in the allo group compared to those in the second auto group.	VERY LOW
median PFS											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	Serious[1]	none	25	85	-	median PFS was 31 months in the second auto group and not reached in the allo group.	VERY LOW
median EFS											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	Serious[1]	none	25	85	-	median EFS was 6 months greater in patients in the allo group compared to those in the second	VERY LOW

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Allo	second auto	Relative (95% CI)	Absolute	Quality
										auto group.	
median OS											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	Serious[1]	none	25	85	-	median OS was 58 months in the second auto group and not reached in the allo group	VERY LOW
TRM											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	Serious[1]	none	25	85	-	TRM was 11% greater in patients in the allo group compared to those in the second auto group.	VERY LOW

1 *1 imprecision due to small sample size*
2

Table 29: GRADE profile: Which patients with myeloma should be considered for allogeneic stem cell transplantation (allo versus second auto in relapsed myeloma patients with Durie-Salmon stage III myeloma)?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Allo	second auto	Relative (95% CI)	Absolute	Quality
relapse											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none	152	137	-	Allotransplant was associated with a high risk of relapse compared to autotransplant (HR 3.05, 95% CI 2.20-4.22)	LOW

imprecision due to small sample size

Cost effectiveness evidence

A literature review of published cost-effectiveness analyses did not identify any relevant papers for this topic. Whilst there were potential cost implications of making recommendations in this area, other questions in the guideline were agreed as higher priorities for economic evaluation. Consequently no further economic modelling was undertaken for this question.

Recommendations	**When assessing whether people with myeloma are suitable for an allogeneic stem cell transplant, take into account:** • **whether the person has chemosensitive disease** • **how many previous lines of treatment they have had** • **whether a fully human leukocyte antigen (HLA) matched donor is available** • **how graft-versus-host disease (GvHD) and other complications will get worse with age** • **the risk of higher transplant-related mortality and morbidity, versus the potential for long-term disease-free survival** • **improving outcomes with other newer treatments** • **the person's understanding of the risks and benefits.**
Relative value placed on the outcomes considered	The Guideline Committee considered the outcomes of overall survival, progression free survival, health-related quality of life, treatment related mortality and morbidity, patient/carer/family acceptability, adverse events and patient reported outcome measures to be the most relevant in determining whether allogeneic stem cell transplant was effective in specific subgroups of patients. No evidence was identified for the outcomes treatment related morbidity, health-related quality of life, adverse events, patient reported outcome measures and patient/carer/family acceptability. When drafting the recommendations the Guideline Committee considered overall survival and progression free survival to be the most important.
Quality of the evidence	The quality of the evidence was assessed using GRADE methodology and appropriate NICE checklists. Using these methods it was determined that the quality of the evidence was very low to low for all outcomes. The evidence was from observational studies with small sample sizes and the results were inconsistent between the studies. The reviewer highlighted to the Guideline Committee that the conflicting results between the studies were likely to be due to differences between the studies themselves such as different patient selections, conditioning regimens, GvHD prophylaxis and variation in length of follow up. The Guideline Committee noted that the papers in the evidence review did not include the full range of current therapies which is particularly relevant given the recent advances in treatment. The number of myeloma patients undergoing allogeneic transplant are small, with a minimum of 5 years of study required to obtain sufficient numbers and follow-up and hence by the end of the study period the drugs used are less clinically relevant.

	Given the limitations with the current evidence base, the Guideline Committee agreed to recommend further research into the benefits of allogeneic stem cell transplantation. The Guideline Committee noted that in order to get sufficient numbers, allogeneic transplantation would need to be included as an option within other mainstream trials.
Trade off between clinical benefits and harms	As the evidence was of weak quality and inconsistent, the Guideline Committee were unable to make recommendations for which patients should be considered for allogeneic stem cell transplant and so instead made recommendations on what factors to take into account when assessing the suitability of allogeneic stem cell transplantation. The factors to be considered were based on those reported in the evidence as well as the Guideline Committee's experience. The Guideline Committee concluded that the benefits would be that allogeneic transplant will not be offered to inappropriate patients and as such there would be an avoidance of early mortality and morbidity in these individuals. The Guideline Committee agreed that the recommendations would also result in the selection of the most appropriate patients for consideration of allogeneic transplant and the promotion of a good patient understanding of the issues. The Guideline Committee agreed that there would not be any harms as their recommendation is to make better informed decisions, thereby avoiding harm.
Trade-off between net health benefits and resource use	The Guideline Committee noted that no relevant published economic evaluations had been identified and no additional economic analysis had been undertaken in this area. The Guideline Committee concluded that there were unlikely to be a cost consequences associated with these recommendations because no interventions had actually been recommended - just what factors to consider when assessing suitability.
Other considerations	The Guideline Committee discussed possible change in practice and concluded that any change would be small. The recommendations would result in a more conservative approach to allogeneic transplant in that inappropriate patients would not be transplanted.

1

Research recommendation	**Research is needed into the effectiveness of combined autologous-allogeneic stem cell transplantation compared with autologous stem cell transplantation, plus consolidation and maintenance treatment in chemosensitive patients at first response or first relapse. Outcomes of interest are progression-free survival, overall survival, transplant-related mortality, quality of life, early and late toxicity including graft-versus-host-disease (GvHD) and resource use. This research should be included as an option in appropriate mainstream clinical trials for myeloma.**
Why is this important?	There are conflicting data from a small number of studies on long-term survival following auto/allo stem cell transplantation compared with autologous stem cell transplantation. These studies were performed before thalidomide, bortezomib and lenalidomide were used as myeloma treatments. These drugs produce better responses and also have the capacity to affect immunological responses after the transplant. Research is needed to see if there is a role for auto/allo stem cell transplant in the ongoing treatment of myeloma.

6.2 Primary plasma cell leukaemia

Plasma cell leukaemia (PCL) is an aggressive type of myeloma characterised by the presence of a large number of malignant plasma cells in the peripheral blood. It is further classified into primary PCL, when it occurs at diagnosis or secondary PCL when it develops as a terminal phase of relapsed refractory myeloma. Primary PCL runs a more aggressive course than myeloma with poor response to conventional chemotherapy and a significantly shorter lifespan with a median survival of only 7 months. In view of the rarity of primary PCL, no large scale clinical trials have been conducted and most information about its management comes from case reports or small series from retrospective studies. Consequently, the clinical approach to the management of patients with primary plasma cell leukaemia remains variable.

Clinical question: What are the most effective treatments for patients with primary plasma cell leukaemia?

Clinical evidence (see also Appendix G)

See Tables 30-38.

Overall survival and progression-free survival

Very low quality evidence from 7 observational studies reporting on overall survival (OS) and progression-free survival (PFS) in primary plasma cell leukemia (pPCL) following treatment with autologous transplant (Drake et al., 2010; Mahindra et al., 2012), allogeneic transplant (Mahindra et al., 2012; Landsburg et al, 2014), lenalidomide (Musto et al., 2014), bortezomib-based regimens (Katodritou et al., 2014), bortezomib/thalidomide/lenalidomide (Talamo et al., 2012) and total therapy protocol (Usmai et al., 2012) was identified. Median OS ranged from 18 to 28 months across the studies and OS at 3 years ranged from 39 to 65%. Median PFS ranged from 10 to 14.3 months across the studies and PFS at 3 years ranged from 20 to 34%.

Median OS was lowest at 18 months in patients (n=18) treated with bortezomib-based regimens (Katodritou et al., 2014). In a study of bortezomib, thalidomide or lenalidomide-based regimes (Talamo et al., 2012) median OS and PFS was 21 and 10 months respectively with treatment. However the sample size was small (n=12) and it is unclear how many pPCL patients were on each treatment. A study of 27 patients on total therapy protocols reported similar results with a median OS 22 months and median PFS 10 months (Usmani et al., 2012). There was heterogeneity in the treatment protocols but with successive TT protocols there was no advance in OS or PFS. A study exploring lenalidomide reported the greatest median OS of 28 months and PFS of 14 months (Musto et al., 2014). However this study of 23 patients has not been peer-reviewed (published as a letter to the editor) and the authors have conflicts of interest and so the validity of the data is questioned. OS and PFS in patients that had undergone transplant were investigated in 2 studies. Drake et al. (2010) examined autologous transplant in 272 patients and reported a median OS of 25.7 months and OS at 3 years was 39.5%. Median PFS was 14.3 months. Mahindra et al. (2012) examined both autologous and allogeneic transplant in 97 and 50 patients, respectively. OS at 3 years was 39% for allogeneic transplant and 64% for autologous transplant. PFS at 3 years was 20% for allogeneic transplant and 34% for autologous transplant. To what extent the OS and PFS associated with transplant is related to the treatment itself or to the patient selection for transplant is unclear as the studies are retrospective cohort studies probably have high patient selection bias in that transplanted patients are generally younger and with better performance status than non transplanted patients.

Overall survival was compared in transplanted (n=23: 21 auto, 2 allo) and non-transplanted (n=50) patients in one study (Pagano et al, 2011). Median overall survival was 29 months longer in transplanted patients compared to non-transplanted patients. In another study progression-free survival was compared in transplanted (n=9: 8 auto, 1 allo) and non-transplanted (n=14) patients (Musto et al, 2014). Progression free survival was 25 months longer in transplanted patients compared to non-transplanted patients.

Overall response rate

Very low quality evidence from 5 observational studies reporting on overall response rate (ORR) in pPCL following treatment with allogeneic transplant (Charbonnier et al., 2014; Landsburg et al, 2014), bortezomib (D'Arena et al., 2012; Katodritou et al., 2014; Pagano et al., 2011), thalidomide (Pagano et al., 2011), bortezomib+thalidomide (Pagano et al., 2011) and lenalidomide (Musto et al., 2014) was identified. ORR ranged from 45 to 89%.

ORR ranged from 71% to 88% in two observational studies of 24 patients that had undergone allogeneic transplant (Charbonnier et al., 2014; Landsburg et al, 2014). However Charbonnier et al. (2014) was published as a conference poster abstract and so full details of the study are outstanding and we await publication of the complete study to assess the study quality and validity. Bortezomib was associated with an ORR of 79% in a study of 29 patients (D'Arena et al., 2012) and 89% in a study of 18 patients (Katodritou et al., 2014). However bortezomib was administered in various combinations to different patients in both these studies. Bortezomib was also used in another study of 4 patients (Pagano et al., 2011) and here the ORR was lower at 50%. Pagano also assessed thalidomide (5 patients) and here the ORR was also low at 45%. But in patients that received both bortezomib and thalidomide (n=10) ORR was much higher at 80%. A study exploring lenalidomide reported an ORR of 74% (Mutso et al., 2014). However this study of 23 patients has not been peer-reviewed and the authors have conflicts of interest and so the validity of this data is questioned.

Adverse events

Very low quality evidence from 4 observational studies reporting on adverse events in plasma cell leukemia following treatment with allogeneic transplant (Charbonnier et al., 2014; Mahindra et al., 2012), bortezomib (D'Arena et al., 2012) and lenalidomide (Musto et al., 2014) was identified.

Graft-versus host disease (GvHD) was reported in patients receiving allogeneic transplant. The incidence of acute GvHD was 28% in a retrospective study of 50 patients (Mahindra et al., 2012), 29% in a retrospective series of 7 patients (Landsburg et al, 2014) and 35% in a prospective study of 17 patients (Charbonnier et al., 2014). The incidence of chronic GvHD was 26% in a retrospective study of 50 patients (Mahindra et al., 2012), 29% in a retrospective series of 7 patients (Landsburg et al, 2014) and 20% in a prospective study of 17 patients (Charbonnier et al., 2014). Treatment related mortality occurred in 2/7 (29%) of patients treated with allogeneic transplant in Landsburg et al (2014).

Various toxicities were reported in patients receiving chemotherapy regimes. In a study of 29 patients receiving bortezomib grade 3–4 haematological toxicities were reported in 20% of patients and grade 3–4 non-haematological toxicities were reported in 55% of patients (D'Arena et al., 2012). In a study of 23 patients receiving lenalidomide grade 3–4 haematological toxicities were reported in 48% of patients and grade 3–4 non-haematological toxicities were reported in 52% of patients (Musto et al., 2014).

Health-related quality of life

No evidence was found for this outcome.

Table 30: GRADE profile: What are the most effective treatments for patients with primary plasma cell leukaemia (autologous transplant)?

Quality assessment							Summary of findings		
No of studies	**Design**	**Limitations**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**No of patients**	**Effect**	**Quality**
overall survival									
2	observational studies	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	369	Median OS: 25.7 Months OS at 3 years 40-64%	VERY LOW
progression free survival									
2	observational studies	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	369	Median PFS: 14.3 Months PFS at 3 years 34%	VERY LOW
Overall response rate									
0									
Adverse events									
0									
HRQOL									
0									

1 retrospective case series

Table 31: GRADE profile: What are the most effective treatments for patients with primary plasma cell leukaemia (allogeneic transplant)?

Quality assessment							Summary of findings		
No of studies	**Design**	**Limitations**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**No of patients**	**Effect**	**Quality**
overall survival									
1	observational	serious[1]	no serious	no serious	no serious	none	50	OS at 3 years	VERY

Quality assessment							Summary of findings		
	studies		inconsistency	indirectness	imprecision			39%	LOW
progression free survival									
1	observational studies	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	50	PFS at 3 years 20%	VERY LOW
Overall response rate									
1	observational studies	Serious[2]	no serious inconsistency	no serious indirectness	no serious imprecision	none	17	ORR: 88%	VERY LOW
Adverse events									
2	observational studies	serious[1,2]	no serious inconsistency	no serious indirectness	no serious imprecision	none	67	Incidence of acute GvHD: 28-35% Incidence of chronic GvHD: 20-26%	VERY LOW
HRQOL									
0									

1 *1 retrospective case series; 2poster conference abstract*

2 **Table 32: GRADE profile: What are the most effective treatments for patients with primary plasma cell leukaemia (transplant versus**
3 **no transplant)?**

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	**Design**	**Limitations**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**no transplant**	**transplant**	**Relative (95% CI)**	**Absolute**	**Quality**
overall survival											
1	observational studies	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	50	23	-	Median overall survival was 29 months longer in transplanted	VERY LOW

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	no transplant	transplant	Relative (95% CI)	Absolute	Quality
										patients	
progression free survival											
1	observational studies	serious[2]	no serious inconsistency	no serious indirectness	no serious imprecision	none	14	9	-	Progression-free survival was 25 months longer in transplanted patients	LOW

1 *1 retrospective case series; 2 published as letter: not peer-reviewed. Conflicts of interest.*

2 **Table 33: GRADE profile: What are the most effective treatments for patients with primary plasma cell leukaemia (bortezomib)?**

Quality assessment							Summary of findings		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	No of patients	Effect	Quality
overall survival									
1	observational studies	serious[1]	serious inconsistency[2]	no serious indirectness	no serious imprecision	none	18	Median OS: 18 months	VERY LOW
progression free survival									
0									
Overall response rate									
3	observational studies	Serious[1]	serious inconsistency[2]	no serious indirectness	no serious imprecision	none	51	ORR: 50-89%	VERY LOW
Adverse events									
1	observational studies	serious[1]	serious inconsistency[2]	no serious indirectness	no serious imprecision	none	29	Grade3-4 hematological toxicities: 20% of patients Grade3-4 non-	VERY LOW

Quality assessment							Summary of findings		
								hematological toxicities: 55% of patients	
HRQOL									
0									

1 *1 retrospective case series; 2 not consistent treatment combinations*

2 **Table 34: GRADE profile: What are the most effective treatments for patients with primary plasma cell leukaemia (thalidomide)?**

Quality assessment							Summary of findings		
No of studies	**Design**	**Limitations**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**No of patients**	**Effect**	**Quality**
overall survival									
0									
progression free survival									
0									
Overall response rate									
1	observational studies	Serious[1]	serious inconsistency	no serious indirectness	no serious imprecision	none	5	ORR: 45%	VERY LOW
Adverse events									
0									
HRQOL									
0									

3 *1 retrospective case series*

4 **Table 35: GRADE profile: What are the most effective treatments for patients with primary plasma cell leukaemia (bortezomib plus**
5 **thalidomide)?**

Quality assessment							Summary of findings		
No of studies	**Design**	**Limitations**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**No of patients**	**Effect**	**Quality**
overall survival									

Quality assessment							Summary of findings		
0									
progression free survival									
0									
Overall response rate									
1	observational studies	Serious[1]	serious inconsistency	no serious indirectness	no serious imprecision	none	10	ORR: 80%	VERY LOW
Adverse events									
0									
HRQOL									
0									

1 *1 retrospective case series*

2 **Table 36: GRADE profile: What are the most effective treatments for patients with primary plasma cell leukaemia (bortezomib or**
3 **thalidomide or lenalidomide)?**

Quality assessment							Summary of findings		
No of studies	**Design**	**Limitations**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**No of patients**	**Effect**	**Quality**
overall survival									
1	observational studies	serious[1]	no serious inconsistency	no serious indirectness	serious imprecision[2]	none	12	Median OS: 21 months	VERY LOW
progression free survival									
1	observational studies	serious[1]	no serious inconsistency	no serious indirectness	serious imprecision[2]	none	12	Median PFS: 10 months	VERY LOW
Overall response rate									
0									
Adverse events									
0									
HRQOL									
0									

4 *1 retrospective case series; 2 small population and unclear how many patients in each regime*

Table 37: GRADE profile: What are the most effective treatments for patients with primary plasma cell leukaemia (lenalidomide)?

Quality assessment							Summary of findings		
No of studies	**Design**	**Limitations**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**No of patients**	**Effect**	**Quality**
overall survival									
1	observational studies	serious[1]	no serious inconsistency	no serious indirectness	No serious imprecision	none	23	Median OS: 28 months	VERY LOW
progression free survival									
1	observational studies	serious[1]	no serious inconsistency	no serious indirectness	No serious imprecision	none	23	Median PFS: 14 months	VERY LOW
Overall response rate									
1	observational studies	Serious[1]	serious inconsistency	no serious indirectness	no serious imprecision	none	23	ORR: 74%	VERY LOW
Adverse events									
1	observational studies	serious[1]	serious inconsistency[2]	no serious indirectness	no serious imprecision	none	23	Grade3-4 hematological toxicities: 48% of patients Grade3-4 non-hematological toxicities: 52% of patients	VERY LOW
HRQOL									
0									

1 published as letter: not peer-reviewed. Conflicts of interest.

Table 38: GRADE profile: What are the most effective treatments for patients with primary plasma cell leukaemia (total therapy protocol)?

Quality assessment							Summary of findings		
No of studies	**Design**	**Limitations**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**No of patients**	**Effect**	**Quality**
overall survival									

Quality assessment							Summary of findings		
1	observational studies	serious[1]	serious inconsistency[2]	no serious indirectness	no serious imprecision	none	27	Median OS: 22 Months	VERY LOW
progression free survival									
1	observational studies	serious[1]	serious inconsistency[2]	no serious indirectness	no serious imprecision	none	27	Median PFS: 10 Months	VERY LOW
Overall response rate									
0									
Adverse events									
0									
HRQOL									
0									

1 *1 retrospective case series; 2 not consistent treatment protocols across the population*

2

1 Cost effectiveness evidence

2 A literature review of published cost-effectiveness analyses did not identify any relevant
3 papers for this topic. Whilst there were potential cost implications of making
4 recommendations in this area, other questions in the guideline were agreed as higher
5 priorities for economic evaluation. Consequently no further economic modelling was
6 undertaken for this question.

7

Recommendations	**Consider bortezomib-based and/or lenalidomide-based combination induction chemotherapy for people with primary plasma cell leukaemia.** **Consider high-dose melphalan-based autologous stem cell transplantation for people with primary plasma cell leukaemia if they are suitable.**
Relative value placed on the outcomes considered	The Guideline Committee considered the outcomes of overall survival, progression-free survival, health-related quality of life and adverse events to be the most relevant in identifying the most effective treatments for patients with primary plasma cell leukaemia Of these, evidence was found for overall survival, progression-free survival and adverse events but no evidence was identified for health-related quality of life. When drafting the recommendations the Guideline Committee considered overall survival and progression-free survival to be the most important. The outcome overall response rate was also reported in the evidence. Even though this was not specified in the review question, the Guideline Committee agreed that it provided additional evidence on the effectiveness of the interventions and therefore used it when drafting their recommendations.
Quality of the evidence	The quality of the evidence was assessed using GRADE methodology and appropriate NICE checklists. Using these methods it was determined that the quality of the evidence for all interventions and all outcomes was very low. The data was mostly non-comparative and consisted of case series of small sample size (due to the rarity of primary plasma cell leukaemia). In addition the transplant studies had a patient selection bias. As a result the Guideline Committee made 'consider' recommendations rather than 'offer'. In addition, there was treatment heterogeneity of induction treatments in many of the studies. As such it was unclear which regimen was most effective and the Guideline Committee were only able to recommend bortezomib and/or lenalidomide-based treatments rather than specific regimens for induction therapy. As the Guideline Committee were unable to recommend specific interventions due to the heterogeneity of induction treatments used in the studies the Guideline Committee made a research recommendation to investigate the most effective induction therapy. The Guideline Committee also included maintenance and consolidation strategies in this research recommendation as no clinical evidence was identified for primary plasma cell leukaemia. Although the Guideline Committee made a research recommendation for an RCT they decided not to list particular interventions that should be included in such an RCT as the rapid evolution of emerging

	therapies would make the recommendation out of date quickly. The Guideline Committee also made a recommendation for further research into autologous and allogeneic transplantation in people with primary plasma cell leukaemia. The Guideline Committee agreed that although the evidence indicated that these were effective treatment options, there is a lack of data on long term outcomes such as survival benefit and toxicity. The Guideline Committee recommended a prospective observational registration study as it was agreed that an RCT was not feasible because of the small number of potentially eligible patients.
Trade off between clinical benefits and harms	The Guideline Committee made a recommendation on the use of autologous transplantation based on the available clinical evidence. There was limited, conflicting data about the effectiveness and safety of allogeneic transplantation in primary plasma cell leukaemia. The Guideline Committee considered that there could potentially be a subgroup of patients who might benefit from this treatment but were unable to define which patients would be included in this group based on the available evidence. Therefore no recommendations were made for the use of allogeneic transplant. The Guideline Committee concluded that effective treatment of primary plasma cell leukaemia would result in an increased consistency in the treatment of primary plasma cell leukaemia across different centres/trusts and improvements in progression free survival and overall survival. The Guideline Committee acknowledged that treatment related morbidity and mortality may increase, but it was agreed that the benefits outweighed the harms as the morbidity is usually manageable and mortality is less than 5%. Furthermore the Guideline Committee agreed that survival for people with primary plasma cell leukaemia would be poor without these treatments.
Trade-off between net health benefits and resource use	The Guideline Committee noted that no relevant published economic evaluations had been identified and no additional economic analysis had been undertaken in this area. The Guideline Committee thought the recommendations may result in some additional costs from the potential earlier use of bortezomib and lenalidomide. There may also be some savings through a reduction of in-patient costs by avoiding inappropriate regimens. The cost consequences of the transplant recommendations were also discussed amongst the Guideline Committee but it was agreed that this was probably already happening in most centres. Therefore overall it was felt that the recommendations would be cost neutral. Furthermore the Guideline Committee highlighted that primary plasma cell leukaemia only affects 30-50 patients per year and so the cost impact of any recommendations would be small.

1

Research recommendations	**A randomised controlled trial should be carried out, to investigate the most effective induction, consolidation and maintenance strategy in people with primary plasma cell leukaemia. Outcomes of interest are progression-free survival, overall survival and quality of life.**
Why is this important?	Primary plasma cell leukaemia is a rare condition and there are currently no comparative studies of different treatments.

2

Research recommendations	**A non-interventional, prospective observational registration study on the efficacy, safety and long term outcomes of both autologous and allogeneic stem cell transplantation in patients with primary plasma cell leukaemia. Outcomes of interest are progression-free survival, overall survival and quality of life.**
Why is this important?	Primary plasma cell leukaemia is a rare condition and there are currently no comparative studies of different treatments. However retrospective studies appear to show a benefit for transplantation over other treatments. The number of patients that would be available for this type of study would be insufficient for an RCT.

References

Attal M, Harousseau JL, Stoppa AM, et al. (1996). A prospective, randomized trial of autologous bone marrow transplantation and chemotherapy in multiple myeloma. Intergroupe Français du Myélome. New England Journal of Medicine, 335, 91-97.

Barlogie B, Zangari M, Bolejack V, et al. (2006). Superior 12-year survival after at least 4-year continuous remission with tandem transplantations for multiple myeloma. Clinical Lymphoma & Myeloma, 6, 469-474.

Bladé J, San-Miguel JF, Fontanillas M, et al. (1996). Survival of multiple myeloma patients who are potential candidates for early high-dose therapy intensification/ autotransplantation and who were conventionally treated. Journal of Clinical Oncology, 14, 2167-2173

Björkstrand B, Iacobelli S, Hegenbart U, et al. (2011) Tandem autologous/reduced-intensity conditioning allogeneic stem-cell transplantation versus autologous transplantation in myeloma: long-term follow-up. J Clin Oncol.;29(22):3016-22.

Bruno B, Rotta M, Patriarca F, et al. (2007) A comparison of allografting with autografting for newly diagnosed myeloma. New England Journal of Medicine, 356: 1110-1120.

Charbonnier A, Michalet M, Xhaard A, et al. (2014) Allogeneic Hematopoietic Stem Cell Transplantation (Allo-Hsct) for Primary Plasma Cell Leukemia (Ppcl): A Prospective Study of Ifm Group. Haematologica, 99: 165.

Child JA, Morgan GJ, Davies FE, et al. (2003). High-dose chemotherapy with hematopoietic stem-cell rescue for multiple myeloma. New England Journal of Medicine, 348, 1875-1883.

Corso A, Mangiacavalli S, Cocito F, et al. (2013) Long Term Evaluation of the Impact of Autologous Peripheral Blood Stem Cell Transplantation in Multiple Myeloma: A Cost-Effectiveness Analysis. PLoS ONE 8(9): e75047.

D'Arena G, Valentini CG, Pietrantuono G, et al. (2012) Frontline chemotherapy with bortezomib-containing combinations improves response rate and survival in primary plasma cell leukemia: a retrospective study from GIMEMA Multiple Myeloma Working Party. Annals of Oncology, 23: 1499-1502.

Drake MB, Iacobelli S, van Biesen A, et al. (2010) Primary plasma cell leukemia and autologous stem cell transplantation. Haematologica, 95: 804-809.

Efebera YA, Qureshi SR, Cole SM, et al. (2010) Reduced-intensity allogeneic hematopoietic stem cell transplantation for relapsed multiple myeloma. Biology of Blood & Marrow Transplantation, 16: 1122-1129.

Facon T, Mary JY, Hulin C, 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, 1209-1218.

Fermand JP, Katsahian S, Divine M, 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. Journal of Clinical Oncology, 23, 9227-9233.

Fermand JP, Ravaud P, Chevret S, et al. (1998). High-dose therapy and autologous peripheral blood stem cell transplantation in multiple myeloma: up-front or rescue treatment? Results of a multicenter sequential randomized clinical trial. Blood, 92, 3131-3136.

Freytes CO, Vesole DH, LeRademacher J, et al. (2014) Second transplants for multiple myeloma relapsing after a previous autotransplant-reduced-intensity allogeneic vs autologous transplantation. Bone Marrow Transplantation, 49: 416-421.

Gahrton G, Iacobelli S, Björkstrand B, et al. (2013) Autologous/reduced-intensity allogeneic stem cell transplantation vs autologous transplantation in multiple myeloma: long-term results of the EBMT-NMAM2000 study. Blood. 20;121(25):5055-63.

Garban F, Attal M, Michallet M, et al. (2006) Prospective comparison of autologous stem cell transplantation followed by dose-reduced allograft (IFM99-03 trial) with tandem autologous stem cell transplantation (IFM99-04 trial) in high-risk de novo multiple myeloma. Blood, 107: 3474-3480.

Gulbrandsen N, Wisløff F, Nord E, et al. (2001). 'Cost-utility analysis of high-dose melphalan with autologous blood stem cell support vs. melphalan plus prednisone in patients younger than 60 years with multiple myeloma.' European Journal of Haematology, 66, 328-336.

Katodritou E, Terpos E, Kelaidi C, et al. (2014) Treatment with bortezomib-based regimens improves overall response and predicts for survival in patients with primary or secondary plasma cell leukemia: Analysis of the Greek myeloma study group. American Journal of Hematology, 89: 145-150.

Krishnan A, Pasquini B, Logan B, et al. (2011) Autologous haemopoietic stem-cell transplantation followed by allogeneic or autologous haemopoietic stem-cell transplantation in patients with multiple myeloma (BMT CTN0102): a phase 3 biological assignment trial. Lancet Oncol.;12(13):1195-1203.

Landsburg DJ (2014). Melphalan/total body irradiation-conditioned myeloablative allogeneic hematopoietic cell transplantation for patients with primary plasma cell leukemia. Clinical Lymphoma, Myeloma and Leukemia, 14, e225-e228.

Levy V, Katsahian S, Fermand JP, et al. (2005). A meta-analysis on data from 575 patients with multiple myeloma randomly assigned to either high-dose therapy or conventional therapy. Medicine, 84, 250-260.

Lokhorst HM, van der Holt B, Cornelissen JJ, et al. (2012) Donor versus no-donor comparison of newly diagnosed myeloma patients included in the HOVON-50 multiple myeloma study. Blood; 119(26):6219-6225.

Mahindra A, Kalaycio ME, Vela-Ojeda J, et al. (2012) Hematopoietic cell transplantation for primary plasma cell leukemia: results from the Center for International Blood and Marrow Transplant Research. Leukemia, 26: 1091-1097.

Musto P, Simeon V, Martorelli MC, et al. (2014) Lenalidomide and low-dose dexamethasone for newly diagnosed primary plasma cell leukemia. Leukemia. 28(1), 222-225.

Pagano L, Valentini CG, De SV, et al. (2011) Primary plasma cell leukemia: a retrospective multicenter study of 73 patients. Annals of Oncology, 22: 1628-1635.

Palumbo A (2004). Intermediate-dose melphalan improves survival of myeloma patients aged 50 to 70: Results of a randomized controlled trial. Blood, 104, 3052-3057.

Patriarca F, Einsele H, Spina F, et al. (2012) Allogeneic stem cell transplantation in multiple myeloma relapsed after autograft: a multicenter retrospective study based on donor availability. Biology of Blood & Marrow Transplantation, 18: 617-626.

Qazilbash MH, Saliba R, De LM, et al. (2006) Second autologous or allogeneic transplantation after the failure of first autograft in patients with multiple myeloma. Cancer, 106: 1084-1089.

Rosinol L, Perez-Simon JA, Sureda A, et al (2008). A prospective PETHEMA study of tandem autologous transplantation versus autograft followed by reduced intensity conditioning allogeneic transplantation in newly diagnosed multiple myeloma. Blood; 112: 3591–3593.

Shimoni A, Hardan I, Ayuk F, et al. (2010) Allogenic hematopoietic stem-cell transplantation with reduced-intensity conditioning in patients with refractory and recurrent multiple myeloma: long-term follow-up. Cancer, 116: 3621-3630.

Sonneveld P, Holt B, Segeren CM, et al. (2007). Intermediate-dose melphalan compared with myeloablative treatment in multiple myeloma: long-term follow-up of the Dutch Cooperative Group HOVON 24 trial. Haematologica, 92, 928-935.

Talamo G, Dolloff NG, Sharma K, et al. (2012) Clinical features and outcomes of plasma cell leukemia: A single-institution experience in the era of novel agents. Rare Tumors, 4: 123-126.

Van Agthoven M, Segeren CM, Buijt I, et al. (2004). A cost-utility analysis comparing intensive chemotherapy alone to intensive chemotherapy followed by myeloablative chemotherapy with autologous stem-cell rescue in newly diagnosed patients with stage II/III multiple myeloma: a prospective randomised phase III study. European Journal of Cancer, 40: 1159-1169.

Usmani SZ, Nair B, Qu P, et al. (2012) Primary plasma cell leukemia: clinical and laboratory presentation, gene-expression profiling and clinical outcome with Total Therapy protocols. Leukemia, 26: 2398-2405.

7 Managing acute renal disease caused by myeloma

Myeloma causes renal impairment in a range of ways that include deposits of myeloma associated proteins in the kidney (cast nephropathy, amyloid or light chain deposition); high calcium levels; infection and drug associated toxicity. It is reported that up to 40% of myeloma patients will have a degree of renal impairment at diagnosis and up to 10% of patients will require renal replacement therapy (dialysis).

Renal impairment can occur at any time throughout the disease course and it is estimated that up to 50% of patients will be affected. The presence of renal impairment has a significant negative impact on the ability to effectively treat myeloma as chemotherapy drugs often require dose changes or are associated with increased toxicity. Dialysis dependence is associated with a particularly poor overall survival for myeloma patients.

Myeloma-induced renal failure is a medical emergency requiring immediate management to avoid long-term dialysis. A number of approaches have been developed to try to reverse renal impairment and/or protect the kidneys from further damage. These approaches include hydration, plasmapheresis, high cut-off dialysis, chemotherapy and supportive treatments. There is variation in which of these approaches are currently used and uncertainty over their effectiveness

Clinical question: What is the optimal management of acute renal disease in patients with myeloma?

Clinical evidence (see also Appendix G)

See Tables 39-53.

Bortezomib-containing regimens + G-CSF, melphalan and auto-SCT versus VAD, VAD-like or TCED chemotherapy + G-CSF, melphalan and auto-SCT

The overall response rate prior to auto-SCT, overall response rate day +100 post auto-SCT and event-free survival were significantly better in the bortezomib group, whereas survival, relapse/progression day +100 post auto-SCT and post transplant toxicity and supportive treatment did not differ between the treatment groups (1 study [Breitkreutz, 2014], N = 27; very low quality).

Bortezomib-based regimens versus lenalidomide-based regimens

The complete renal response rate, major renal response rate, and time to major renal response were significantly better in the bortezomib group, whereas survival, early deaths, myeloma response, best eGFR and any renal response rate did not differ between the treatment groups (1 study [Dimopoulos 2013], N = 71; very low quality).

Bortezomib-based regimens versus thalidomide-based regimens

The major renal response rate, any renal response rate, survival, early deaths, myeloma response, and best eGFR did not differ between the treatment groups (1 study [Dimopoulos 2013], N = 105; very low quality).

Chemotherapy with thalidomide-based regimens versus chemotherapy with lenalidomide-based regimens

The major renal response rate, any renal response rate, time to major renal response, survival, early deaths, myeloma response, and best eGFR did not differ between the treatment groups (1 study [Dimopoulos 2013], N = 90; very low quality).

Dexamethasone, thalidomide and/or bortezomib versus VAD, VAD-like, melphalan plus dexamethasone or dexamethasone alone

Time to reversal of renal failure was significantly better in the dexamethasone, thalidomide and/or bortezomib group, whereas the reversal of renal failure rate and myeloma response rate did not differ between the treatment groups (1 study [Kastritis 2007], N = 41; very low quality).

Melphalan, prednisone, bortezomib and thalidomide + maintenance with bortezomib and thalidomide (VMPT-VT) versus bortezomib, melphalan and prednisone without maintenance (VMP)

In patients with eGFR ≤ 30, the complete myeloma response rate, myeloma response rate, time to first myeloma response, duration of myeloma response, reversal of renal impairment rate, progression-free survival, 2-year overall survival, discontinuation due to adverse events and adverse events rates did not differ between the treatment groups, apart from neutropenia, which was experienced significantly more in the VMPT-VT group (1 study [Morabito et al. 2011], N = 30; very low quality).

In patients with eGFR 31-50, myeloma response rate, and progression-free survival were significantly better in the VMPT-VT group, where as discontinuation due to adverse events was significantly higher in the VMPT-VT group also, with the complete myeloma response rate, time to first myeloma response, duration of myeloma response, and adverse events rates not differing between the treatment groups (1 study [Morabito et al. 2011], N = 110; very low quality).

In patients with eGFR ≤ 50, the myeloma response rate, complete myeloma response rate, and progression-free survival were significantly better in the VMPT-VT group, whereas the time to first myeloma response, duration of myeloma response, reversal of renal impairment rate, discontinuation due to adverse events and adverse events rates did not differ between the treatment groups (1 study [Morabito et al. 2011], N = 140; very low quality).

Bortezomib and dexamethasone-containing regimens versus thalidomide or lenalidomide-based regimens with dexamethasone and/or cyclophosphamide or melphalan (IMiDs-based chemotherapy)

The major renal response rate and time to major renal response were significantly better in the bortezomib-based group whereas the complete renal response rate did not differ between the treatment groups (1 study [Roussou 2010], N = 64; very low quality).

Bortezomib and dexamethasone-containing regimens versus VAD or VAD-like regimens, melphalan plus dexamethasone (conventional chemotherapy)

The major renal response rate, any renal response rate and time to major renal response were significantly better in the bortezomib-based group whereas the complete renal response rate did not differ between the treatment groups (1 study [Roussou 2010], N = 49; very low quality).

VAD or VAD-like regimens, melphalan plus dexamethasone (conventional chemotherapy) versus thalidomide or lenalidomide-based regimens with high-dose dexamethasone and/or cyclophosphamide or melphalan (IMiDs-based chemotherapy)

The any renal response rate was significantly better in the IMiDs-based group whereas the major renal response rate, complete renal response rate and time to major renal response did not differ between the treatment groups (1 study [Roussou 2010], N = 79; very low quality).

Chemotherapy with bortezomib, doxorubicin and dexamethasone; melphalan/ASCT + maintenance bortezomib (PAD) versus vincristine, doxorubicin and dexamethasone; melphalan/ASCT + maintenance thalidomide (VAD)

The myeloma response after 1-3 cycles of induction therapy, best myeloma response achieved any time during the trial treatment, 3-year progression-free survival, and 3-year overall survival were significantly better in the PAD group whereas renal function (creatinine level and clearance), renal response after 3 cycles of induction therapy, and adverse events (type and frequency) did not differ between the treatment groups (1 study [Scheid et al 2014], N = 81; very low quality).

Chemotherapy with melphalan, prednisone and bortezomib (VMP) versus melphalan and prednisone (MP)

In patients with eGFR ≤ 30, the complete myeloma response rate, myeloma response rate, time to progression, and overall survival did not differ between the treatment groups (1 study [Dimopoulos et al 2009], N = 34; very low quality).

In patients with eGFR 31-50, the complete myeloma response rate, myeloma response rate, and time to progression were significantly better in the VMP group, with overall survival differing between the treatment groups (1 study [Dimopoulos et al 2009], N = 191; very low quality).

In patients with eGFR ≤ 50, the myeloma response rate, complete myeloma response rate, time to progression and time to reversal of renal impairment were significantly better in the VMP group, whereas the reversal of renal impairment rate and overall survival did not differ between the treatment groups (1 study [Dimopoulos et al 2009], N = 225; very low quality).

Chemotherapy with bortezomib versus dexamethasone

The time to progression was significantly longer in the bortezomib group, whereas overall survival did not differ significantly between the treatment groups (1 study [San-Miguel et al. 2008], N = 120; very low quality).

Chemotherapy with melphalan, prednisone and thalidomide versus cyclophosphamide, dexamethasone and thalidomide

The ‘at least a very good partial myeloma response rate’, ‘at least partial myeloma response rate’, event-free survival, overall survival, neutropenia and infection with febrile neutropenia (including mortality thereof) were significantly worse in MPT-GRF < 40 group, compared to MPT-GRF ≥ 40, TCD-GRF < 40 group, and TCD-GRF ≥ 40 groups whereas the myeloma complete response rate, anaemia, thrombocytopenia, embolism, peripheral neuropathy, infection without neutropenia and gastrointestinal adverse effects did not differ significantly between the 4 treatment groups. Moreover, in patients with GFR ≥ 40, serum creatinine did not differ after 2, 4, 6, and 8 cycles between the treatments, whereas in patients with GFR < 40, serum creatinine was significantly higher in the MPT group after 2, 4, 6, and 8 cycles compared to the TCD group (1 study [Song 2012], N = 157; very low quality).

Plasmapheresis + chemotherapy with melphalan and prednisone versus chemotherapy with melphalan and prednisone

Survival was longer and renal function was either similar (hypercalcaemia, hyperuricaemia) or better (creatinine, oliguric at presentation and polyuric after treatment) after treatment with plasmapheresis and chemotherapy compared to chemotherapy alone (1 study [Abdulrahman 2003], N = 29; very low quality).

Plasmapheresis + chemotherapy with melphalan and prednisone or VAD versus chemotherapy with melphalan and prednisone or VAD

The composite outcome (death, dialysis dependence and an estimated GFR < 0.29 mL • s-2 • m-2) and its constituent parts did not differ after treatment with either plasmapheresis and chemotherapy or chemotherapy alone (1 study [Clark et al. 2005], N = 97; very low quality).

Health-related quality of life

No evidence was found.

Table 39: GRADE profile: What is the optimal management of myeloma-induced acute renal disease (Bortezomib-containing regimens + G-CSF, melphalan and auto-SCT' versus 'VAD, VAD-like or TCED chemotherapy + G-CSF, melphalan and auto-SCT)?

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Bortezomib chemo-therapy	VAD, VAD-like, or TCED chemo-therapy	Effect	Quality
Survival (follow-up: Bortezomib 53 months; VAD, VAD-like or TCED 84 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	13	14	The groups did not differ significantly	VERY LOW
Overall response rate prior to auto-SCT (follow-up: Bortezomib 53 months; VAD, VAD-like or TCED 84 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	13	14	Significantly better in bortezomib group	VERY LOW
Overall response rate day +100 post auto-SCT (follow-up: Bortezomib 53 months; VAD, VAD-like or TCED 84 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	13	14	Significantly better in bortezomib group	VERY LOW
Event-free survival (follow-up: Bortezomib 53 months; VAD, VAD-like or TCED 84 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	13	14	Significantly better in bortezomib group	VERY LOW
Relapse/progression day +100 post auto-SCT (follow-up: Bortezomib 53 months; VAD, VAD-like or TCED 84 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	13	14	The groups did not	VERY LOW

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Bortezomib chemo-therapy	VAD, VAD-like, or TCED chemo-therapy	Effect	Quality
									differ significantly	
Post transplant toxicity and supportive treatment (follow-up: Bortezomib 53 months; VAD, VAD-like or TCED 84 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	13	14	The groups did not differ significantly	VERY LOW

1 *1 Breitkreutz (2014);2 Unsure if the patients had acute renal disease; 3 Low number of events.*

2 **Table 40: GRADE profile: What is the optimal management of myeloma-induced acute renal disease (chemotherapy with bortezomib-**
3 **based regimens' versis 'chemotherapy with lenalidomide-based regimens')?**

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Bortezomib-based chemo-therapy	Lenalidomide-based chemo-therapy	Effect	Quality
Complete renal response (CR; median follow-up = 17.5 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	43	28	Significantly better in bortezomib group	VERY LOW
Major renal response (CR + PR; median follow-up = 17.5 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	43	28	Significantly better in bortezomib group	VERY LOW

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Bortezomib-based chemo-therapy	Lenalidomide-based chemo-therapy	Effect	Quality
Any renal response (at least minor response; median follow-up = 17.5 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	43	28	The groups did not differ significantly	VERY LOW
Time to major renal response (median follow-up = 17.5 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	43	28	Significantly better in bortezomib group	VERY LOW
Best eGRF (median follow-up = 17.5 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	43	28	The groups did not differ significantly	VERY LOW
Survival (median follow-up = 17.5 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	43	28	The groups did not differ significantly	VERY LOW
Early deaths (median follow-up = 17.5 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	43	28	The groups did not differ significantly	VERY LOW
Myeloma response (median follow-up = 17.5 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	43	28	The groups did not differ significantly	VERY LOW

1 *1 Dimopoulos (2013); 2 Unclear of the patients had "myeloma-induced acute renal disease"; 3 Low number of events.*

Table 41: GRADE profile: What is the optimal management of myeloma-induced acute renal disease (chemotherapy with bortezomib-based regimens' versus 'chemotherapy with thalidomide-based regimens')?

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Bortezomib-based chemo-therapy	Thalidomide-based chemo-therapy	Effect	Quality
Major renal response (CR + PR; median follow-up = 17.5 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	43	62	The groups did not differ significantly	VERY LOW
Any renal response (at least minor response; median follow-up = 17.5 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	43	62	The groups did not differ significantly	VERY LOW
Best eGRF (median follow-up = 17.5 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	43	62	The groups did not differ significantly	VERY LOW
Survival (median follow-up = 17.5 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	43	62	The groups did not differ significantly	VERY LOW
Early deaths (median follow-up = 17.5 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	43	62	The groups did not differ significantly	VERY LOW
Myeloma response (median follow-up = 17.5 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	43	62	The groups did not differ significantly	VERY LOW

1 Dimopoulos (2013); 2 Unclear of the patients had "myeloma-induced acute renal disease". 3 Low number of events.

Table 42: GRADE profile: What is the optimal management of myeloma-induced acute renal disease (chemotherapy with thalidomide-based regimens' versus 'chemotherapy with lenalidomide-based regimens')?

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Thalidomide-based chemo-therapy	Lenalidomide-based chemo-therapy	Effect	Quality
Major renal response (CR + PR; median follow-up = 17.5 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	62	28	The groups did not differ significantly	VERY LOW
Any renal response (at least minor response; median follow-up = 17.5 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	62	28	The groups did not differ significantly	VERY LOW
Time to major renal response (median follow-up = 17.5 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	62	28	The groups did not differ significantly	VERY LOW
Best eGRF (median follow-up = 17.5 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	62	28	The groups did not differ significantly	VERY LOW
Survival (median follow-up = 17.5 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	62	28	The groups did not differ significantly	VERY LOW
Early deaths (median follow-up = 17.5 months)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	62	28	The groups did not differ significantly	VERY LOW
Myeloma response (median follow-up = 17.5 months)										

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Thalidomide-based chemo-therapy	Lenalidomide-based chemo-therapy	Effect	Quality
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	62	28	The groups did not differ significantly	VERY LOW

1 *1 Dimopoulos (2013); 2 Unclear of the patients had "myeloma-induced acute renal disease". 3 Low number of events.*

2 **Table 43: GRADE profile: What is the optimal management of myeloma-induced acute renal disease ('chemotherapy with**
3 **dexamethasone and thalidomide and/or bortezomib' versus 'chemotherapy with VAD, VAD-like, melphalan plus**
4 **dexamethasone or dexamethasone alone')?**

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Dexameth-asone + thalidomide and/or bortezomib	VAD, VAD-like, melphalan plus dexameth-asone or dexameth-asone alone	Effect	Quality
Reversal of renal failure (follow-up not reported)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	15	26	The groups did not differ significantly	VERY LOW
Time to reversal of renal failure (follow-up not reported)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	15	26	Dexa-methasone + thalidomide and/or bortezomib significantly faster	VERY LOW

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Dexameth-asone + thalidomide and/or bortezomib	VAD, VAD-like, melphalan plus dexameth-asone or dexameth-asone alone	Effect	Quality
Myeloma response (CR+PR; follow-up not reported)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	15	26	The groups did not differ significantly	VERY LOW

1 Kastritis (2007); 2 Unclear of the patients had "myeloma-induced acute renal disease". 3 Low number of events.

Table 44: GRADE profile: What is the optimal management of myeloma-induced acute renal disease ('chemotherapy with melphalan, prednisone, bortezomib and thalidomide + maintenance with bortezomib and thalidomide (VMPT-VT)' versus 'chemotherapy with bortezomib, melphalan and prednisone without maintenance (VMP)')?

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	VMPT-VT	VMP	Effect	Quality
Patients with eGFR ≤ 30: Myeloma response rate (median follow-up = 21.6 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	11	19	The groups did not differ significantly	VERY LOW
Patients with eGFR ≤ 30: Complete myeloma response rate (median follow-up = 21.6 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	11	19	The groups did not differ significantly	VERY LOW
Patients with eGFR ≤ 30: Time to first myeloma response (median follow-up = 21.6 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	11	19	The groups did not differ significantly	VERY LOW

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	VMPT-VT	VMP	Effect	Quality
Patients with eGFR ≤ 30: Duration of myeloma response (median follow-up = 21.6 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	11	19	The groups did not differ significantly	VERY LOW
Patients with eGFR ≤ 30: Reversal of renal impairment (median follow-up = 21.6 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	11	19	The groups did not differ significantly	VERY LOW
Patients with eGFR ≤ 30: Progression-free survival (median follow-up = 21.6 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	11	19	The groups did not differ significantly	VERY LOW
Patients with eGFR ≤ 30: 2-year overall survival (median follow-up = 21.6 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	11	19	The groups did not differ significantly	VERY LOW
Patients with eGFR ≤ 30: Adverse events (median follow-up = 21.6 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	11	19	The groups did not differ significantly in any adverse event rates, including discontinuation due to adverse events, apart from neutropenia	VERY LOW

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	VMPT-VT	VMP	Effect	Quality
									which was experienced significantly more in the VMPT-VT group.	
Patients with eGFR 31-50: Myeloma response rate (median follow-up = 21.6 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	52	58	VMPT-VT significantly better	VERY LOW
Patients with eGFR 31-50: Complete myeloma response rate (median follow-up = 21.6 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	52	58	The groups did not differ significantly	VERY LOW
Patients with eGFR 31-50: Time to first myeloma response (median follow-up = 21.6 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	52	58	The groups did not differ significantly	VERY LOW
Patients with eGFR 31-50: Duration of myeloma response (median follow-up = 21.6 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	52	58	The groups did not differ significantly	VERY LOW
Patients with eGFR 31-50: Progression-free survival (median follow-up = 21.6 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	52	58	VMPT-VT significantly better	VERY LOW
Patients with eGFR 31-50: Adverse events (median follow-up = 21.6 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	52	58	The groups did not differ	VERY LOW

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	VMPT-VT	VMP	Effect	Quality
									significantly in any adverse event rates, but significantly more VMPT-VT patients discontinued treatment due to adverse events.	
Patients with eGFR ≤ 50: Myeloma response rate (median follow-up = 21.6 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	63	77	VMPT-VT significantly better	VERY LOW
Patients with eGFR ≤ 50: Complete myeloma response rate (median follow-up = 21.6 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	63	77	VMPT-VT significantly better	VERY LOW
Patients with eGFR ≤ 50: Time to first myeloma response (median follow-up = 21.6 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	63	77	The groups did not differ significantly	VERY LOW
Patients with eGFR ≤ 50: Duration of myeloma response (median follow-up = 21.6 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	63	77	The groups did not differ significantly	VERY LOW
Patients with eGFR ≤ 50: Reversal of renal impairment (median follow-up = 21.6 months)										

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	VMPT-VT	VMP	Effect	Quality
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	63	77	The groups did not differ significantly	VERY LOW
Patients with eGFR ≤ 50: Progression-free survival (median follow-up = 21.6 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	63	77	VMPT-VT significantly better	VERY LOW
Patients with eGFR ≤ 50: Adverse events (median follow-up = 21.6 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	63	77	The groups did not differ significantly in any adverse event rates, including discontinuation due to adverse events.	VERY LOW

1 Morabito (2011); 2 Unclear risk of patient selection, no blinding details reported. 3 Unclear of the patients had "myeloma-induced acute renal disease"; 4 Low number of events.

Table 45: GRADE profile: What is the optimal management of myeloma-induced acute renal disease ('bortezomib and dexamethasone-containing regimens' versus 'chemotherapy with thalidomide or lenalidomide-based regimens with high-dose dexamethasone and/or cyclophosphamide or melphalan (IMiDs-based chemotherapy)')?

Quality assessment	Summary of findings		
	No of patients	Effect	Quality

No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Bortezomib-based chemo-therapy	IMiDs-based chemo-therapy		
Major renal response (PR + CR; follow-up not reported)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	17	47	Bortezomib-based significantly better	VERY LOW
Complete renal response										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	17	47	The groups did not differ significantly	VERY LOW
Time to major renal response (follow-up not reported)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	17	47	Bortezomib-based significantly faster	VERY LOW

1 Roussou (2010); 2 Unclear of the patients had "myeloma-induced acute renal disease". 3 Low number of events.

Table 46: GRADE profile: What is the optimal management of myeloma-induced acute renal disease ('bortezomib and dexamethasone-containing regimens' versus 'chemotherapy with VAD or VAD-like regimens, melphalan plus dexamethasone (conventional chemotherapy)')?

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Bortezomib-based chemo-therapy	Conventional chemo-therapy	Effect	Quality
Any renal response (at least minor response; follow-up not reported)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	17	32	Bortezomib-based significantly better	VERY LOW

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Bortezomib-based chemo-therapy	Conventional chemo-therapy	Effect	Quality
Major renal response (PR + CR; follow-up not reported)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	17	32	Bortezomib-based significantly better	VERY LOW
Complete renal response										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	17	32	The groups did not differ significantly	VERY LOW
Time to major renal response (follow-up not reported)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	17	32	Bortezomib-based significantly faster	VERY LOW

1 Roussou (2010); 2 Unclear of the patients had "myeloma-induced acute renal disease"; 3 Low number of events.

Table 47: GRADE profile: What is the optimal management of myeloma-induced acute renal disease ('chemotherapy with VAD or VAD-like regimens, melphalan plus dexamethasone (conventional chemotherapy)' versus 'chemotherapy with thalidomide or lenalidomide-based regimens with high-dose dexamethasone and/or cyclophosphamide or melphalan (IMiDs-based chemotherapy)')?

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Conventional chemo-therapy	IMiDs-based chemo-therapy	Effect	Quality
Any renal response (at least minor response; follow-up not reported)										
1	observational	no serious	no serious	no serious	very serious	none	32	47	IMiDs-based	VERY

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Conventional chemo-therapy	IMiDs-based chemo-therapy	Effect	Quality
	study[1]	limitations	inconsistency	indirectness[2]	imprecision[3]				significantly better	LOW
Major renal response (PR + CR; follow-up not reported)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	32	47	The groups did not differ significantly	VERY LOW
Complete renal response										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	32	47	The groups did not differ significantly	VERY LOW
Time to major renal response (follow-up not reported)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness[2]	very serious imprecision[3]	none	32	47	The groups did not differ significantly	VERY LOW

1 *1 Roussou (2010); 2 Unclear of the patients had "myeloma-induced acute renal disease". 3 Low number of events.*

2 **Table 48: GRADE profile: What is the optimal management of myeloma-induced acute renal disease ('chemotherapy with**
3 **bortezomib, doxorubicin and dexamethasone; melphalan/ASCT + maintenance bortezomib (PAD)' versus 'chemotherapy**
4 **with vincristine, doxorubicin and dexamethasone; melphalan/ASCT + maintenance thalidomide (VAD)')?**

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	PAD	VAD	Effect	Quality
Renal function after induction (creatinine level and clearance; follow-up not reported)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	36	45	The groups did not differ significantly	VERY LOW

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	PAD	VAD	Effect	Quality
Renal response after 3 cycles of induction therapy (follow-up not reported)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	36	45	The groups did not differ significantly	VERY LOW
Myeloma response after 1-3 cycles of induction therapy (follow-up not reported)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	36	45	PAD significantly better	VERY LOW
Best myeloma response achieved any time during trial treatment (follow-up not reported)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	36	45	PAD significantly better	VERY LOW
3-year progression-free survival (follow-up not reported)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	36	45	PAD significantly better	VERY LOW
3-year overall survival (follow-up not reported)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	36	45	PAD significantly better	VERY LOW
Adverse events (follow-up not reported)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	36	45	The groups did not differ significantly in frequency or type of adverse	VERY LOW

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	PAD	VAD	Effect	Quality
									events.	

1 Scheid (2014); 2 Unclear risk of patient selection, no blinding details reported. 3 Unclear of the patients had "myeloma-induced acute renal disease"; 4 Low number of events.

Table 49: GRADE profile: What is the optimal management of myeloma-induced acute renal disease ('chemotherapy with melphalan, prednisone, and bortezomib (VMP)' versus 'chemotherapy with melphalan and prednisone (MP)')?

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	VMP	MP	Effect	Quality
Patients with eGFR ≤ 30: Myeloma response rate (median follow-up = 25.9 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	19	15	The groups did not differ significantly	VERY LOW
Patients with eGFR ≤ 30: Complete myeloma response rate (median follow-up = 25.9 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	19	15	The groups did not differ significantly	VERY LOW
Patients with eGFR ≤ 30: Time to progression (median follow-up = 25.9 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	19	15	The groups did not differ significantly	VERY LOW
Patients with eGFR ≤ 30: Overall survival (median follow-up = 25.9 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	19	15	The groups did not differ significantly	VERY LOW
Patients with eGFR 31-50: Myeloma response rate (median follow-up = 25.9 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	92	99	VMP significantly	VERY LOW

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	VMP	MP	Effect	Quality
									better	
Patients with eGFR 31-50: Complete myeloma response rate (median follow-up = 25.9 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	92	99	VMP significantly better	VERY LOW
Patients with eGFR 31-50: Time to progression (median follow-up = 25.9 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	92	99	VMP significantly better	VERY LOW
Patients with eGFR 31-50: Overall survival (median follow-up = 25.9 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	92	99	The groups did not differ significantly	VERY LOW
Patients with eGFR ≤ 50: Myeloma response rate (median follow-up = 25.9 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	111	114	VMP significantly better	VERY LOW
Patients with eGFR ≤ 50: Complete myeloma response rate (median follow-up = 25.9 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	111	114	VMP significantly better	VERY LOW
Patients with eGFR ≤ 50: Reversal of renal impairment rate (median follow-up = 25.9 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	111	114	The groups did not differ significantly	VERY LOW
Patients with eGFR ≤ 50: Time to reversal of renal impairment (median follow-up = 25.9 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	111	114	VMP significantly better	VERY LOW

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	VMP	MP	Effect	Quality
Patients with eGFR ≤ 50: Time to progression (median follow-up = 25.9 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	111	114	VMP significantly better	VERY LOW
Patients with eGFR ≤ 50: Overall survival (median follow-up = 25.9 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	111	114	The groups did not differ significantly	VERY LOW

1 Dimopoulos (2009); 2 Unclear risk of patient selection, no blinding details reported. 3 Unclear of the patients had "myeloma-induced acute renal disease"; 4 Low number of events.

Table 50: GRADE profile: What is the optimal management of myeloma-induced acute renal disease ('chemotherapy with bortezomib' versus 'chemotherapy with dexamethasone')?

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Bortezomib	Dexameth-asone	Effect	Quality
Time to progression (median follow-up ≤ 22 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	58	62	Bortezomib significantly better	VERY LOW
Overall survival (median follow-up ≤ 22 months)										
1	Randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	58	62	The groups did not differ significantly	VERY LOW

1 San-Miguel (2008); 2 Unclear risk of patient selection, no blinding details reported. 3 Unclear of the patients had "myeloma-induced acute renal disease"; 4 Low number of events.

Table 51: GRADE profile: What is the optimal management of myeloma-induced acute renal disease ('chemotherapy with melphalan, prednisone, and thalidomide (MPT)' versus 'chemotherapy with cyclophosphamide, dexamethasone and thalidomide (TCD)')?

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	MPT: Divided into MPT-GFR < 40 and MPT-GFR ≥ 40	TCD: Divided into TCD-GFR < 40 and TCD-GFR ≥ 40	Effect	Quality
Myeloma complete response rate (median follow-up = 36 months)										
1	Observational study[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	30/44	38/45	The groups did not differ significantly	VERY LOW
At least very good partial myeloma complete response rate (median follow-up = 36 months)										
1	Observational study[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	30/44	38/45	MPT-GRF < 40 significantly worse than the other 3 groups	VERY LOW
Event-free survival (median follow-up = 36 months)										
1	Observational study[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	30/44	38/45	MPT-GRF < 40 significantly worse than the other 3 groups	VERY LOW
Overall survival (median follow-up = 36 months)										
1	Observational study[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	30/44	38/45	MPT-GRF < 40 significantly worse than the other 3 groups	VERY LOW
Serum creatinine (median follow-up = 36 months)										
1	Observational study[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	30/44	38/45	GFR ≥ 40: MPT = TCD after 2, 4, 6 and 8 cycles;	VERY LOW

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	MPT: Divided into MPT-GFR < 40 and MPT-GFR ≥ 40	TCD: Divided into TCD-GFR < 40 and TCD-GFR ≥ 40	Effect	Quality
									GRF < 40: Significantly higher in MPT after 2, 4, 6 and 8 cycles	
Haematological adverse effects (median follow-up = 36 months)										
1	Observational study[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	30/44	38/45	Neutropenia: MPT-GRF < 40 significantly worse than the other 3 groups; Anaemia and thrombocytopenia : The groups did not differ significantly	VERY LOW
Non-haematological adverse effects (median follow-up = 36 months)										
1	Observational study[1]	serious limitations[2]	no serious inconsistency	no serious indirectness[3]	very serious imprecision[4]	none	30/44	38/45	Infection with febrile neutropenia and mortality due to this: MPT-GRF < 40 significantly worse than the other 3 groups; Embolism, peripheral neuropathy,	VERY LOW

Quality assessment							Summary of findings			
							No of patients			
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	MPT: Divided into MPT-GFR < 40 and MPT-GFR ≥ 40	TCD: Divided into TCD-GFR < 40 and TCD-GFR ≥ 40	Effect	Quality
									infection without neutropenia and gastrointestinal: The groups did not differ significantly	

1Unclear if Song et al (2012) is a retrospective study or RCT; if RCT no details reported about patient selection/allocation methods; 2 Unclear risk of patient selection, no blinding details reported. 3 Unclear of the patients had "myeloma-induced acute renal disease"; 4 Low number of events.

Table 52: GRADE profile: What is the optimal management of myeloma-induced acute renal disease ('plasmapheresis plus chemotherapy with melphalan and prednisone' versus 'chemotherapy with melphalan and prednisone')?

Quality assessment							Summary of findings			
							No of patients		Effect	
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Plasmapheresis + chemo-therapy	Chemo-therapy	Relative (95% CI)	Quality
Survival (follow-up not reported)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness	very serious imprecision [2]	none	15	14	Significantly longer in plasmapheresis group	VERY LOW
Renal function (follow-up not reported)										
1	observational study[1]	no serious limitations	no serious inconsistency	no serious indirectness	very serious imprecision [2]	none	15	14	Similar or significantly better in plasmapheresis group	VERY LOW

1 Abdulrahman (2003); 2 Low number of events.

Table 53: GRADE profile: What is the optimal management of myeloma-induced acute renal disease ('plasmapheresis plus chemotherapy with melphalan and prednisone or with VAD' versus 'chemotherapy with melphalan and prednisone or VAD')?

Quality assessment							Summary of findings			
							No of patients		Effect	
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Plasmapheresis + chemo-therapy	Chemo-therapy	Relative (95% CI)	Quality
Composite outcome (death, dialysis dependence and an estimated GFR < 0.29 mL • s-2 • m-2) and its constituent parts (6 month follow-up)										
1	randomised trial[1]	serious limitations[2]	no serious inconsistency	no serious indirectness	very serious imprecision[3]	none	58	39	No difference between the groups	VERY LOW

1 Clark (2005); 2 No blinding. 3 Low number of events.

Cost effectiveness evidence (see also Appendix F)

The following databases were searched for economic evidence relevant to the clinical question: MEDLINE, EMBASE, COCHRANE, NHS EED. Studies conducted from any OECD countries were considered (Guidelines Manual 2014). 463 possibly relevant papers were identified. Of these, 1 full paper relating to this topic was obtained for appraisal. This paper (Grima et al. 2011) was included in the current review of published economic evidence for this topic (Table 54).

The study was a cost-effectiveness analysis of high cut-off haemodialysis (HCO-HD) versus standard haemodialysis (HD) in patients with myeloma complicated by dialysis dependant renal failure secondary to myeloma kidney. The study reported the results in terms of cost per Quality Adjusted Life Year (QALY) gained and considered a NHS and Personal Social Services (PSS) perspective.

Grima et al is deemed directly applicable to the decision problem that we are evaluating. This is because it took a NHS and PSS perspective and reported health outcomes in terms of QALYs. Both costs and outcomes were also discounted at an annual rate of 3.5%.

Potentially serious limitations were identified with Grima et al. Most notably, a potential conflict of interest as the study was funded by a manufacturer of HCO-HD. Uncertainty around the effectiveness of HCO-HD compared to HD was also not adequately explored. The range of tested values given for the difference in the percentage of patients recovering renal function between the two interventions was inadequately narrow, given the uncertainty around this parameter. There was also inadequate exploration around other key parameters.

The base case suggested that using HCO-HD over HD would lead to total cost savings of £6,500 and 0.75 additional QALYs per patient (HCO-HD dominant). This result was robust to all but one of the deterministic sensitivity analyses reported. Probabilistic sensitivity analysis suggested the results were robust with 99.7% of iterations being cost effective at a threshold of £20,000 per QALY. Over 80% of iterations were also cost-saving and health improving. These results should be interpreted with caution given the issues with the sensitivity analyses.

1 **Table 54: Modified GRADE profile: included economic studies**

Study	Population	Comparators	Costs	Effects	Incr costs	Incr effects	ICER	Uncertainty	Applicability	Limitations
Grima et al (2011) UK	Hypothetical cohort of patients with myeloma complicated by dialysis dependant renal failure secondary to myeloma kidney.	Standard Haemodialysis (HD)	£31,345	1.07	Reference			Probabilistic sensitivity analyses: HCO-HD was below less than £20,000 per additional QALY in all deterministic sensitivity analyses and dominant in all but one when compared to HD. Probabilistic sensitivity analysis estimated that there was a 99.7% probability that HCO-HD was cost effective at a willingness to pay of £20,000 per QALY and >80% probability of being both more effective and cost saving when compared to HD.	Directly Applicable	Potentially Serious Limitations The study did not adequately investigate some important parameters during sensitivity analysis Potential conflict of interest as the study was funded by the manufacturer of HCO-HD
		High cut-off haemodialysis (HCO-HD)	£24,845	1.82	-£6500	0.75	HCO-HD dominant			
	Comments:									

2

1

Recommendations	**Consider immediately starting a bortezomib- and dexamethasone-based combination regimen for people with untreated, newly diagnosed, myeloma-induced acute renal disease.** **If a bortezomib-based combination regimen is unsuitable for people with untreated, newly diagnosed, myeloma-induced acute renal disease, consider immediately starting a thalidomide- and dexamethasone-based combination regimen[a].** **Do not perform plasma exchange for myeloma-induced acute renal disease.**
Relative value placed on the outcomes considered	The Guideline Committee considered the outcomes of improvement in renal function, recovery from dialysis, rate of dialysis, overall survival, progression free survival, health related quality of life and adverse events to be the most relevant to identify the optimal management of acute renal disease in patients with myeloma. When drafting recommendations the Guideline Committee considered improvement in renal function, overall survival and progression-free survival to be the most important outcomes. Evidence was found for all outcomes except health-related quality of life.
Quality of the evidence	The quality of the evidence was assessed by GRADE and was of very low quality for all reported outcomes. The Guideline Committee noted three major issues with the evidence. Firstly, for the retrospective studies, the treatments varied considerably within the treatment groups. Secondly, for all of the comparisons, the evidence consisted of only one study, and thirdly the sample sizes were generally very small. As a result of these limitations and the very low quality evidence the Guideline Committee made 'consider' recommendations rather than 'offer'. The Guideline Committee noted that evidence was only available for patients with untreated, newly diagnosed myeloma-induced acute renal disease and that the evidence did not cover patients with relapsed myeloma.
Trade off between clinical benefits and harms	The Guideline Committee decided to only make recommendations for the patient group for which there was evidence. They agreed not to make recommendations for patients with relapsed myeloma based on their clinical experience because the optimal treatment for renal disease in these patients would depend on a number of different factors that are specific to the individual patient, including previous treatments. Based on the evidence, the Guideline Committee agreed to recommend a bortezomib and dexamethasone-based regimen for

a At the time of consultation (August 2015), thalidomide in combination with dexamethasone did not have a UK marketing authorisation for this indication. The prescriber should follow relevant professional guidance, taking full responsibility for the decision. Informed consent should be obtained and documented. See the General Medical Council's Prescribing guidance: prescribing unlicensed medicines for further information.

	patients with untreated, newly diagnosed myeloma-induced acute renal disease, and a thalidomide and dexamethasone-based regimen if bortezomib-based treatment is unsuitable. The Guideline Committee noted that there is biological evidence that renal impairment will become permanent if treatment is not commenced promptly for myeloma-induced acute renal disease. They therefore agreed to recommend that treatment is commenced immediately. The Guideline Committee noted that whilst there are phase II data on the utility of lenalidomide- and bendamustine-based regimens in this patient group, there are no comparative data. Consequently it was not possible to make recommendations on the use of these interventions in the management of myeloma-induced acute renal disease. Based on the evidence, the Guideline Committee noted that the study showing a positive effect of plasmapheresis involved the use of melphalan and prednisolone which would not currently form part of standard treatment. The other study did not show any effect over chemotherapy alone. In addition based on their clinical experience, the Guideline Committee identified that there was the risk of adverse events associated with this intervention and resource use and capacity issues. On this basis the Guideline Committee agreed to recommend that plasmapheresis should not be performed for myeloma-induced acute renal disease. The Guideline Committee concluded that appropriate management of acute renal disease would result in a number of benefits, including the avoidance of long-term dialysis and a consequent improvement in health-related quality of life, longer overall survival, and a restoration of renal function, which would enable patients to qualify for clinical trials. The Guideline Committee agreed that potential harms were toxicity and adverse events from the treatments. Overall the Guideline Committee agreed that the benefits of the recommendations in terms of improved patient outcomes outweighed the potential harms.
Trade-off between net health benefits and resource use	The economic evidence review identified one previous economic evaluation for this topic. This study compared high cut-off haemodialysis (HCO-HD) to standard haemodialysis (HD) in patients with myeloma complicated by dialysis dependant renal failure taking a NHS and PSS perspective. The study was deemed directly applicable with potentially serious limitations. The authors reported there was strong evidence that HCO-HD was cost effective with all but one sensitivity analysis showing it as both health improving and cost saving. The results of the model were sensitive to the difference in the ‘percentage of patients recovering renal function’ between the two interventions. Weak, non-comparative evidence was used to inform this parameter in the economic analysis. The Guideline Committee agreed, in the absence of evidence from the accompanying clinical evidence review that the estimate in the base case was likely to be much higher than the true value. There was considered to be an inadequate exploration of uncertainty around this parameter with a one way sensitivity analysis only decreasing the base case value of recovery of renal function in the HCO-HD group by three percentage points. The range used in

	the probabilistic sensitivity analysis around this variable was also considered too narrow. It was therefore unclear what effect a smaller difference would have on the outcomes of the economic analysis. With weak evidence and inadequate exploration of uncertainty around this key parameter the Guideline Committee felt unable to use this evidence in informing their recommendations. The GC therefore made no recommendations on HCO-HD. The Guideline Committee estimated that the recommendations will lead to an overall cost-saving because of the avoidance of long-term dialysis for a number of patients and because the additional costs associated with the use of bortezomib would only affect a small subset of the patient population.

References

Abdulrahman IS (2003) A prospective study of renal failure in multiple myeloma: A promising role for plasmapheresis. HAEMA, 6: 358-365.

Breitkreutz I (2014) Bortezomib improves outcome after SCT in multiple myeloma patients with end-stage renal failure. Bone Marrow Transplantation, 49: 1371-1375

Clark WF, Stewart AK, Rock GA, et al. (2005) Plasma exchange when myeloma presents as acute renal failure: a randomized, controlled trial.[Erratum appears in Ann Intern Med. 2007 Mar 20;146(6):471; PMID: 17402169], [Summary for patients in Ann Intern Med. 2005 Dec 6;143(11):I20; PMID: 16330784]. Annals of Internal Medicine, 143: 777-784.

Dimopoulos MA, Richardson PG, Schlag R, et al. (2009) VMP (Bortezomib, Melphalan, and Prednisone) is active and well tolerated in newly diagnosed patients with multiple myeloma with moderately impaired renal function, and results in reversal of renal impairment: cohort analysis of the phase III VISTA study. Journal of clinical.oncology, 27: 6086-6093.

Dimopoulos MA (2013) The role of novel agents on the reversibility of renal impairment in newly diagnosed symptomatic patients with multiple myeloma. Leukemia, 27: 423-429.

Grima DT, Airia P, Attard C, et al. (2011) 'Modelled cost-effectiveness of high cut-off haemodialysis compared to standard haemodialysis in the management of myeloma kidney' Current Medical Research & Opinion 27(2): 383-391.

Kastritis EA (2007) Reversibility of renal failure in newly diagnosed multiple myeloma patients treated with high dose dexamethasone-containing regimens and the impact of novel agents. Haematologica, 92: 546-549.

Morabito F, Gentile M, Mazzone C, et al. (2011) Safety and efficacy of bortezomib-melphalan-prednisone-thalidomide followed by bortezomib-thalidomide maintenance (VMPT-VT) versus bortezomib-melphalan-prednisone (VMP) in untreated multiple myeloma patients with renal impairment. Blood, 118: 5759-5766.

Roussou M (2010) Reversibility of renal failure in newly diagnosed patients with multiple myeloma and the role of novel agents. Leukemia Research, 34: 1395-1397.

San-Miguel JF, Richardson PG, Sonneveld P, et al. (2008) Efficacy and safety of bortezomib in patients with renal impairment: results from the APEX phase 3 study. Leukemia, 22: 842-849.

Scheid C, Sonneveld P, Schmidt W et al. (2014) Bortezomib before and after autologous stem cell transplantation overcomes the negative prognostic impact of renal impairment in

newly diagnosed multiple myeloma: A subgroup analysis from the HOVON-65/GMMG-HD4 trial. Haematologica, 99: 148-154.

Song MK (2012) Cyclophosphamide-containing regimen (TCD) is superior to melphalan-containing regimen (MPT) in elderly multiple myeloma patients with renal impairment. Annals of Hematology, 91: 889-896.

8 Preventing and managing bone disease

8.1 Preventing bone disease

Bone disease remains the most common presenting feature of myeloma. Clinical features of bone disease resulting from myeloma may take the form of bone pain, bone fractures occurring spontaneously or following minimal trauma (pathological fractures), spinal cord compression, high calcium in the blood (hypercalcaemia) with possible consequent renal damage, and development of holes in the bones (lytic lesions). These features are usually named collectively as skeletal related events (SREs).

Combination chemotherapy is the primary management of patients with symptomatic myeloma and this may have a beneficial effect on SREs as a result of treating the myeloma. A number of clinical trials have also examined the efficacy of other treatment measures that can specifically prevent and/or treat SREs. Bisphosphonates, a class of drugs that inhibit osteoclastic activity, were the first bone directed therapy shown in randomised clinical trials to improve SREs in patients with myeloma. Bisphosphonate therapy is now commonly used to prevent bone disease in symptomatic patients, however some aspects of its use remain unclear. These include type of bisphosphonate, treatment duration and frequency, their use in patients with smouldering myeloma and alternative treatment options in patients who could not tolerate the bisphosphonate therapy. Also the use of some bisphosphonates can cause complications such as osteonecrosis of the jaw (ONJ). Alternatives/adjuncts to bisphosphonates include calcium supplements, vitamin D supplements, osteoclast inhibitors such as denosumab, bone anabolic therapy and exercise but there is variation in which of these is used and uncertainty over which is most effective.

Clinical question: What is the most effective method of preventing bone disease in patients with myeloma?

Clinical evidence (see also Appendix G)

See Tables 55-58.

Overall survival (OS)

Pooled results of 12 RCTs (2292 patients) in Mhaskar et al provide low quality evidence suggesting that bisphosphonates do not improve OS when compared with placebo or no treatment (HR 0.96; 95% CI 0.82 - 1.13). However, there was statistically significant heterogeneity among the included RCTs (I2 = 55%, P = 0.01).

Results from the Mhaskar et al (2012) Cochrane review network meta-analyses which included all studies that examined overall survival (12 RCTs comparing bisphosphonate with placebo or no treatment, and 2 RCTs with a different bisphosphonate as a comparator) demonstrated that zoledronate is superior to placebo and etidronate in improving OS. Meta-analyses of 14 RCTs (4766 patients) showed superior OS with zoledronate compared with etidronate (HR 0.43, 95% CI 0.16 to 0.86) and placebo (HR 0.61, 95% CI 0.28 to 0.98). However, there was no difference between zoledronate and other bisphosphonates.

Results from Henry et al provide moderate quality evidence of increased overall survival in myeloma patients receiving denosumab compared to those receiving zoledronic acid (HR 2.26; 95% CI 1.13 - 4.50).

Progression-free survival (PFS)

Pooled analysis of 4 RCTs (364 patients) in Mhaskar et al provide very low quality evidence suggesting that bisphosphonates do not improve PFS when compared with placebo or no treatment (HR 0.70; 95% CI 0.41 - 1.19).

Skeletal-related events (SRE)

Pooled analysis of 7 RCTs (1116 patients) in Mhaskar et al provides moderate quality evidence of a beneficial effect of bisphosphonates compared with placebo or no treatment in preventing pathological vertebral fractures (RR 0.74; 95% CI 0.62 - 0.89; p=0.001). Results also demonstrated an effect of bisphosphonates on the prevention of total skeletal-related events (7 RCTs, 1497 patients) (RR 0.80; 95% CI 0.72 - 0.89; p<0.0001). There was uncertainty whether bisphosphonates were more or less effective than placebo or no treatment in reducing nonvertebral fractures (6 RCTs, 1389 patients) (RR 1.03; 95% CI 0.68 - 1.56).

Results from network meta-analyses in Mhaskar et al found no evidence for superiority of any specific bisphosphonate for preventing skeletal related events. However, a head-to-head comparative study of the effects of zoledronic acid versus clodronic acid (Morgan et al., 2011) provides moderate quality evidence demonstrating that treatment with zoledronic acid is superior to clodronic acid with regards to preventing skeletal-related events. Fewer patients in the zoledronic acid group had vertebral fractures than did those in the clodronic acid group (5% vs. 9%, p=0.0008), other fractures (5% vs. 7%, p=0.04), and new osteolytic lesions (5% vs. 10%, p<0.0001).

Results from Henry et al provide moderate quality evidence that there is uncertainty about whether the time to first on-study SRE is longer with denosumab or zoledronic acid (HR 1.03; 95% CI 0.68 - 1.57).

Incidence of hypercalcemia (≥ 2.65 mmol/L)

Pooled analysis of 8 RCTs (1934 patients) in Mhaskar et al provide moderate quality evidence of uncertainty in relative effectiveness of bisphosphonates compared with placebo or no treatment in reducing hypercalcemia (RR 0.79; 95% CI 0.56 - 1.11). The 95% confidence interval of the effective estimate includes both significant benefit with bisphosphonates and no difference between the treatments.

Pain

Pooled analysis of 8 RCTs (1281 patients) in Mhaskar et al provide very low quality evidence that demonstrated a beneficial effect of bisphosphonates compared with placebo or no treatment on amelioration of pain (RR 0.75; 95% CI 0.60 - 0.95; p=0.01). However, there was statistically significant heterogeneity among the included RCTs (I2 = 63%, P = 0.008) more than likely due the variation in the pain reporting methods and quality of included studies.

Adverse events

Osteonecrosis of the jaw (ONJ)

ONJ was at reported a rate of 0.8% with bisphosphonate treatment but no cases were reported with placebo or no treatment in a systematic review of 3 RCTs including 736 patients (Mhaskar et al). The pooled results do not show a statistically significant increase in frequency of ONJ with the use of bisphosphonates compared with placebo or no treatment (RR 3.99; 95% CI 0.44 - 5.84), this was due to the very low event rate for ONJ in these studies which is why the evidence is considered low quality.

Two RCTs with bisphosphonate as the comparator also reported estimates of ONJ. In the RCT by Morgan et al (Morgan 2010), zoledronate was associated with higher rates of ONJ (35/983 (4%)) than clodronate (3/979 (< 1%)). In the RCT by Gimsing et al, ONJ was reported in 2 of 252 (0.79%) patients receiving 30mg of pamidronate compared with 8 of 250 (3.2%) patients receiving 90mg of pamidronate (Gimsing 2010).

Even though only 5 RCTs reported ONJ, a growing number of ONJ case reports and observational studies evaluating ONJ have been published in recent years and these studies were included in the data extracted for the Cochrane review which found that the rates of ONJ in observational studies (9 studies, 1400 patients) (table 5) ranged from 0% to 51% (the quality of this evidence is very low). The highest frequencies of ONJ were seen in studies that used a combination of pamidronate and zoledronate (range 5% to 51%). Zoledronate was associated with ONJ in 3% to 11% of cases. Pamidronate related frequencies of ONJ ranged from 0% to 18%.

Gastrointestinal symptoms

The pooled results (6 RCTs, 1689 patients) in Mhaskar et al provide low quality evidence that showed no statistically significant increase in frequency of gastrointestinal symptoms with the use of bisphosphonates compared with placebo or no treatment (RR 1.23; 95% CI 0.95 - 1.60), although the confidence intervals for the effect estimate include the possibility that bisphosphonates are associated with an increased rate of gastrointestinal symptoms.

One RCT with bisphosphonate as the comparator also reported estimates of GI symptoms (Morgan 2010). In this study 24 of 981 (2.4%) patients enrolled in the zoledronate arm had GI symptoms, and 30 of 979 (3.1%) patients receiving clodronate had GI symptoms.

Hypocalcaemia

The pooled results (3 RCTs, 1002 patients) in Mhaskar et al provide very low quality evidence of uncertainty about the relative frequency of hypocalcaemia with the use of bisphosphonates compared with placebo or no treatment (RR 2.19; 95% CI 0.49 - 9.74).

One RCT with bisphosphonate as the comparator also reported estimates of hypocalcaemia (Terpos 2003). In this study none of the 23 patients enrolled in the pamidronate arm had hypocalcaemia, while 2 of 19 patients receiving ibandronate did.

Renal dysfunction

The pooled results (2 RCTs, 414 patients) in Mhaskar et al provide low quality evidence of uncertainty about the relative frequency of renal dysfunction with the use of bisphosphonates compared with placebo or no treatment (the pooled mean difference in serum creatinine was -0.36 (95%CI -9.75 to 9.03).

One RCT with bisphosphonate as the comparator also reported estimates of renal failure (Morgan 2010). In this study 57 of 983 (5.8%) patients enrolled in the zoledronate arm had renal failure, while 60 of 979 (6.1%) patients receiving clodronate had renal failure.

The network meta-analysis in Mhaskar et al did not show any differences in the incidence of osteonecrosis of the jaw, hypocalcaemia, renal dysfunction and gastrointestinal toxicity between the bisphosphonates used.

The study by Henry et al reported on adverse events but these were reported for the whole population and not by tumour type and so there is no evidence from this study regarding occurrence of adverse events in myeloma patients. For the whole population patients in both treatment groups (denosumab or zoledronic acid) experienced similar rates of overall adverse events. Hypocalcaemia occurred more frequently with denosumab (10.8% vs. 5.8%), acute phase reactions after the first dose occurred more frequently with zoledronic

acid (14.5% vs. 6.9%), renal adverse events occurred more frequently with zoledronic acid (10.9% vs. 8.3%) and elevations in serum creatinine occurred more frequently with zoledronic acid (23.9% vs. 16.5%).

Need for radiotherapy

No evidence was found for this outcome.

Quality life

No evidence was found for this outcome.

Table 55: GRADE summary of findings table (benefits): What is the most effective method of preventing bone disease in patients with myeloma (bisphosphonates versus placebo or no treatment)? (from Mhaskar et al., 2012)

Note: not all studies included patients with lytic lesions or did not specify bone disease in inclusion criteria

Summary of findings			
No of patients	**Effect**		**Quality**
	Relative (95% CI)	**Absolute**	
Overall mortality			
2292 (12 studies)	HR 0.96 (0.82 to 1.13)	530 per 1000 with control, 504 per 1000 (449 to 561) with bisphosphonate	low[1,2,3]
Progression free survival			
364 (4 studies)	HR 0.70 (0.41 to 1.19)	350 per 1000 with control, 260 per 1000 (162 to 401) with bisphosphonate	very low[1,4]
Vertebral fractures			
1389 (6 studies)	RR 0.74 (0.62 to 0.89)	350 per 1000 with control, 259 per 1000 (217 to 311) with bisphosphonate	moderate[1,6]
Non vertebral fractures			
1389 (6 studies)	RR 1.03 (0.68 to 1.56)	140 per 1000 with control, 144 per 1000 (95 to 218)with bisphosphonate	moderate[1,7]
Skeletal-related events			
1497 (7 studies)	RR 0.80 (0.72 to 0.89)	303 per 1000 with control, 245 per 1000 (218 to 279) with bisphosphonate	moderate[1,8]
Pain			
1281 (8 studies	RR 0.75 (0.6 to 0.95)	500 per 1000 with control, 375 per 1000 (300 to 475) with bisphosphonate	very low[9,10, 11]
Hypercalcaemia			
1934 (8 studies)	RR 0.79 (0.56 to 1.11)	100 per 1000 with control, 87 per 1000 (61 to 124) with bisphosphonate	moderate[1]

1 Only 37% (6/16) of trials had adequate allocation concealment. Only 18% (3/16) of trials reported methods of randomization. Similarly, 18% (3/16) of trials reported blinding procedures and personnel who were blinded to the intervention assignment. However, sensitivity analyses based on allocation concealment and description of randomization method didn't change the estimates. Hence, the assessment of studies limitations may represent the poor quality of reporting rather than true biased estimates.

2 I2 = 55%. The pooled estimate is driven by studies by Aviles et al and Belch et al; when we removed these RCTs pooled estimates remained the same but heterogeneity disappeared.
3 The overall mortality data were extractable from 11 of 16 studies. Also, note that overall mortality data denotes the mortality rates, i.e. the number of events refers to the number of deaths.
4 The progression-free survival data could be extracted from only 4 of 16 studies.
5 We have denoted only medium risks in controls for statistically nonsignificant outcomes while denoting low, medium and high risks in controls for statistically significant outcomes.
6 Data related to patients with vertebral fractures were extractable from only 7 of 16 RCTs.
7 Data related to patients with nonvertebral fractures were extractable from only 6 of 16 RCTs.
8 Skeletal-related events data were extractable from only 7 of 16 RCTs.
9 Only 37% (6/16) of trials had adequate allocation concealment. Only 18% (3/16) of trials reported methods of randomization. Similarly, 18% (3/16) of trials reported blinding procedures and personnel who were blinded to the intervention assignment.
10 There was variation in the pain scales used to measure pain; 11 pain relief as defined by the study authors

Table 56: GRADE summary of findings table (harms): What is the most effective method of preventing bone disease in patients with myeloma (bisphosphonates versus placebo or no treatment)? (from Mhaskar et al., 2012).

Note: not all studies included patients with lytic lesions or did not specify bone disease in inclusion criteria

Summary of findings				
No of patients	**Effect**		**Quality**	**Comments**
	Relative (95% CI)	**Absolute**		
Gastrointestinal toxicity				
1689 (6 RCTs)	RR 1.23 (0.95 to 1.6)	86/836 (10.3%) with control, 110/853 (12.9%) with bisphosphonate	low	Limitations in design: serious 1 Serious imprecision 2
Hypocalcaemia				
1002 (3 RCTs)	RR 2.19 (0.49 to 9.74)	2/451 (0.4%) with control, 5/462 (1.1%) with bisphosphonate	Very low	Limitations in design: serious 1 Very serious imprecision 3 Reporting bias 4
Osteonecrosis of jaw				
913 (3 RCTs)	RR 3.99 (0.44 to 35.84)	0/370 (0%) with control, 3/366 (0.8%) with bisphosphonate	Low	Limitations in design: serious 1 Reporting bias 4
1400 (9 observational studies)	-	ONJ incidence range: 0% to 51%	Very low	reporting bias reduced effect for RR >> 1 or RR <<

Summary of findings				
				15 dose response gradient6
Renal dysfunction				
414 (2RCTs)	-	Mean difference: −0.36 (−9.75 to 9.03)	Low	Limitations in design: serious 1 Reporting bias 7

**The basis for the assumed risk (e.g. the median control group risk across studies) is provided in footnotes. The corresponding risk (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI).*

1 Only 37% (6/16) of trials had adequate allocation concealment. Only 18% (3/16) of trials reported methods of randomization. Similarly, 18% (3/16) of trials reported blinding procedures and personnel who were blinded to the intervention assignment. However, sensitivity analyses based on allocation concealment and description of randomization method didn't change the estimates. Hence, the assessment of studies' limitations may represent the poor quality of reporting rather than true biased estimates. Nonetheless, it should be noted that some authors would not downgrade evidence regarding treatment-related harms based on quality of randomization process.

2 The pooled estimate has a wide confidence interval.

3 All the RCTs have estimates with wide confidence intervals.

4 Data related to patients with hypocalcaemia and ONJ was extractable from only 3 of 16 RCTs.

5 ONJ was observed in case control, case series and prospective observational studies and RCTs. Very few studies included consecutive prospective cohort with clear diagnostic criteria and blinded assessment of radiological findings. Therefore, while ONJ is considered a real adverse event, the exact incidence or risk is difficult to assess.

6 While some studies indicate dose response, it could be that ONJ is related to the type of bisphosphonate. So far, no ONJ has been observed in the studies of clodronate.

7 Data related to patients with renal dysfunction were extractable from only 2 of 16 RCTs.

CI: Confidence interval; RR: Risk ratio; ONJ: Osteonecrosis of the jaw

Table 57: GRADE profile: What is the most effective method of preventing bone disease in patients with myeloma (zoledronic acid versus clodronic acid?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	**Design**	**Limitations**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**zoledronic acid**	**clodronic acid**	**Relative (95% CI)**	**Absolute**	**Quality**
Incidence of skeletal related events (follow-up median 3.7 years)											
1	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	265/981 (27%)	346/979 (35.3%)	HR 0.74 (0.62 to 0.87)	78 fewer per 1000 (from 38 fewer to 117	MODERATE

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	zoledronic acid	clodronic acid	Relative (95% CI)	Absolute	Quality
										fewer)	

1 *1 Performance bias and detection bias as study is open-label and not blinded*

2

3 **Table 58: GRADE profile: What is the most effective method of preventing bone disease in patients with myeloma (denosumab**
4 **versus zoledronic acid)?**

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	denosumab	zoledronic acid	Relative (95% CI)	Absolute	Quality
time to first on-study SRE (Better indicated by higher values)											
1	randomised trials	no serious limitations	no serious inconsistency	no serious indirectness	serious[1]	none	93	87	HR of 1.03 95% CI, 0.68 to 1.5	Not reported	MODERATE
overall survival (Better indicated by lower values)											
1	randomised trials	no serious limitations	no serious inconsistency	no serious indirectness	serious[1]	none	93	87	HR of 2.26 (95% CI, 1.13 to 4.50	Not reported	MODERATE

5 *1 no absolute data reported for myeloma*
6

7

Cost effectiveness evidence (see also Appendix F)

The following databases were searched for economic evidence relevant to the PICO: MEDLINE, EMBASE, COCHRANE, NHS EED. Studies conducted from any OECD countries were considered (Guidelines Manual 2014). 463 possibly relevant papers were identified. Of these, 2 full papers relating to this topic were obtained for appraisal. Both papers identified used nearly identical models with differing costs to represent the perspective of a UK and a Canadian healthcare system. Therefore only one paper (Delea et al. 2012) was included in the current review of published economic evidence for this topic (Table 59).

The study was a cost-effectiveness analysis of zoledronic acid (ZOL) versus clodronic acid (CLO) for patients receiving first-line treatment for Stage I-III myeloma. The study reported the results in terms of cost per Quality Adjusted Life Year (QALY) gained and considered a NHS and Personal Social Services (PSS) perspective.

Delea et al is deemed directly applicable to the decision problem that we are evaluating. This is because it took a NHS+PSS perspective and reported health outcomes in terms of QALYs. In addition, quality of life states were scored directly by the relevant patient group using the EQ-5D health questionnaire and valued using UK population preferences.

Potentially serious limitations were identified with Delea et al. Most notably, a potential conflict of interest was identified as the study was funded by and the majority of authors owned stock options in the manufacturer of ZOL (Novartis Pharmaceuticals Corporation).Uncertainty around the utility values for both ZOL and CLO were also not appropriately captured in sensitivity analyses and the range of deterministic sensitivity analyses performed was inadequate.

The base case suggested that treating with ZOL over CLO would cost £5443 per QALY gained although this varied from ZOL being dominant (less costly, more effective) to £19,378 per QALY gained during deterministic sensitivity analysis. All deterministic sensitivity analyses resulted in an ICER below £20,000 per QALY. The analysis was conducted at a time when ZOL was still ‘on patent’. At the time this evidence review was conducted this was no longer the case and the cost of ZOL was likely to be significantly less than that used in the analysis. When a 50% reduction in drug cost was assumed, a conservative estimate of current NHS+PSS costs, ZOL became dominant (both cost saving and health improving) compared to CLO.

Deterministic and probabilistic sensitivity analyses suggested this result was robust with ZOL having a 90% and 94% probability of being cost-effective at a willingness to pay threshold of £20,000 and £30,000 respectively although uncertainty around utility values for the interventions were not adequately captured.

Table 59: Modified GRADE profile: included economic studies

Study	Population	Comparators	Costs	Effects	Incr costs	Incr effects	ICER	Uncertainty	Applicability	Limitations
Delea et al. 2012 UK	Patients receiving first-line treatment for newly diagnosed Stage I-III myeloma	Clodronic acid 1600mg daily (CLO)	£8176	2.68 QALYs	Reference			One-way Sensitivity Analysis One-way sensitivity analyses were conducted with incremental cost per QALY of ZOL compared to CLO ranging from ZOL being dominant (less costly and more effective) and £19,378 per QALY Probabilistic Sensitivity Analysis There was a 90% and 94% probability that ZOL was cost-effective at a willingness-to-pay threshold of £20,000 and £30,000 respectively.	Directly Applicable	Potentially Serious Limitations .
		Zoledronic acid 4mg intravenously every 3-4 weeks (ZOL)	£9829	2.99 QALYs	£1653	0.30 QALYs	£5443 per QALY			
	Comments:									

Recommendations	**To prevent bone disease, offer people with myeloma:** • **zoledronic acid, or** • **disodium pamidronate, if zoledronic acid is contraindicated or not tolerated, or** • **sodium clodronate, if zoledronic acid and disodium pamidronate are contraindicated, not tolerated or not suitable.** **Consider immediately referring people with myeloma for dental assessment and treatment before starting zoledronic acid or disodium pamidronate.** **For people who need urgent myeloma treatment, consider referring for dental assessment and treatment as soon as possible after they start treatment**
Relative value placed on the outcomes considered	The Guideline Committee considered the outcomes of skeletal related events, adverse events, quality of life, overall survival, progression-free survival, pain, need for radiotherapy and hypercalcaemia to be the most relevant in identifying the most effective method of preventing bone disease in patients with myeloma. Of these, evidence was found for all outcomes except need for radiotherapy and quality of life. The Guideline Committee considered skeletal related events and osteonecrosis of the jaw to be the most important as they have the most impact on patients quality of life, and overall survival when determining their recommendations.
Quality of the evidence	The clinical evidence for overall survival ranged from low to moderate quality as assessed by GRADE. There was moderate quality evidence for skeletal related events and very low to low quality evidence for osteonecrosis of the jaw. The Guideline Committee noted that the Mhasker et al 2012 systematic review had included bisphosphonates that are known to be ineffective at preventing SREs in myeloma patients (etidronate and ibandronate). They considered that the inclusion of these drugs had probably diluted the overall results and made the other bisphosphonates appear to be less effective than they really are. The health economic evidence comparing bisphosphonates was assessed as directly applicable but with potential serious limitations because there were potential conflicts of interest and the uncertainty has not been adequately captured. The Guideline Committee noted that no clinical evidence had been found on the optimal duration and frequency of bisphosphonate treatment. They therefore recommended further research was conducted in this area.
Trade off between clinical benefits and harms	The Guideline Committee agreed, based on the evidence, that zoledronic acid was the more effective than sodium clodronate for preventing bone disease. It was also the opinion of the Guideline Committee, based on their clinical experience, that zoledronic acid was more effective than disodium pamidronate. This was because it has been shown to have statistically significant overall and progression free survival advantage whereas only a trend has been shown for disodium pamidronate. It was noted that disodium pamidronate and sodium clodronate were also effective to a lesser extent. The Guideline Committee acknowledged that no direct evidence had been found comparing disodium pamidronate and

	sodium clodronate. However, based on their clinical experience they agreed that disodium pamidronate was more effective than sodium clodronate and therefore should be recommended as second option if zoledronic acid is contraindicated or not tolerated. The Guideline Committee also agreed to recommend sodium clodronate as a third option for those people who require an oral option. The Guideline Committee noted, based on the evidence, that since denosumab had shown a survival advantage compared with bisphosphonates, it could be a potential treatment option for people who were intolerant to bisphosphonates. However, they agreed that the evidence was limited (coming from 1 trial of 180 patients), no adverse event data were available and there was no evidence on cost effectiveness. Given this and the fact that denosumab does not have a licence for use in patients with myeloma the Guideline Committee decided not to make any recommendations for this intervention. No evidence was identified that examined the interventions of calcium supplements, vitamin D supplements, bone anabolic therapy or exercise for preventing bone disease in myeloma patients. Therefore the Guideline Committee were not able to make any recommendations on these interventions. The Guideline Committee considered the potential benefits would be that patients with myeloma would be given the most efficacious bisphosphonate therapy, leading to a reduction in skeletal related events, reduced pain and improved quality of life and improved overall survival. However the Guideline Committee acknowledged, based on the evidence, that use of bisphosphonates is associated with adverse effects, specifically osteonecrosis of the jaw (particularly with intravenous bisphosphonates), hypocalcaemia and renal dysfunction. The Guideline Committee considered, based on their clinical experience, that dental extractions are a risk factor for developing osteonecrosis of the jaw so it was important for patients to have a dental assessment and treatment before beginning bisphosphonate treatment. Since the impact of bisphosphonates is higher when treatment is started sooner, the Guideline Committee agreed that referral for such a dental assessment should be immediate. The Guideline Committee also acknowledged that there are clinical situations in which treatment with bisphosphonates takes priority over dental assessment and intervention.
Trade-off between net health benefits and resource use	The Guideline Committee noted that one published economic evaluation had been identified comparing the cost effectiveness of zoledronic acid and sodium clodronate. The results of the probabilistic sensitivity analysis in Delea et al. 2012 showed that zoledronic acid had a 90% probability of being cost effective at a willingness to pay threshold of £20,000/QALY. Although there were some limitations with this paper, it was considered to be methodologically sound with robust results. The Guideline Committee agreed that zoledronic acid was the most cost effective option for preventing bone disease in patients with myeloma. The recommendation will lead to a reduction in use of less costly disodium pamidronate and sodium clodronate and a reduction in treatment costs associated with their adverse events. The Guideline Committee acknowledge that there will be a greater

	cost through an increase in urgent referrals for dental assessment in either primary or secondary care. This will be balanced out against a reduction in costs associated with a lower incidence of osteonecrosis of the jaw. No published economic evaluations were identified on the other interventions of interest.
Other	The Guideline Committee noted that not all centres currently have pathways set up for urgent dental evaluation and treatment prior to starting bisphosphonate treatment.

Research recommendation	**A randomised controlled trial should be carried out, comparing monthly zoledronic acid indefinitely with zoledronic acid for fixed duration in patients with myeloma. Outcomes of interest are skeletal-related events, progression-free survival, overall survival, utility of bone biomarkers, incidence of osteonecrosis of the jaw, quality of life and resource use.**
Why is this important?	There is good quality evidence to support the use of zoledronic acid to prevent bone disease in people with myeloma. However, the optimal frequency and duration of treatment is not clearly defined and needs further research, particularly given the quality-of-life implications for people needing regular, life-long visits to hospital.

8.2 Managing non-spinal bone disease

Bone pain, pathological fractures, lytic bone lesions and hypercalcaemia are the main skeletal related events (SREs) in non-spinal bone disease resulting from myeloma. The management of these SREs is multidimensional and depends on several factors including site and extent of bone involvement, symptoms, performance status, co-morbidities and life expectancy in addition to patient circumstances and preferences.

One or more modalities of treatment, in addition to combination chemotherapy, are usually required to treat non spinal bone disease. These may include radiotherapy, osteoclast inhibitors such as bisphosphonates and orthopaedic surgical intervention. However, there may be uncertainty around the optimal treatment and/or combinations of treatment. Treatment decisions will involve multidisciplinary professionals including a clinical haematologist, clinical oncologist, radiologist, orthopaedic surgeon, pain control and palliative care specialist, physiotherapist and clinical nurse specialist.

Clinical question: What are the most effective treatments (other than chemotherapy) for non-spinal bone disease in patients with myeloma (including radiotherapy and surgical intervention)?

Clinical evidence (see also Appendix G)

See Tables 60-61.

Radiotherapy

Very low quality evidence from one observational study of radiotherapy for non-spinal bone disease in 27 patients with multiple myeloma was identified (Catell et al., 1998). The study aimed to examine the effectiveness of radiotherapy to the symptomatic portion of a long bone for palliation. The outcome assessed was progressive disease and it was found that 15% of patients developed progressive disease.

Surgery

Very low quality evidence from three observational studies of surgery for non-spinal bone disease in patients with multiple myeloma was identified (Chang et al., 2001; Natarajan et al., 2007; Papagelopoulos et al., 1997). Using data from all 3 studies the complication rate from surgery was 25.9%; the main issues being intra-operative complications and wound related complications. From 2 studies the implant failure rate was low (6.9%) and there was improvement in both pain (45 – 91% of patients reporting complete pain relief) and ambulatory status (40 – 64% of patients not requiring support for moving around/walking).

Two studies assessed overall survival post surgery. One study of 22 patients (Chang et al, 2001) found the mean overall survival to be 19 months (range 3 – 60 months). Another study of 46 patients (Papagelopoulos et al., 1997) found the median overall survival to be 18 months (range 7 days – 19.9 years).

One study of 9 patients (Natarajan et al., 2007) assessed functional outcome which was determined to be good or excellent in 67% of patients.

Interventional pain management, bisphosphonates, denosumab and supportive care

We did not find evidence for these interventions.

Table 60: GRADE profile: What are the most effective treatments for non-spinal bone disease in patients with myeloma (radiotherapy)?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	radiotherapy	control	Relative (95% CI)	Absolute	Quality
progressive disease											
1	observational studies	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	4/27 (14.8%)	n/a	-	-	VERY LOW

1 retrospective case series (no comparator); 2 small sample size limits precision of results

Table 61: GRADE profile: What are the most effective treatments for non-spinal bone disease in patients with myeloma (orthopaedic surgery)?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limit-ations	Inconsis-tency	Indirectness	Impre-cision	Other considerations	orthopaedic surgery	control	Relative (95% CI)	Absolute	Quality
overall survival											
2	observational studies	serious[1]	serious[2]	no serious indirectness	no serious imprecision	none	68	n/a	-	Study 1 (n=22): mean overall survival 19 months (range 3 – 60 months) Study 2 (n=46): median overall survival 18 months (range 7 days – 19.9 years)	VERY LOW
implant failure											
2	observational studies	serious[1]	serious[2]	no serious indirectness	no serious imprecision	none	5/72 (6.9%)	n/a	-	-	VERY LOW
complication rate											

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limit-ations	Inconsis-tency	Indirectness	Impre-cision	Other considerations	orthopaedic surgery	control	Relative (95% CI)	Absolute	Quality
3	observational studies	serious[1]	serious[2]	no serious indirectness	no serious imprecision	none	21/81 (25.9%)	n/a	-	-	VERY LOW
pain relief											
2	observational studies	serious[1]	serious[2]	no serious indirectness	no serious imprecision	none	67	n/a	-	Complete pain relief: 45 – 91%	VERY LOW
ambulatory status											
2	observational studies	serious[1]	serious[2]	no serious indirectness	no serious imprecision	none	57	n/a	-	Full weight bearing/used no support: 40 – 64%	VERY LOW
functional outcome											
1	observational studies	serious[1]	no serious inconsis-tency	no serious indirectness	serious3	none	9	n/a	-	Functional outcome was good or excellent in 67% of patients	VERY LOW

1 *1 retrospective case series (no comparator); 2 the different studies use different surgical methods; 3 small sample size limits precision of results*

2

1 Cost effectiveness evidence

2 A literature review of published cost-effectiveness analyses did not identify any relevant
3 papers for this topic. Whilst there were potential cost implications of making
4 recommendations in this area, other questions in the guideline were agreed as higher
5 priorities for economic evaluation. Consequently no further economic modelling was
6 undertaken for this question.

7

Recommendations	**Offer people with myeloma and non-spinal bone disease who have not already started bisphosphonates:** • **zoledronic acid, or** • **disodium pamidronate, if zoledronic acid is contraindicated or not tolerated, or** • **sodium clodronate, if zoledronic acid and disodium pamidronate are contraindicated, not tolerated or not suitable.** **Assess the risk of fracture (in line with the NICE guideline on assessing the risk of fragility fractures in osteoporosis) in people with myeloma and non-spinal bone disease.** **Consider surgical stabilisation followed by radiotherapy for non-spinal bones that have fractured or are at high risk of fractures.** **Consider radiotherapy for non-spinal bones that have fractured or are at high risk of fracture if surgical intervention is unsuitable or not immediately needed.** **Consider radiotherapy for people with myeloma and non-spinal bone disease who need additional pain relief if:** • **chemotherapy and initial pain management has not led to prompt improvement in pain control.** • **chemotherapy is unsuitable and current pain medication is not working**. **Consider re-treatment with radiotherapy if pain recurs or if there is regrowth of a previously treated lesion.** **Consider seeking advice from or referral to specialists in palliative care or pain medicine for people with complex non-spinal bone disease.**
Relative value placed on the outcomes considered	The Guideline Committee considered the outcomes of health related quality of life, progression free survival, overall survival, adverse events, pain control, mobility/dependency and patient expectation to be the most relevant to identify the most effective treatments for non-spinal bone disease in patients with myeloma. Of these, evidence was found for all outcomes except health related quality of life, adverse events and patient expectation. When drafting the recommendations the Guideline Committee considered progression free survival, overall survival and mobility/dependency to be the most important outcomes for the intervention bisphosphonates. For radiotherapy the Guideline Committee considered pain control and progression free survival to be the most important outcomes. For surgery the Guideline

	Committee considered pain relief and mobility to be the most important outcomes, together with a low risk of operative failure. The Guideline Committee considered the additional outcomes of stabilisation and prevention of fracture. These were identified as outcomes, following discussion with expert advisors and consensus among the Guideline Committee, for non-spinal bone disease which is not amenable to surgery and also for those patients who are unsuitable for surgery.
Quality of the evidence	The quality of the evidence was assessed using GRADE methodology and appropriate NICE checklists to be of very low quality for all outcomes. The Guideline Committee noted that only 4 studies had been included in the evidence review. A number of studies concerning potentially relevant interventions were excluded as the population was mixed including both spinal and non-spinal bone disease and it was not possible to extract the data specifically for non-spinal bone disease. Furthermore many studies examining the relevant interventions were not specific to myeloma - myeloma patients were often included in larger studies with other malignancies and no sub-group analysis for myeloma was provided. As a consequence there was very little directly applicable evidence for the interventions of interest. The Guideline Committee noted that the 4 studies that were included in the evidence review were small retrospective case series. They were observational studies with no comparative data. In addition, evidence was only identified for two of the interventions of interest – radiotherapy and surgery. The Guideline Committee noted that the evidence for radiotherapy came from a single small non-randomised study on radiotherapy with no useful clinical outcomes. The evidence for surgery came from three small, case series which were non-comparative. Owing to these issues the Guideline Committee agreed not to use this evidence to make recommendations. Instead they based their recommendations on their clinical expertise and experiences as well as information provided by the expert advisors. The Guideline Committee made a research recommendation for radiotherapy as although there was Guideline Committee agreement (informed by expert advice) about the role for radiotherapy in the management of non-spinal bone disease, there was uncertainty on the optimal schedule. Given the lack of evidence on supportive care as an intervention to manage non-spinal bone disease, the Guideline Committee also recommended research into the most effective strategy early (‘pro-active’) versus late (‘reactive’) to involve specialist palliative care for pain relief.
Trade off between clinical benefits and harms	Based on their clinical experience the Guideline Committee agreed that it was important to assess fracture risk in people with non-spinal bone disease as this would affect their treatment options. Based on expert advice and their clinical experience the Guideline Committee acknowledged that there was a role for surgery in the stabilisation of non-spinal bones that have fractured or that are at high risk of fracture. Based on expert advice and their experience the Guideline Committee acknowledged that the main evidence base for the use of radiotherapy in bone disease comes from people with other solid tumours. In these people, surgical stabilisation is the most

	effective treatment for bones that have fractured or are at high risk of fracture. However, for some people and in some anatomical sites, surgery is not suitable, in which case radiotherapy is an option. Radiotherapy is also effective at managing pain where chemotherapy and initial pain management have failed. The Guideline Committee noted that whilst the evidence for the use of radiotherapy in the management of non-spinal bone disease comes from solid tumours, there is recognition that myeloma is more radiosensitive than most solid tumours. Given this the Guideline Committee agreed to recommend that radiotherapy is used in these instances. Based on expert advice, the Guideline Committee also noted that the dose of radiotherapy normally used in myeloma means retreatment is feasible. Consequently they decided to recommend this. Despite the lack of evidence on supportive care, the Guideline Committee agreed, based on their clinical experience that advice should be sought or patients referred to specialist palliative care/pain medicine in complex cases as these patients may have different pain control needs. The Guideline Committee concluded that the potential benefits would be improved pain control, improved quality of life, reduced fracture risk, improved healing and reduced analgesic toxicity. The Guideline Committee considered that the potential harms could be a risk of radiation toxicity as well as risks associated with surgical intervention. The Guideline Committee agreed that the recommendations take into account that surgery may not be suitable for all patients and where this is the case radiotherapy has been recommended instead, thereby reducing the potential risks associated with surgery. In addition, the Guideline Committee considered that radiation toxicity is lower outside the spine and skull, meaning re-treatment of long bones is a practical option.
Trade-off between net health benefits and resource use	The Guideline Committee noted that no relevant published economic evaluations had been identified and no additional economic analysis had been undertaken in this area. The Guideline Committee agreed that there are unlikely to be additional costs as a result of the recommendations made on surgery and radiotherapy as these are already a standard of care for the management of pain and fracture. The Guideline Committee noted that although radiotherapy re-treatment is not currently standard practice, it is likely to be single fraction so any increase in cost will be small. Referral to specialist palliative care and pain medicine is currently variable. The Guideline Committee's recommendations are likely to result in some people being referred earlier (with an associated increase in costs). However these costs may be offset by reduced length of stay, avoidance of later complications and emergency admissions.

1

Research	**A randomised controlled trial should be carried out comparing**

recommendation	**single versus multiple fraction radiotherapy in patients with myeloma and non-spinal bone disease who are not amenable to surgery. Outcomes of interest: fracture, pain, quality of life, progression free survival, overall survival.**
Why is this important?	Myeloma is a more radio sensitive tumour than most solid tumours and as such radiotherapy can provide pain relief and tumour destruction which may assist in bone healing in people with non-spinal bone disease. However, there is no evidence or consensus on the optimal dose and fractionation schedule.

Research recommendation	**A randomised controlled trial should be carried out comparing early (pro-active) referral to specialist palliative care versus standard care (reactive referral) for management of non-spinal bone related pain in patients with myeloma. Outcomes of interest: pain, quality of life, progression free survival, overall survival, carer experience, resource use.**
Why is this important?	Unlike many other cancers, myeloma is not curable so palliation is very important. Access to specialist palliative care services for people with myeloma has increased but there is no clear guidance on when to refer. Studies in other cancers have shown patient experience is improved (in terms of quality of life and reduction in number of inappropriate interventions) and there is a survival benefit from early referral to specialist palliative care services. However the potential health economic consequences of doing this could be substantial. Further research is needed to help resolve this uncertainty.

8.3 Managing spinal bone disease

When myeloma affects the vertebral spine, it can sometimes lead to collapse of one or more vertebrae, which can cause very serious consequences including severe pain. Spinal bone disease may also be associated with soft tissue growth into surrounding tissues. Where there is spinal cord compression, weakness or paralysis of the lower limbs and loss of bladder and bowel control can rapidly become permanent without urgent treatment, and this has devastating consequences.

The core aims of the management of spinal bone disease in myeloma are decompression, stabilization and pain control. Decompression is covered by recommendations made in the NICE guideline on metastatic spinal cord compression. Interventions used for stabilisation and pain control are drugs (analgesics), radiotherapy, external bracing/orthotics, vertebroplasty, balloon kyphoplasty and in severe cases, open spinal surgery. The involvement of multiple vertebrae further complicates the clinical decisions to be made.

There is uncertainty over the effectiveness of the different treatments for management of spinal bone disease in patients with myeloma and the circumstances and order in which they should be used.

There is variation across the UK in terms of access to specialist spinal surgery, including rehabilitation and also to cement augmentation (vertebroplasty and kyphoplasty). The optimal configuration of local and regional haematology services for the management of myeloma is covered in section 5.1.

Clinical question: Excluding chemotherapy, which treatments are effective for spinal bone disease in patients with myeloma, and in which circumstances and order should they be offered?

Clinical evidence (see also Appendix G)

See Tables 62-67.

Bisphosphonates

One systematic review and network meta-analysis of bisphosphonates for the prevention of skeletal-related events in myeloma (20 RCTs, 6692 patients) was identified (Mhaskar et al., 2012). In six trials it was specified that the inclusion criteria included the presence of at least one osteolytic lesion. However, it was not specified if the lesions were spinal or non-spinal, which limits relevance to the review question.

Pooled results showed no direct effect of bisphosphonates on overall survival compared with placebo or no treatment (HR 0.96, 95% CI 0.82 to 1.13; P = 0.64). However, there was a statistically significant heterogeneity among the included RCTs (I2 = 55%, P = 0.01) for OS (Low quality).

Pooled analysis did not demonstrate a beneficial effect of bisphosphonates compared with placebo or no treatment in improving PFS (HR 0.70, 95% CI 0.41 to 1.19; P = 0.18) There was no heterogeneity among trials reporting PFS estimates (I2 = 35%, P = 0.20) (Very low quality).

Pooled analysis demonstrated a beneficial effect of bisphosphonates compared with placebo or no treatment on prevention of pathological vertebral fractures (RR 0.74, 95% CI 0.62 to 0.89; I2 = 7%) (moderate quality), skeletal-related events (SRE) (RR 0.80, 95% CI 0.72 to 0.89; I2 = 2%) (moderate quality) and amelioration of pain (RR 0.75, 95% CI 0.60 to 0.95; I2 = 63%) (very low quality).

The network meta-analysis did not show any difference in the incidence of osteonecrosis of the jaw (5 RCTs, 3198 patients) between bisphosphonates. Rates of osteonecrosis of the jaw in observational studies (9 studies, 1400 patients) ranged from 0% to 51% (very low quality). The pooled results (6 RCTs, 1689 patients) showed no statistically significant increase in frequency of gastrointestinal symptoms with the use of bisphosphonates compared with placebo or no treatment (RR 1.23, 95% CI 0.95 to 1.60; P = 0.11) (low quality).

The pooled results (3 RCTs, 1002 patients) showed no statistically significant increase in frequency of hypocalcaemia with the use of bisphosphonates compared with placebo or no treatment (RR 2.19, 95% CI 0.49 to 9.74). The network meta-analysis did not show any differences in the incidence of hypocalcaemia, renal dysfunction and gastrointestinal toxicity between the bisphosphonates used (low quality).

Denosumab

One randomised trial including 180 myeloma patients with at least 1 bone metastases or osteolytic lesion compared denosmab with zoledronic acid (Henry et al., 2011). The effect of denosumab on time to first on-study skeletal-related event (including fracture and spinal cord compression) relative to zoledronic acid resulted in a HR of 1.03 (95% CI: 0.68 to 1.57) (low quality).

An ad hoc analysis examining overall survival demonstrated an HR of 2.26 (95% CI: 1.13 to 4.50) (low quality).

Vertebral augmentation (kyphoplasty/vertebroplasty)

Very low quality evidence from one randomised trial of 134 patients (49 with multiple myeloma) compared balloon kyphoplasty with non-surgical management for painful vertebral body compression fractures (Berenson et al., 2011). Back-specific functional status (as measured by the Roland-Morris disability questionnaire) at 1 month was reduced in the kyphoplasty group by 8.3 points (95% CI -6.4 to -10.2), and by 0.1 points (95% CI -0.8 to 1)

in the control group. Patients in the kyphoplasty group also had significant improvements in quality of life, back pain and performance status, which were not seen in the control group. One patient in the kyphoplasty group had cement leakage and device-related vertebral compression fracture.

Very low quality evidence from one pooled analysis of case series of kyphoplasty (nine studies) and vertebroplasty (12 studies) or both (two studies) was identified, including a total of 923 patients (Khan et al., 2014). There was a decrease in pain from baseline across all time periods (≤1 week, 1 week to 1 year, >1 year). There were no differences between kyphoplasty and vertebroplasty studies in terms of mean pain reduction from baseline to the three time periods presented. There was no significant decrease in disability scores (as measured by the Owestry Disability Index) from baseline to any of the time periods. The most common complication was new vertebral fractures at untreated vertebral bodies. This occurred in 7.3% (42/576) of vertebroplasty patients and 6.8% (25/367) kyphoplasty patients (p=0.78).

Low quality evidence from three further case series (Erdem et al., 2013a; Simony et al, 2014; Ha et al, 2015) of vertebral augmentation in 424 myeloma patients reports typical reduction in pain from baseline to 1-month post-op of around 4 points (on a scale of 0-10) ($p<0.001$). One study (Erdem et al., 2013a) reports that no significant differences in pain improvements between the type of procedure performed (kyphoplasty versus vertebroplasty or kyphoplasty+vertebroplasty) for pain relief or improvement in activity.

One observational study including 39 patients with myeloma undergoing percutaneous vertebroplasty reported median overall survival of 20 months (range 2-91), with estimated 5-year survival of 40% (Chew et al., 2011) (very low quality).

Two observational studies (total 77 patients) of radio-frequency targeted vertebral augmentation in multiple myeloma both reported reductions in mean pain scores and improvements in disability post-procedure (Erdem et al., 2013b; Orgera et al., 2014). 5 patients (6.5%) had cement leakage (very low quality). One study reported that there were significant differences in pain reduction and complications between radiofrequency ablation and vertebroplasty compared with vertebroplasty alone (Orgera et al., 2014) (very low quality).

Surgery

Very low quality evidence from three observational studies of surgical intervention for myeloma bone disease (including both spinal and non-spinal disease) was identified (Zadnik et al., 2015; Zeifang et al., 2005; Utzschneider et al., 2011). Surgical interventions included posterior decompression-stabilisation, decompression alone, and endoprosthesis. Median survival was 3.9 years and 6.6 years. The most common adverse event related to wound complications.

Radiotherapy

Very low quality evidence from three observational studies of radiotherapy for skeletal lesions in multiple myeloma was identified (Budak et al., 1991; Yaneva et al., 2006; Balducci et al., 2011). Two studies reported median overall survival of 36 months and 32 months. Three studies reported that 55% (248/521) of patients reported good or complete relief of pain after treatment. One study reported that 78% (62/79) of patients reported improvements in motor function. Grade 3 or 4 adverse events were reported in 0.8% (3/371) patients.

Table 62: GRADE profile: Excluding chemotherapy, which treatments are effective for spinal bone disease in patients with myeloma (vertebroplasty versus kyphoplasty)?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Vertebro-plasty	Kypho-plasty	Relative (95% CI)	Absolute	Quality
Pain (from baseline up to 1 week post-procedure) (measured with: Visual Analogue Scale; Brief Pain Inventory; SF-36; Better indicated by lower values)											
11[1]	observational studies	serious[2]	no serious inconsistency	no serious indirectness	no serious imprecision	none	Not reported	Not reported	-	For vertebroplasty and kyphoplasty: Mean pain reduction 4.8±0.56	VERY LOW
Pain (from baseline to >1yr post-procedure) (measured with: Visual Analogue Scale; Brief Pain Inventory; SF-36; Better indicated by lower values)											
14[1]	observational studies	serious[2]	no serious inconsistency	no serious indirectness	no serious imprecision	none	Not reported	Not reported	-	For vertebroplasty and kyphoplasty: Mean pain reduction 4.4±0.48	VERY LOW
Activities of daily living (change from baseline up to 1 week post-procedure) (measured with: Owestry Disability Index; scale 0-100; Better indicated by lower values)											
3[1]	observational studies	serious[2]	no serious inconsistency	no serious indirectness	no serious imprecision	none	Not reported	Not reported	-	Mean decrease 39.2 (16.3 to 75) P=0.37	VERY LOW
Activities of daily living (change from baseline to >1 year post-procedure) (measured with: Owestry Disability Index; scale 0-100; Better indicated by lower values)											
4[1]	observational	serious[2]	no serious	no serious	no serious	none	Not	Not	-	Mean	VERY

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Vertebro-plasty	Kypho-plasty	Relative (95% CI)	Absolute	Quality
	studies		inconsistency	indirectness	imprecision		reported	reported		decrease 46.5 (14.5 to 75) P=0.88	LOW
Infection											
1[3]	observational studies	serious[2]	no serious inconsistency	no serious indirectness	no serious imprecision	none	1/576 (0.2%)	0/367 (0%)	P=0.64	-	VERY LOW
Pulmonary embolism											
1[3]	observational studies	serious[2]	no serious inconsistency	no serious indirectness	no serious imprecision	none	0/576 (0%)	1/367 (0.3%)	P=0.21		VERY LOW
Myocardial Infarction											
1[3]	observational studies	serious[3]	no serious inconsistency	no serious indirectness	no serious imprecision	none	0/576 (0%)	1/367 (0.3%)	P=0.21		VERY LOW
Vertebral compression fracture at untreated levels											
1[3]	observational studies	serious[2]	no serious inconsistency	no serious indirectness	no serious imprecision	none	42/576 (7.3%)	25/367 (6.8%)	P=0.78		VERY LOW
Neurologic symptoms requiring revision surgery											
1[3]	observational studies	serious[2]	no serious inconsistency	no serious indirectness	no serious imprecision	none	0/576 (0%)	2/367 (0.5%)	P=0.08		VERY LOW
Transient perioperative pain											
1[3]	observational studies	serious[2]	no serious inconsistency	no serious indirectness	no serious imprecision	none	4/576 (0.7%)	2/367 (0.5%)	P=0.78		VERY LOW
Spinal cord compression											
0	no evidence										
Progression-free survival											
0	no evidence										
Overall survival (Kaplan-Meier curve)											

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Vertebro-plasty	Kypho-plasty	Relative (95% CI)	Absolute	Quality
1[4]	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	serious[5]	none	39	n/a		Median survival= 20 months (range 2-91)	VERY LOW
Performance status											
0	no evidence										
Dependency											
0	no evidence										
Health-related quality of life											
0	no evidence										
Pain (at 1 month) (follow-up 1 months; measured with: Visual Acuity Scale; range of scores: 0-10; Better indicated by lower values)											
1[6]	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none	351		-	Mean reduction 4.2 (4.0 to 4.5)[7]	LOW
Improvement in activity (Proportion of patients scoring 0-1 (no limitations); range of scores 0-6; Better indicated by lower values)											
1[6]	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none	354			28% at baseline vs 59% post-procedure	LOW

1 As reported in systematic review by Khan et al. (2014); 2 Prospective and retrospective case series. Studies differed in adjunctive therapy, disease stage and other factors. Small sample size in individual studies; 3 As reported in systematic review by Khan et al. (2014). Number of participants not reported; 4 Chew et al. (2011); 5 Small number of participants with Myeloma (n=39) limits precision of results; 6 Erdem et al. (2013a); 7 Average reduction of pain from baseline to 1 month

Table 63: GRADE profile: Excluding chemotherapy, which treatments are effective for spinal bone disease in patients with myeloma (balloon kyphoplasty for painful vertebral compression fractures)?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Balloon kyphoplasty	Non-surgical management	Relative (95% CI)	Absolute	Quality
Vertebral collapse											
0	no evidence										
Spinal cord compression											
0	no evidence										
Health-related quality of life (follow-up 1 month; measured with: SF-36 Physical components scale; range of scores: 0-100; Better indicated by higher values)											
1[1]	randomised trials	serious[2]	no serious inconsistency	serious[3]	serious[4]	none	65	52	-	MD 8.4 higher (7.7 to 9.1 higher)5	VERY LOW
Progression-free survival											
0	no evidence										
Overall survival (mortality rate)											
1[1]	randomised trials	serious[2]	no serious inconsistency	serious[3]	serious[4]	none	29/108 (26.9%)6	6/26 (23.1%)	RR 1.16 (0.54 to 2.51)	37 more per 1000 (from 106 fewer to 348 more)	VERY LOW
Performance status (follow-up 1 month; measured with: Karnofsky performance status; range of scores: 0-100; Better indicated by higher values)											
1[1]	randomised	serious[2]	no serious	serious[3]	serious[4]	none	65	52	-	MD 15.3	VERY

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Balloon kyphoplasty	Non-surgical management	Relative (95% CI)	Absolute	Quality
	trials		inconsistency							higher (13.5 to 17.1 higher)[5]	LOW
Quality of life (follow-up 1 month; measured with: SF-36 mental components scale; range of scores: 0-100; Better indicated by higher values)											
1[1]	randomised trials	serious[2]	no serious inconsistency	serious[3]	serious[4]	none	65	52	-	MD 11.1 higher (10.7 to 11.5 higher)[5]	VERY LOW
Pain control (follow-up 7 days; measured with: Numerical rating scale; range of scores: 0-10; Better indicated by lower values)											
1[1]	randomised trials	serious[2]	no serious inconsistency	serious[3]	serious[4]	none	65	52	-	MD 3.5 lower (3.8 to 3.2 lower)[7]	VERY LOW
Pain control (follow-up 1 month; measured with: Numerical rating scale; range of scores: 0-10; Better indicated by lower values)											
1[1]	randomised trials	serious[2]	no serious inconsistency	serious[3]	serious[4]	none	65	52	-	MD 3.3 lower (3.6 to 3.0 lower) [7]	VERY LOW
Reduced activity days caused by back pain (follow-up 1 month; Better indicated by lower values)											
1[1]	randomised trials	serious[2]	no serious inconsistency	serious[3]	serious[4]	none	65	52	-	MD 6.3 lower (6.8 to 5.8 lower)[5]	VERY LOW
Back-specific physical functioning (follow-up 1 month; measured with: Roland-Morris Disability Questionnaire (RDQ); range of scores: 0-24; Better indicated by lower values)											

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Balloon kyphoplasty	Non-surgical management	Relative (95% CI)	Absolute	Quality
1[1]	randomised trials	serious[2]	no serious inconsistency	serious[3]	serious[4]	none	65	52	-	MD 8.4 lower (7.6 to 9.2 lower) [5]	VERY LOW
Dependency											
0	no evidence										
Adverse events (follow-up 1 month; Adverse events in first month)											
1[1]	randomised trials	serious[2]	no serious inconsistency	serious[3]	serious[4]	none	26/70 (37.1%)	19/64 (29.7%)	RR 1.25 (0.77 to 2.03)	74 more per 1000 (from 68 fewer to 306 more)	VERY LOW
Serious adverse events (serious AEs after 1 month until study end)											
1[1]	randomised trials	serious[2]	no serious inconsistency	serious[3]	serious[4]	none	37/70 (52.9%)	8/26 (30.8%)	RR 1.72 (0.93 to 3.19)	222 more per 1000 (from 22 fewer to 674 more)	VERY LOW
Pain (follow-up 3 months; assessed with Visual Analogue Scale 0 to 10; better indicated by lower score)											
1[8]	observational study	serious[9]	no serious inconsistency	no serious indirectness	serious[10]	none	69	n/a	-	Mean pain score decreased from 7.9 at baseline to 2.5 post-	VERY LOW

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Balloon kyphoplasty	Non-surgical management	Relative (95% CI)	Absolute	Quality
										procedure	

1 Berenson et al. (2011); 2 Sponsors of the study (Medtronic Spine LLC) contributed to study design, data collection and analysis.; 3 68% of kyphoplasty group and 56% of control group had cancer diagnosis other than myeloma which limits relevance of study to the review question; 4 Small sample size limits precision of results; 5 Mean change in intervention group. Statistically significant difference at one month in comparison with control group; 6 Intervention group includes kyphoplasty + crossover patients; 7 Difference in change from baseline between control and kyphoplasty group; 8 Papanastassiou et al. (2014); 9 Retrospective case series; 10 Small sample size (n=69) limits precision of results

Table 64: GRADE profile: Excluding chemotherapy, which treatments are effective for spinal bone disease in patients with myeloma (radiofrequency targeted vertebral augmentation)

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Radiofrequency targeted vertebral augmentation	control	Relative (95% CI)	Absolute	Quality
Vertebral collapse											
0	no evidence										
Spinal cord compression											
0	no evidence										
Health-related quality of life											
0	no evidence										
Progression-free survival											
0	no evidence										
Overall survival											
0	no evidence										
Performance status											
0	no evidence										
Pain control at 6 months versus baseline (assessed with Visual Analogue Scale, 0-10; better indicated by lower value)											
1[1]	observational	no serious	no serious	no serious	serious[2]	none	41	n/a	-	Mean	VERY

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	Radiofrequency targeted vertebral augmentation	control	Relative (95% CI)	Absolute	Quality
	studies	limitations	inconsistency	indirectness						decrease 5.6±2.8	LOW
Pain control at 24h post-procedure versus baseline (assessed with Visual Analogue Scale, 0-10; better indicated by lower value)											
1[3]	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	serious[2]	none	36	n/a	-	Mean score decrease from 9.1±0.9 to 3.4±1.2[4]	VERY LOW
Adverse events (Cement leakage)											
2[5]	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	serious[2]	none	5/77 (6.5%)	n/a	-	-	VERY LOW
Patient activity (Proportion of patients with fully unassisted ambulation at baseline and 6-months)											
1[1]	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	serious[2]	none	41	n/a	-	Increased from 31% to 63%	VERY LOW
Disability at 24h post-procedure versus baseline (measured with: Roland-Morris disability questionnaire; range of scores: 0-24; Better indicated by lower values)											
1[3]	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	serious[2]	none	36	n/a	-	Mean score decrease from 19.8 ±1.5 to 9.6 ±1.24	VERY LOW
Dependency											
0	no evidence										

1 *1 Erdem et al. (2013b); 2 Small number of participants limits precision of results; 3 Orgera et al. (2014); 4 Mean score for RFA vertebroplasty (no difference between RFA and*
2 *no-RFA vertebroplasty); 5 Erdem et al. (2013b); Orgera et al. (2014)*

Table 65: GRADE profile: Excluding chemotherapy, which treatments are effective for spinal bone disease in patients with myeloma (surgery)?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	spinal surgery	control	Relative (95% CI)	Absolute	Quality
Vertebral collapse											
0	no evidence										
Spinal cord compression											
0	no evidence										
Health-related quality of life											
0	no evidence										
Progression-free survival											
0	no evidence										
Overall survival											
2[1]	observational studies	serious[2]	no serious inconsistency	serious[3]	serious[4]	none	159	n/a	-	Median OS 3.9y and 4.7y across studies	VERY LOW
Performance status											
0	no evidence										
Adverse events											
2[1]	observational studies	serious[2]	no serious inconsistency	serious[5]	serious[4]	none	39/129 (30.2 %)	n/a	-		VERY LOW
Pain control											
0	no evidence										
Activities of living/mobility											
0	no evidence										
Dependency											

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	spinal surgery	control	Relative (95% CI)	Absolute	Quality
0	no evidence										

1 Zeifang et al. (2005); Utzschneider et al. (2011); 2 Retrospective case series; 3 Survival not reported separately for spinal and non-spinal surgery. Cohort in Utzschneider (2011) dates back to 1980 which limits relevance to current UK practice; 4 Small sample size limits precision; 5 Complication not reported separately for spinal and non-spinal surgery patients in Utzschneider (2011)

Table 66: GRADE profile: Excluding chemotherapy, which treatments are effective for spinal bone disease in patients with myeloma (radiotherapy)?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	radiotherapy	control	Relative (95% CI)	Absolute	Quality
Vertebral collapse											
0	no evidence										
Spinal cord compression											
0	no evidence										
Health-related quality of life											
0	no evidence										
Progression-free survival											
0	no evidence										
Overall survival											
2[1]	observational studies	serious[2]	no serious inconsistency	serious[3]	no serious imprecision	none	319	n/a	-	Median OS 3.0 years and 2.7 years	VERY LOW
Performance status											
0	no evidence										

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	radiotherapy	control	Relative (95% CI)	Absolute	Quality
Adverse events (Grade 3-4)											
3[4]	observational studies	serious[2]	no serious inconsistency	serious[3]	no serious imprecision	none	3/371 (0.8%)	n/a	-	-	VERY LOW
Pain relief (proportion of patients with good/complete relief of pain)											
3[4]	observational studies	serious[2]	no serious inconsistency	serious[3]	no serious imprecision	none	284/521 (54.5%)	n/a	-	-	VERY LOW
Activities of daily living/mobility (proportion of patients reporting improvement in motor function)											
1[5]	observational studies	serious[2]	no serious inconsistency	serious[3]	serious[6]	none	62/79 (78%)	n/a	-	-	VERY LOW
Dependency											
0	no evidence										

1 Budak et al. (1991); Yaneva et al. (2006); 2 Non-comparative retrospective case series; 3 Outcomes not reported separately for spinal and non-spinal bone disease. Patients with spinal cord compression included in Budach et al. (1991); 4 Budach et al. (1991); Yaneva et al. (2006); Balducci et al. (2011); 5 Yaneva et al. (2006); 6 Small sample size limits precision

Table 67: GRADE profile: Excluding chemotherapy, which treatments are effective for spinal bone disease in patients with myeloma (denosumab versus zoledronic acid in patients with myeloma and at least one osteolytic lesion)?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	denosumab	zoledronic acid	Relative (95% CI)	Absolute	Quality
time to first on-study SRE (Better indicated by higher values)											
1[1]	randomised trials	no serious limitations	no serious inconsistency	serious[2]	Serious[3]	none	93	87	HR of 1.03 95% CI, 0.68 to 1.5	Not reported	LOW
overall survival (Better indicated by lower values)											

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	denosumab	zoledronic acid	Relative (95% CI)	Absolute	Quality
1[1]	randomised trials	no serious limitations	no serious inconsistency	serious[2]	serious3	none	93	87	HR of 2.26 (95% CI, 1.13 to 4.50	Not reported	LOW

1 *1 Henry et al. (2011); 2 Included patients had ≥1 osteolytic lesion – it is not specified if these lesions were vertebral or non-vertebral; 3 no absolute data reported for myeloma.*
2 *Small sample size and wide confidence intervals reduces precision.*

3

Cost effectiveness evidence (see also Appendix B)

There is uncertainty around whether balloon kyphoplasty (BKP) and vertebroplasty (VP) are cost effective when compared to non-surgical management (NSM). Upfront treatment costs will be higher with both BKP and VP but they could lead to dramatic improvements in quality of life and reduced resource use post-treatment.

The aim of the economic analysis was to assess the cost effectiveness of BKP and VP compared to NSM for the treatment of vertebral compression fractures (VCFs) patients with myeloma.

Economic evidence statement

A systematic literature review was performed to assess the current economic literature for this topic. The review identified 463 possibly relevant economic papers relating to myeloma. Of these, no papers were deemed relevant for this topic and therefore no papers were included in the review of existing economic evidence.

De novo economic analysis

A de novo economic analysis was conducted based upon outcomes and resource use reported in the one RCT identified for vertebral cement augmentation in the accompanying clinical evidence review (Berenson et al., 2011). The trial compared BKP to NSM for the treatment of VCFs in 134 patients with cancer.

The patient group had an average age of 64 years and was 58% male with an average estimated symptomatic fracture age of 3.5 months. The trial included cancers other than myeloma with 62% of the trial population having another cancer diagnosis.

The study had a large amount of crossover with patients randomised to NSM allowed to switch to BKP after one month follow-up. 38 (72%) of the 52 patients randomised to NSM group, who completed one month follow-up, crossed over to BKP. Therefore three groups were presented in the results by the authors: patients randomised to BKP, patients randomised to NSM who ultimately received BKP (crossover) and those who continued with NSM (NSM group). The authors reported no differences in the baseline characteristics of the three groups although differences at time of crossover were not reported.

Clinical input data

All clinical inputs for the model were based on evidence identified in the accompanying evidence review. Low quality evidence did not show any difference in clinical outcomes between VP and BKP. For the base case therefore the clinical outcomes were assumed to be identical between the two interventions. This assumption was explored during sensitivity analysis.

Patient groups

RCTs are conventionally analysed using an intention to treat (ITT) approach to reduce bias due to non-random loss and crossover of participants. As the Berenson et al trial had large crossover the ITT approach may not fully capture the true outcomes of the interventions being considered. The guideline committee therefore considered that an ‘as treated’ comparison comparing all patients who ultimately received BKP to those who remained in NSM would most accurately estimate the difference in effectiveness between the two groups. In the base case an ‘as treated’ approach was taken comparing the BKP and crossover group (cement technique received group) to the NSM group. A secondary analysis based on ITT principles was also conducted.

Utilisation of non-surgical interventions for VCFs at one month

The changes in the use of non-surgical interventions at one month are shown in Table 68.

Table 68: Percentage of patients in model cohort receiving non-surgical interventions at baseline and percentage point change in utilisation between base-line and one month follow-up.

	Percent baseline	BKP	NSM
Walking Aids	33%	-9.0%	1.5%
Bracing	14%	-12.7%	-1.4%
Wheelchair	6%	-4.8%	-2.0%
Physical Therapy	14%	-10.4%	-3.6%
Any Medication	86%	-40.5%	-17.0%
Radiation Therapy	4%	-0.9%	11.3%

Future VCFs

Whilst further VCFs are common in patients receiving both cement techniques and NSM, the accompanying systematic review found no evidence on whether there was a difference in the incidence of future VCFs between the interventions. In lieu of evidence it was assumed that the incidence between the groups was identical.

Adverse events

Device related adverse events were observed during the trial in the BKP group. Whilst the costs and quality of life detriments of these adverse events were not explicitly considered in the economic evaluation, costs attributable to adverse events of surgery were included (discussed later).

Survival

Survival for the economic analysis was taken from a prospective observational study of outcomes and survival in 39 patients with myeloma receiving VP in an NHS setting (Chew et al., 2011).

Time horizon

Time horizons of one and five years were used in the economic evaluation. The one year time horizon was the more conservative scenario as it was closer to the time period covered in the trial. Further assumptions were required for the five year time horizon as no evidence was identified around the effectiveness of BKP or VP post one year. Two alternative scenarios were investigated in the five year time horizon analysis based on the guideline committee's clinical experience. In the first scenario (used in base case) it was assumed that the quality of life difference between the groups at one year would be maintained over the entirety of the five year time horizon. In the second scenario it was assumed that in the group with the highest quality of life the difference would taper down at a constant rate until equal to the comparison group at five years.

The analysis also conservatively assumed that the difference in costs between the two groups would be identical after the first year. A sensitivity analysis was run for this model though that also assumed that the difference in costs not attributable to cement techniques, during the first year, would continue in all years.

Quality of life

The main measure of health related quality of life (HRQoL) in the trial was the Short Form (36) Health Survey (SF-36) physical component summary score (PCS). The change in SF-36 PCS from baseline for the BKP and NSM group, and from time of treatment for crossover

group, is shown in Table 69. These were given a normal distribution and varied across their reported range during probabilistic sensitivity analysis (PSA).

Table 69: Change in SF-36 PCS score following treatment

Follow-up	1 Month	3 Month	6 Month	12 Month
BKP	9.2	9.6	8.8	10.6
Crossover	8.8	10.8	10.4	10.6
NSM	-0.2	1.2	-0.8	1.2
Cement Technique Received	9.0	10.1	9.4	10.6
NSM-ITT	5.7	7.5	7.4	8.3

To conform to the NICE reference case, changes in the SF-36 PCS were converted to UK population preference EQ-5D weights using a mapping algorithm (Ara & Brazier, 2008). The estimated EQ-5D scores are shown in Table 70. Note that it was assumed that all patients started with a baseline QoL weight of 0.4392 based on the pre-treatment mean EQ-5D score from 11 consecutive patients receiving VP in an NHS setting (Chew, O'Dwyer, & Edwards, 2013).

Table 70: Estimated EQ-5D scores following treatment

Follow-up	Baseline	1 Month	3 Month	6 Month	12 Month
BKP	0.4392	0.4667	0.4679	0.4655	0.4709
Crossover	0.4392	0.4657	0.4717	0.4705	0.4709
NSM	0.4392	0.4386	0.4428	0.4368	0.4428
Cement Technique Received	0.4392	0.4662	0.4693	0.4674	0.4709
NSM-ITT	0.4392	0.4563	0.4617	0.4613	0.4643

Costs

Costs were inflated to 2014 prices, using the hospital & community health services (HCHS) index (Curtis, 2014). All costs are presented in Table 71.

Treatment costs

The costs of VP were taken from 11 consecutive patients receiving VP for spinal metastases at one NHS hospital. Chew et al estimated an average cost of £2213.25 per patient. This consisted of a cost of £744 for the VP kit and other costs of £1469. Treatment costs other than the kit cost were assumed to be identical for both VP and BKP.

The cost of the BKP kit was taken from NICE TA279 looking at BKP and VP in the treatment of osteoporotic vertebral compression fractures.

Non-surgical management costs

The annual cost of analgesic medication was taken from a study estimating the costs associated with VCFs from an NHS perspective using Hospital Episode Statistics and Personal Social Services Research Unit data (Puffer, et al., 2004).

Radiation therapy costs were taken from a cost-effectiveness analysis of zoledronic acid in the prevention of skeletal related events for patients with bone metastases secondary to advanced renal cell carcinoma (Botteman et al., 2011).

Bracing costs of £500 were estimated using correspondence with one NHS trust. Costs of wheelchair and walking aids were taken from PSSRU data (Curtis, 2014). Physical therapy

costs were estimated from NHS Reference Costs. Six appointments were assumed equal to a cost of £312 (Department of Health, 2015).

These costs were applied to both arms of the model in line with utilisation reported in Table 68.

The guideline committee felt that the clinical trial may not have adequately captured all relevant resource use. It was thought that there would be additional resource use associated with doctor and nurse time fitting, adjusting and advising on bracing and wheelchair use and time spent tailoring pharmaceutical treatment for pain. By virtue of greater utilisation in the NSM arm the underestimate would be larger than for the cement technique arms. Therefore, threshold sensitivity analysis was performed around the non-treatment costs to estimate the additional cost needed in the NSM arm to reduce the cost per QALY to the £20,000 NICE threshold. During PSA an non-specific cost was added to the NSM arm ranging from £0 to an upper estimate of £3552 equal to the total annual healthcare related cost of VCFs (Puffer et al., 2004).

Imaging costs

Costs of imaging pre-treatment were not included in this de novo economic evaluation as these were assumed to be performed as part of a patient's regular follow-up and would be identical between the two groups.

Table 71: Unit costs

Cost item	Value	Source	PSA Distribution
Total Cost BKP	£3369	(Chew et al., 2013)	Gamma(α=119.7, β=34.0)
Total Cost VP	£2213	(Chew et al., 2013)	Gamma(α=35.4, β=62.5)
Annual cost pharmaceutical treatment	£132	(Puffer et al., 2004)	Triangular(£66,£264)
Annual cost radiotherapy	£431	(Botteman et al., 2011)	Triangular(£216,£863)
Annual cost bracing	£500	NHS Correspondence	Uniform(£250,£1000)
Annual cost wheelchair	£91	(Curtis, 2014)	Triangular(£46,£182)
Annual cost walking aids	£91	(Curtis, 2014)	Triangular(£46,£182)
Annual cost physical therapy	£312	(Department of Health, 2015)	Gamma(α=25.1, β=12.4)
Annual Non-specific NSM costs	£0	Guideline committee estimate	Uniform(£0,£3552)

Discounting

All costs and QALYs were discounted at 3.5% per annum as recommended by the NICE Guidelines Manual (National Institute for Health and Care Excellence, 2014).

Results

Deterministic base case results

Table 72 and Table 73 show the base case results at one and five years for BKP and VP respectively. It can be seen that both cement procedures led to an increase in costs and QALYs. Total QALYs are equal between both cement techniques given the assumptions of the model with BKP having higher incremental costs owing to its increased kit cost. In all modelled scenarios the incremental cost effectiveness ratios (ICERs) are above the NICE threshold of £20,000 per QALY although it can be seen that the ICERs are substantially reduced in the five year scenario. It should be noted however that these estimates are likely

to be conservative because of the assumptions described in the previous sections above (particularly in regard to estimates of both incremental QALYs and NSM costs).

Table 72: Base case deterministic results for balloon kyphoplasty

Outcome	BKP	NSM	Incremental
One year time horizon			
Total Cost	£3,485	£304	£3,181
Total QALYs	0.4429	0.4170	0.0260
Cost per QALY gained			**£122,498**
Five year time horizon			
Total Cost	£3,485	£304	£3,181
Total QALYs	1.5678	1.4748	0.093
Cost per QALY gained			**£34,209**

Table 73: Base case deterministic results for vertebroplasty

Outcome	VP	NSM	Incremental
One year time horizon			
Total Cost	£2,329	£304	£2,025
Total QALYs	0.4429	0.4170	0.0260
Cost per QALY gained			**£77,987**
Five year time horizon			
Total Cost	£2,329	£304	£2,026
Total QALYs	1.5678	1.4748	0.093
Cost per QALY gained			**£21,779**

Probabilistic base case results

Table 74 and Table 75 show the base case probabilistic results calculated from the mean results of the PSA. The probabilistic results show an increased cost for NSM whilst the cement technique costs and QALYs for both groups remain consistent compared to the deterministic results. Other than for BKP in the conservative one year time horizon analysis all ICERs are now below the NICE £20,000 threshold. As NSM costs were almost certainly underestimated in the deterministic analysis these results are potentially more reflective of the true cost effectiveness.

Table 74: Base case probabilistic results for balloon kyphoplasty one year time horizon

Outcome	BKP	NSM	Incremental
One year time horizon			
Total Cost	£3,515	£2,191	£1,325
Total QALYs	0.4429	0.4170	0.0259
Cost per QALY gained			**£51,085**
Five year time horizon			
Total Cost	£3,519	£2,172	£1,347
Total QALYs	1.5680	1.4773	0.0908
Cost per QALY gained			**£14,842**

Table 75: Base case probabilistic results for vertebroplasty one year time horizon

Outcome	VP	NSM	Incremental
One year time horizon			

Outcome	VP	NSM	Incremental
Total Cost	£2,338	£2,168	£170
Total QALYs	0.4429	0.4169	0.0260
Cost per QALY gained			**£6,544**
Five year time horizon			
Total Cost	£2,354	£2,166	£188
Total QALYs	1.5681	1.4737	0.0944
Cost per QALY gained			**£1,994**

Deterministic sensitivity analysis

Deterministic sensitivity analysis was carried out to test alternate assumptions and how these influence the results of the economic evaluation. The results of the deterministic sensitivity analysis are shown in Table 76.

Table 76: Deterministic sensitivity analysis results-ICER for alternative assumptions

Modelled scenario	BKP-1 Year	BKP-5 year	VP-1 Year	VP-5 Year
Non-kit cost reduced to £996	£104,276	£29,120	£59,765	£16,690
Mental component added	£39,743	£11,726	£25,302	£7,471
Difference in costs continue post one year	N/A	£30,590	N/A	£18,171
Tapering quality of life after 1 year	N/A	£50,743	N/A	£32,309

Threshold analysis

A threshold analysis was performed to see how much extra NSM needed to cost, per patient, before the ICER reduced below £20,000 per QALY (Table 77). All the additional costs were lower than the upper limit of the PSA range.

Table 77: Additional NSM costs required for ICER to be below £20,000 per QALY

Strategy	1 Year Time Horizon	5 Year Time Horizon
BKP	£2662	£1322
VP	£1506	£166

Further threshold analysis showed that BKP needed to provide an additional 0.054 QALYs over the lifetime of a patient to give the same ICER when compared to VP. Given the assumptions of the model this was irrespective of the time horizon.

ITT Analysis

The ITT analysis did not alter the results, in terms of being above or below £20,000 per QALY, in any scenario.

Probabilistic sensitivity analysis results

The results of the probabilistic sensitivity analysis are shown in the cost effectiveness acceptability curves (CEACs) depicted below. Figure 7 and Figure 8 show the cost-effectiveness results for BKP against NSM at one and five years respectively. It can be seen

that BKP was below the cost-effectiveness threshold in 26.1% of iterations over a one year time horizon while under the five year time horizon this figure increased to 64.2%.

Figure 7: Cost effectiveness acceptability curve for balloon kyphoplasty with a one year time horizon

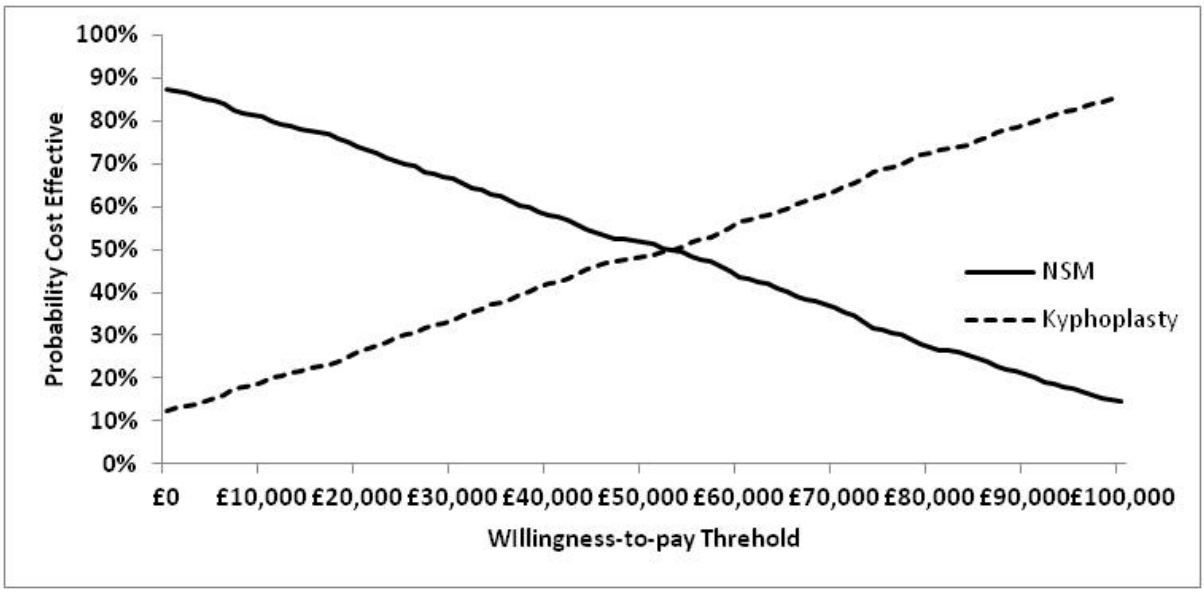


Figure 8: Cost effectiveness acceptability curve for balloon kyphoplasty with a five year time horizon

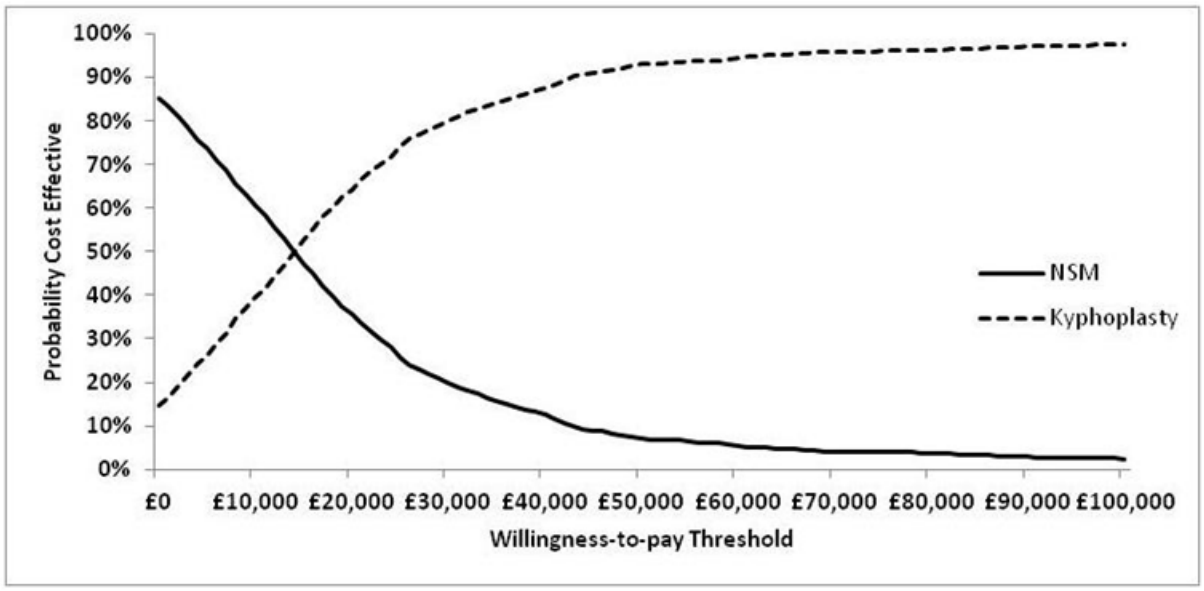


Figure 9 and Figure 10 show the cost-effectiveness results for VP against NSM at one and five years respectively. VP was shown to be below the cost-effectiveness threshold in 59.6% of iterations over a one year time horizon while under the five year time horizon this figure increased to 89.4%.

Figure 9: Cost effectiveness acceptability curve for vertebroplasty with a one year time horizon

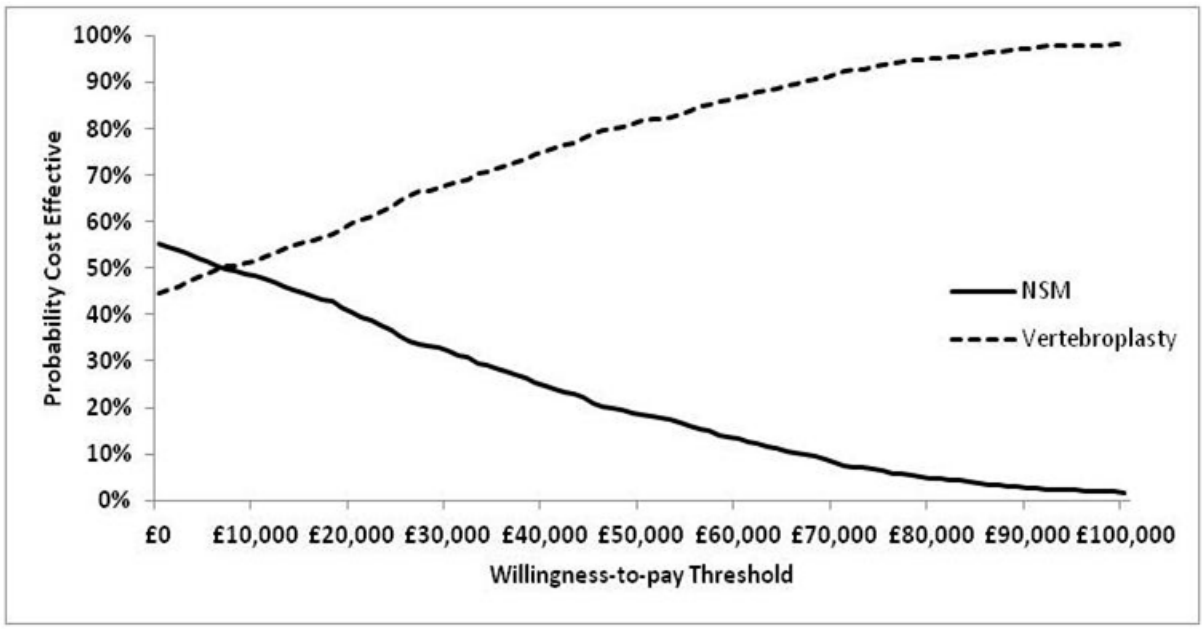
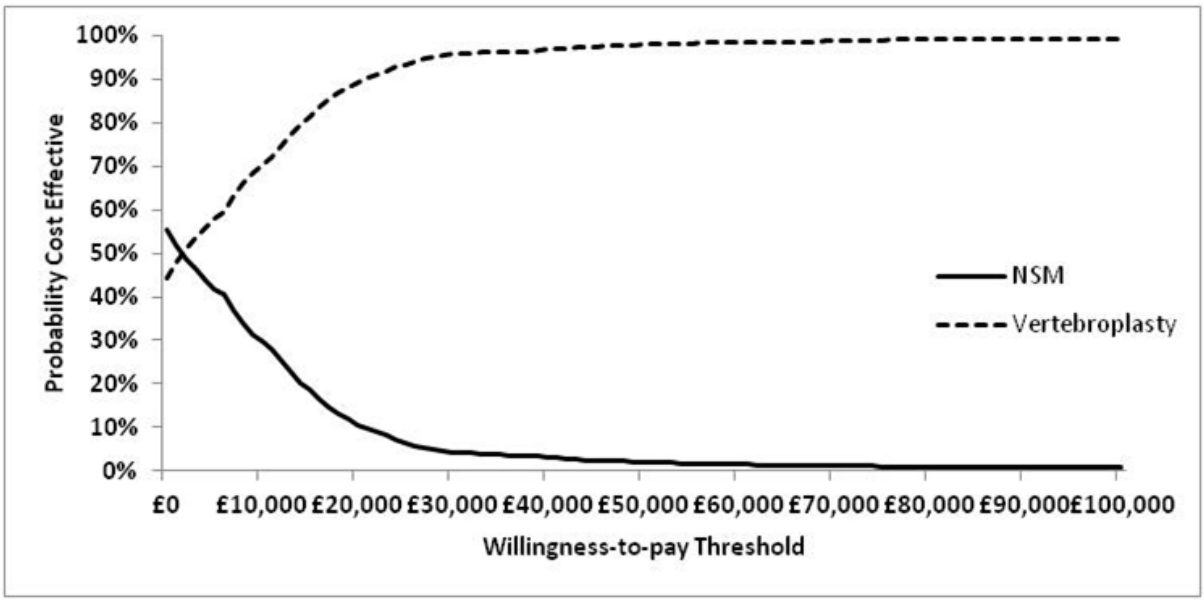


Figure 10: Cost effectiveness acceptability curve for vertebroplasty with a five year time horizon



Conclusions

The results of the base case analysis showed that BKP and VP were not cost effective over a one year time horizon and only VP was cost effective over a five year time horizon. However, when considering the probabilistic results, both cement techniques were shown to be cost effective with a five year time horizon with VP also cost effective under a one year time horizon. Furthermore, during PSA and under a five year time horizon both cement techniques were cost effective in the majority of iterations with VP being cost saving and health improving in 40% of cases.

The results were shown to be particularly sensitive to the costs of NSM. Threshold sensitivity analysis showed that even if our economic analysis only modestly underestimates the true

cost of NSM or the effectiveness of cement techniques then both VP and BKP would likely be cost effective.

Recommendations	**For guidance on treating metastatic spinal cord compression, see the NICE guideline on metastatic spinal cord compression.**
	
	Offer all people with myeloma and spinal bone disease:
	• **bisphosphonates as follows, if not already started:**
	  ○ **zoledronic acid, or**
	  ○ **disodium pamidronate, if zoledronic acid is contraindicated or not tolerated, or**
	  ○ **sodium clodronate, if zoledronic acid and disodium pamidronate are contraindicated, not tolerated or unsuitable**
	• **systemic pain control including when relevant using the NICE guidelines on neuropathic pain and opioids in palliative care.**
	
	Consider the following as adjuncts to other treatments for all people with myeloma and spinal bone disease:
	• **interventional pain management**
	• **bracing.**
	
	In people with radiological evidence of myeloma-related spinal instability, consider immediate intervention with:
	• **spinal surgery, with or without radiotherapy**
	• **cement augmentation, with or without radiotherapy**
	• **radiotherapy alone, if spinal intervention is unsuitable or not currently needed.**
	
	In people with radiological evidence of myeloma-related spinal bone disease without instability, consider:
	• **cement augmentation, with or without radiotherapy**
	• **radiotherapy alone.**

Relative value placed on the outcomes considered	The Guideline Committee considered the outcomes of vertebral collapse, spinal cord compression, health related quality of life, progression free survival, overall survival, performance status, adverse events, pain control, activities of daily living/mobility and dependency to be the most relevant to identify the most effective treatments for spinal bone disease in patients with myeloma. When drafting the recommendations the Guideline Committee considered prevention of vertebral collapse, spinal cord compression, health related quality of life, pain control and activities of daily living/mobility to be the most important outcomes as they are most important to the patient. Evidence was found for all outcomes except vertebral collapse, spinal cord compression and dependency.
Quality of the evidence	Evidence was identified for the interventions radiotherapy, surgery, bisphosphonates, denosumab and vertebral augmentation. No evidence was identified for pain control or bracing. The quality of the evidence was assessed using GRADE methodology and appropriate NICE checklists. Using these methods it was determined that the quality of the evidence ranged from very low to low across all outcomes.

	As the evidence was of low quality and the Guideline Committee had limited knowledge and experience of the management of spinal bone disease, three expert advisors were recruited to provide advice to the group. These were a clinical oncologist, an interventional radiologist and a spinal surgeon. The Guideline Committee noted that the Mhasker et al 2012 systematic review had included bisphosphonates (etidronate and ibandronate) that are known to be less effective in people with myeloma. They considered that the inclusion of these drugs had probably diluted the overall results and made zoledronic acid and disodium pamidronate appear to be less effective than they really are. Given this and the Guideline Committee's clinical experience that bisphosphonates are effective for the management of spinal bone disease the Guideline Committee made an 'offer' recommendation for this intervention for all patients.
Trade off between clinical benefits and harms	Based on the clinical evidence, expert advice and their clinical experience, the Guideline Committee agreed that recommendations were needed for 3 groups – all patients with spinal bone disease; those with spinal instability; those without spinal instability. Whilst the Guideline Committee noted that no clinical evidence had been found on systemic pain control, they agreed, based on their clinical experience, that provision of this intervention was fundamental to the care of patients with spinal bone disease. They therefore agreed to 'offer' systemic pain control to all patients with spinal bone disease. Based on expert advice and the clinical experience of the Guideline Committee, it was agreed to recommend interventional pain control and bracing for all patients with spinal bone disease. For patients with spinal instability, the Guideline Committee recommended the use of spinal surgery or cement augmentation, based on expert advice and their clinical experience that these interventions would improve the stability of the spine. The Guideline Committee were aware that in some people with spinal instability, surgical intervention would not be suitable due to the extent of spinal disease or the comorbidities of the patient. In these situations the Guidleine Committee recommended the use of radiotherapy alone to prevent further progression of spinal disease. The Guideline Committee also agreed that patients with spinal instability would need immediate intervention to prevent neurological damage. For patients without spinal instability the interventions recommended by the Guideline Committee were based on the clinical evidence, expert advice and their clinical experience. However, the Guideline Committee were unable to recommend an optimal sequence of interventions as there was not enough evidence to be able to determine this. The Guideline Committee made recommendations for the use of cement augmentation but did not specify kyphoplasty or vertebroblasty as expert advice suggested that it is not a case of one intervention being better than the other but that each is suitable in different patient circumstances. The Guideline Committee noted, based on the evidence, that denosumab had shown a survival advantage compared with

	bisphosphonates, and could therefore be a potential treatment option for people who were intolerant to bisphosphonates. However, they noted that the evidence was limited (coming from 1 trial of 180 patients). Given this and the fact that denosumab does not have a licence for use in patients with myeloma the Guideline Committee decided not to make any recommendations for this intervention. The Guideline Committee concluded that benefits would be improved and earlier pain control resulting in improved quality of life, a reduction in skeletal related events, an improvement in disease management and improved function and posture. The Guideline Committee acknowledged that there was a risk of radiation toxicity, complications arising from surgery/cement augmentation and a risk of infection as a result of the recommendations. Short term opioid toxicity after interventional procedures was also thought to be a possible harm. The Guideline Committee agreed that the risks of complications from the interventions were small but there would be vast improvements in health related quality of life, pain control and activities of daily living/mobility. Therefore the benefits outweighed the harms.
Trade-off between net health benefits and resource use	The Guideline Committee noted that no relevant economic evaluations were identified for this topic. As this topic was considered a high economic priority a de novo economic analysis was performed. The *de novo* model compared balloon kyphoplasty and vertebroplasty individually to non-surgical management in patients with vertebral compression fractures. It was not deemed appropriate to compare balloon kyphoplasty directly to vertebroplasty given that the suitability of either would depend on the characteristics of the vertebral fracture. In the base case, the model was based on the outcomes of Berenson et al trial, the only RCT identified in the clinical evidence review. Under these outcomes and under the trials one year time horizon neither vertebroplasty nor balloon kyphoplasty were deemed cost effective when compared to non-surgical management with ICERs of £77,987 per QALY and £122,498 per QALY respectively. The Guideline Committee thought basing the economic model solely on outcomes reported in the trial was likely to significantly under estimate two key components of cost effectiveness: the time horizon of the effects of cement augmentation and the resource use in the non-surgical management group. Alternate scenarios were therefore modelled. When the results are extrapolated out to five years following cement augmentation the ICERs decreased to £34,209 and £21,279 for balloon kyphoplasty and vertebroplasty respectively. Under this assumption non-surgical management costs only needed to be underestimated by £1400 for both to be less than £20,000 per QALY - an amount the Guideline Committee considered plausible and less than values reported in economic evaluations of other surgery versus non-surgical spinal interventions. Under the probabilistic results where a proxy from another economic evaluation was used for the non-surgical management costs the ICERs were below £20,000 for all interventions and time horizons other than for balloon kyphoplasty under a one year time horizon.

	The probabilistic sensitivity analysis suggested that under the five year time horizon there was a greater than 50% chance that both cement augmentation interventions were cost effective at a willingness to pay of £20,000 per QALY with 12.5% and 44.7% of iterations being both health improving and cost saving for balloon kyphoplasty and vertebroplasty respectively. The Guideline Committee considered there was moderate evidence that cement augmentation techniques were likely to be cost effective when compared to non-surgical management. Spinal surgery was not covered by the economic model given the different patient group. The Guideline Committee thought that there would be an overall cost saving from recommending spinal surgery as patients would be discharged from hospital earlier and would be more independent without need for a wheelchair.
Other considerations	The Guideline Committee discussed possible change in practice and agreed that as a result of the recommendations, cement augmentation would become more widely and uniformly available than it is currently.

References

Ara, R. & Brazier, J. (2008). Deriving an algorithm to convert the eight mean SF-36 dimension scores into a mean EQ-5D preference-based score from published studies (where patient level data are not available). Value in Health, 11, 1131-1143.

Balducci, M et al. (2011) Impact of radiotherapy on pain relief and recalcification in plasma cell neoplasms: long-term experience. Strahlentherapie und Onkologie; 187(2): 114-119.

Berenson, J., Pflugmacher, R., Jarzem, P., et al. (2011). Balloon kyphoplasty versus non-surgical fracture management for treatment of painful vertebral body compression fractures in patients with cancer: a multicentre, randomised controlled trial. The Lancet Oncology, 12, 225-235.

Botteman, M. F., Meijboom, M., Foley, I., et al (2011). Cost-effectiveness of zoledronic acid in the prevention of skeletal-related events in patients with bone metastases secondary to advanced renal cell carcinoma: application to France, Germany, and the United Kingdom. The European Journal of Health Economics, 12, 575-588.

Budach, V. (1991) Multiple myeloma: Results of radiotherapy in skeletal lesions. A review of 163 patients. Tumor Diagnostik und Therapie; 12(6): 238-243.

Catell, D., Kogen, Z., Donahue, B. et al. (1998) Multiple myeloma of an extremity: must the entire bone be treated? International Journal of Radiation Oncology, Biology, Physics, 40: 117-119.

Chang, S. A., Lee, S. S., Ueng, S. W., et al (2001) Surgical treatment for pathological long bone fracture in patients with multiple myeloma: a retrospective analysis of 22 cases. Chang Gung Medical Journal, 24: 300-306.

Chew, C., O'Dwyer, P. J., & Edwards, R. (2013). Health service cost associated with percutaneous vertebroplasty in patients with spinal metastases. Clinical radiology, 68, 776-779.

Chew, C., Ritchie, M., O'Dwyer, P. J et al (2011). A prospective study of percutaneous vertebroplasty in patients with myeloma and spinal metastases. Clinical radiology, 66, 1193-1196.

Curtis, L. (2014). The Unit Costs of Health and Social Care 2014.

Department of Health (2015). NHS Reference Costs 2013-14.

Delea TE, Rotter J, Taylor M, et al. (2012) 'Cost-effectiveness of zoledronic acid vs clodronic acid for newly-diagnosed multiple myeloma from the United Kingdom healthcare system perspective.' Journal of Medical Economics 15: p454-64.

Erdem, E et al. (2013a) Radiofrequency-targeted vertebral augmentation for the treatment of vertebral compression fractures as a result of multiple myeloma. Spine; 38(15): 1275-1281.

Erdem, E et al. (2013b) Vertebral augmentation in the treatment of pathologic compression fractures in 792 patients with multiple myeloma. Leukemia; 27(12): 2391-2393.

Henry, D. H., Costa, L., Goldwasser, F et al (2011) Randomized, double-blind study of denosumab versus zoledronic acid in the treatment of bone metastases in patients with advanced cancer (excluding breast and prostate cancer) or multiple myeloma. Journal of clinical oncology, 29: 1125-1132.

Khan, OA, Brinjikji, W, and Kallmes, DF. (2014) Vertebral augmentation in patients with multiple myeloma: a pooled analysis of published case series. American Journal of Neuroradiology; 35(1): 207-210.

Mhaskar, R., Redzepovic, J., Wheatley, K., et al (2012) Bisphosphonates in multiple myeloma: a network meta-analysis. [Review][Update of Cochrane Database Syst Rev. 2010;(3):CD003188; PMID: 20238320]. Cochrane Database of Systematic Reviews, 5: CD003188.

Morgan, G. J., Child, J. A., Gregory, W. M., et al. (2012) Effects of zoledronic acid versus clodronic acid on skeletal morbidity in patients with newly diagnosed multiple myeloma (MRC Myeloma IX): secondary outcomes from a randomised controlled trial. The.lancet oncology, 12: 743-752

Natarajan, M. V., Mohanlal, P. & Bose, J. C. (2007) The role of limb salvage surgery and custom mega prosthesis in multiple myeloma. Acta Orthopaedica Belgica, 73: 462-467.

National Institute for Health and Care Excellence (2014). The guidelines manual. London: NICE: 2014. National Institute for Clinical Excellence, 13, 4-2007.

Orgera, G et al. (2014) Percutaneous vertebroplasty for pain management in patients with multiple myeloma: Is radiofrequency ablation necessary? Cardiovascular and Interventional Radiology; 37(1): 203-210.

Papagelopoulos, P. J., Galanis, E. C., Greipp, P. R.et al. (1997) Prosthetic hip replacement for pathologic or impending pathologic fractures in myeloma. Clinical Orthopaedics and Related Research, 192-205.

Papanastassiou, ID et al. (2014) Comparison of Unilateral versus Bilateral Kyphoplasty in Multiple Myeloma Patients and the Importance of Preoperative Planning. Asian Spine Journal; 8(3): 244-252.

Puffer, S., Torgerson, D. J., Sykes, D., et al. (2004). Health care costs of women with symptomatic vertebral fractures. Bone, 35, 383-386.

Simony, A. (2014). Pain reduction after percutaneous vertebroplasty for myeloma-associated vertebral fractures. Danish Medical Journal, 61, A4945.

Utzschneider S., S. (2011) Surgical therapy of skeletal complications in multiple myeloma. International Orthopaedics; 35(8): 1209-1213.

1 Yaneva, MP, Goranova-Marinova, V, & Goranov, S. (2006) Palliative radiotherapy in patients
2 with multiple myeloma. Journal of Balkan Union of Oncology; 11(1): 43-48.

3 Zadnik PL, Goodwin CR, Karami KJ, et al. (2015). Outcomes following surgical intervention
4 for impending and gross instability caused by multiple myeloma in the spinal column. Journal
5 of Neurosurgery Spine, 22, 301-309.

6 Zeifang, F et al. (2005) Long-term survival after surgical intervention for bone disease in
7 multiple myeloma. Annals of Oncology; 16(2): 222-227.

9 Preventing and managing complications

9.1 Preventing infection

Plasma cells are antibody producing cells and are a major component of the immune system. Patients with myeloma have an increased risk of developing infections due to suppression of the immune system caused by the disease and its treatment. It is also known that specific treatments can be associated with specific types of infections. Herpes zoster infections following proteasome inhibitor therapy are common and aciclovir or similar prophylaxis is frequently prescribed to avoid this complication. Often infections can be more difficult to treat in people with myeloma and are one of the commonest causes of death in the first 3 months after diagnosis and during times of active disease.

Possible prophylactic measures include antibiotics, antiviral drugs, antifungal drugs, the use of pre-emptive vaccination (e.g. for flu), the use of growth factors which stimulate aspects of the immune system and regular immunoglobulin replacement therapy. Whilst there may be benefits in terms of reducing the number and severity of infections, there is also a possible risk resulting from drug-related side effects and the development of drug resistance due to overuse. The use of these measures therefore requires clarification as well as the different time points at which they should be used.

Clinical question: What is the most effective prophylactic strategy for infection in patients with myeloma (including immunoglobulin, antibiotics, growth factors and vaccinations)?

Clinical evidence (see also Appendix G)

See Tables 78-85.

Study quality

Four systematic reviews, 5 randomised trials and 2 non randomised comparative studies (1 prospective and 1 retrospective) which met the inclusion criteria were identified.

Due to the nature of the topic, inclusion of studies was not limited to those with exclusively a myeloma population and as such some of the studies included patients with other haematological malignancies, such as lymphoma or leukaemia.

Studies in which neutropenia was the primary outcome of interest were excluded as the prophylactic treatment of neutropenia is covered by current NICE guidance on neutropenic sepsis

Much of the available evidence concentrated on prophylaxis in patients undergoing stem cell transplants with little evidence available relating to patients on active maintenance, relapsed myeloma or myeloma patients off treatment. No studies investigating the effect of prophylactic treatment on hepatitis in patients with myeloma were identified.

Newly diagnosed myeloma patients

Low quality evidence from one randomised trial including 212 patients with newly diagnosed myeloma (Vesole et al, 2012) suggests uncertainty about the effectiveness of prophylactic antibiotics (quinolone/ofloxacin or trimethoprim-sulfamethoxazole) compared to observation alone. The rate of severe bacterial infection was 9.3% with antibiotics versus 15.9% with observation (RR=0.59; 95% C.I. 0.28 to 1.28)

Patients on active therapy or maintenance therapy

Growth Factors

Moderate evidence from one randomised trial including 281 patients undergoing chemotherapy in a high dose Melphalan (HDM) transplant setting (Blijlevens et al, 2013) suggests uncertainty about the effectiveness of prophylactic palifermin compared to placebo for the prevention of oral mucositis. The rate of severe oral mucositis was 38% with palifermin versus 37% with placebo (RR 1.04; 95% C.I. 0.69 to 1.57).

Immunoglobulins

Low quality evidence came from a single randomised trial including 81 patients with myeloma comparing polyvalent intravenous immunoglobulins (IVIG) with placebo, identified in the Raanani et al (2009) systematic review. Low quality evidence suggests uncertainty about the effect of polyvalent IVIG versus placebo in terms on all cause mortality during study follow-up (19% versus 7% respectively; RR 2.67; 95% CI 0.76 to 9.35). Low quality evidence suggests that polyvalent IVIG is effective compared to placebo in preventing major infections (5% versus 24% respectively; RR 0.20; 95% CI 0.05 to 0.86) and clinically documented infections (42% versus 93% respectively; RR 0.45; 95% CI 0.31 to 0.65).

Antibiotics

Low quality evidence came from one randomised trial including 54 patients (Oken et al, 1996) comparing 2 months of trimethoprim-sulfamethoxazole (TMP-SMZ) prophylaxis with no prophylaxis in patients with myeloma. Low quality evidence suggests that TMP-SMZ prophylaxis is effective compared to no prophylaxis in reducing the rate of infection (18% versus 46% respectively; RR 0.39; 95% CI 0.16 to 0.95).

Post autologous transplant myeloma patients

Growth factors

Low quality evidence from one randomised trial including 47 patients (31 with myeloma; Ozkan et al, 2013) suggests uncertainty about whether G-CSF daily versus every other day is the more effective in terms of time to neutrophil engraftment (median was 10 days in both groups; P=0.31); Very low quality evidence from one retrospective study including 117 patients (Cox et al, 2014) reported significantly longer time to neutrophil engraftment in patients receiving delayed G-CSF administration compared with conventional administration (15 days versus 12 days respectively; P<0.0001).

Low quality evidence from one randomised trial including 47 patients (Ozkan et al, 2013) suggests uncertainty about the relative effectiveness of daily G-CSF daily versus every other day for the prevention of blood stream infection (rates were 14% versus 19% respectively; RR 0.74; 95% CI 0.20 to 2.76).

Immunoglobulins

Moderate quality evidence from one systematic review and meta-analysis including a total of 4223 patients (Raanani et al, 2009) reported no significant difference in all cause mortality for patients treated with polyvalent IVIG versus no treatment (1418 patients in 8 trials; 0.99 (0.88 to 1.12) p=0.92). Infection related death did not differ significantly between the groups (275 patients in 3 trials; Risk Ratio 0.64 (0.28 to 1.49) P=0.3).

Moderate quality evidence from one systematic review and meta-analysis including a total of 4223 patients (Raanani et al, 2009) reported significantly more adverse events for patients treated with polyvalent IVIG compared with placebo/no treatment (728 patients in 5 trials; Risk Ratio 8.12 (3.15 to 20.97) P=0.000015).

Anti-fungals

Very low quality evidence from a retrospective study of 104 patients (Orvain et al., 2015) suggests uncertainty about the effectiveness miconazole mucoadhesive buccal tablets compared with oral amphotericin B suspension in reducing hospital stay after stem cell re-infusion (mean 15.3 days versus 16.4 days respectively; p=0.09).

Viral Vaccinations

Varicella zoster vaccine (VZV)

Low quality evidence from two randomised trials including 139 patients with haematological malignancies (Cheuk et al, 2011) suggests uncertainty about the benefit of VZV compared to no vaccine on all cause mortality (Risk Ratio 0.96; 95% CI0.54 to 1.69:P=0.89). Low quality evidence suggests that both systemtic and local adverse events (at the injection site) are more likely with VZV than with no vaccination. Systemic adverse events occurred at a rate of 5% with VZV and local adverse events at a rate of 21%, no adverse events were reported in the no vaccination group.

Influenza Vaccine

Low quality evidence from 2 trials (Cheuk et al, 2011) comparing influenza vaccine to no vaccine in patients with haematological malignancies suggests uncertainty about its effectiveness in preventing infection related mortality (Risk Ratio 0.2 [0.01-3.97] p=0.29). In this analysisLower respiratory tract infections were more likely in the no vaccine group (Risk ratio 0.39; 95% CI [0.19-0.78] p=0.0082). Rates of hospitalisation (Risk ratio 0.17 [0.09-0.31] p<0.00001) were significantly higher in the no vaccine group while the frequency of adverse events (Risk Ratio 35 [4.9-249.8] p=0.00039) were significantly higher in the vaccine group.

Relapsed myeloma patients and myeloma patients currently off treatment

No evidence relating to prophylactic infection strategies for relapsed myeloma patients or those currently off treatment was identified.

Table 78: GRADE profile: What is the most effective prophylactic strategy for infection in patients with myeloma (antibiotics compared to observation for patients with newly diagnosed myeloma)?

Quality assessment							No of patients		Effect		
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Antibiotics**	**Observation**	**Relative (95% CI)**	**Absolute**	**Quality**
Severe Bacterial Infection at 2 months (follow-up 2 months)											
1[3]	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	13/138 (9.4%)	10/63 (15.9%)	RR 0.59 (0.28 to 1.28)	65 fewer per 1000 (from 114 fewer to 44 more)	LOW
Any infection during the first 2 months											
1[3]	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	30/138 (21.7%)	14/63 (22.2%)	RR 0.98 (0.56 to 1.71)	4 fewer per 1000 (from 98 fewer to 158 more)	LOW
Severe infection during the 1st month											
1[3]	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	serious[2]	none	4/138 (2.9%)	3/63 (4.8%)	RR 0.61 (0.14 to 2.64)	19 fewer per 1000 (from 41 fewer to 78 more)	LOW

1 No details provided on randomisation method or blinding; 2 Small sample size; 3 Vesole et al, 2012

Table 79: GRADE profile: What is the most effective prophylactic strategy for infection in patients with myeloma (palifermin compared to placebo for patients undergoing conditioning chemotherapy)?

Quality assessment							No of patients		Effect		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Palifermin	Placebo	Relative (95% CI)	Absolute	Quality
Incidence of ulcerative oral mucositis (follow-up 14 days)											
1[2]	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	79/115 (68.7%)	33/57 (57.9%)	RR 1.19 (0.92 to 1.53)	110 more per 1000 (from 46 fewer to 307 more)	MODERATE
Incidence of severe oral mucositis (follow-up 14 days)											
1[2]	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	44/115 (38.3%)	21/57 (36.8%)	RR 1.04 (0.69 to 1.57)	15 more per 1000 (from 114 fewer to 210 more)	MODERATE
Serious adverse events											
1[2]	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[1]	none	18/109 (16.5%)	3/57 (5.3%)	RR 3.14 (0.96 to 10.21)	113 more per 1000 (from 2 fewer to 485 more)	MODERATE

1 Small sample size, 2 Blijlevens et al, 2013

Table 80: GRADE profile: What is the most effective prophylactic strategy for infection in patients with myeloma (immunoglobulins compared to placebo/no treatment for patients with lymphoproliferative disorders)?

Quality assessment							No of patients		Effect		Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Immuno-globulins	Placebo/No treatment	Relative (95% CI)	Absolute	
All cause mortality (follow-up 1 years[1])											
1[2]	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	none	8/41 (19.5%)	3/41 (7.3%)	RR 2.67 (0.76 to 9.35)	122 more per 1000 (from 18 fewer to 611 more)	LOW
Major Infections											
1[2]	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	none	2/41 (4.9%)	10/41 (24.4%)	RR 0.20 (0.05 to 0.86)	195 fewer per 1000 (from 34 fewer to 232 fewer)	LOW
Clinically documented infection											
1[2]	randomised trials	no serious risk of bias	no serious inconsistency	no serious indirectness	very serious[3]	none	17/41 (41.5%)	38/41 (92.7%)	RR 0.45 (0.31 to 0.65)	510 fewer per 1000 (from 324 fewer to 640	LOW

Quality assessment							No of patients		Effect		Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Immuno-globulins	Placebo/No treatment	Relative (95% CI)	Absolute	
										fewer)	

1 *1 All cause mortality was assessed at 1 year in the two trials for which this outcome was reported; 2 Raanani (2009) systematic review - single MM trial Chapel (1994); 3*
2 *Small sample size*

3 **Table 81: GRADE profile: What is the most effective prophylactic strategy for infection in patients with myeloma (trimethoprim-**
4 **sulfamethoxazole versus no treatment for patients with a confirmed melanoma diagnosis (Oken et al, 1996))?**

Quality assessment							No of patients		Effect		Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Trimethoprim-sulfamethoxazole	No treatment	Relative (95% CI)	Absolute	
Infection Incidence											
1[1]	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[3]	none	5/28 (17.9%)	12/26 (46.2%)	RR 0.39 (0.16 to 0.95)	282 fewer per 1000 (from 23 fewer to 388 fewer)	LOW
Death from infection											
1[1]	randomised trials	serious[2]	no serious inconsistency	no serious indirectness	serious[3]	none	1/28 (3.6%)	4/26 (15.4%)	RR 0.23 (0.03 to 1.94)	118 fewer per 1000 (from 149 fewer to 145	LOW

Quality assessment							No of patients		Effect		Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Trimethoprim-sulfamethoxazole	No treatment	Relative (95% CI)	Absolute	
										more)	

1 *1 Oken et al (1996); 2 No details on randomisation method or blinding; 3 Small sample size*

2 **Table 82: GRADE profile: What is the most effective prophylactic strategy for infection in patients with myeloma (immunoglobulins**
3 **versus placebo or no treatment/different preparation, schedule or dose in patients undergoing hematopoietic stem cell**
4 **transplantation)?**

Quality assessment							No of patients		Effect		Quality
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Immuno-globulins	Placebo/no treatment/different preparation, schedule or dose	Relative 95% CI)	Absolute	
All cause mortality											
8	randomised trials[1]	no serious risk of bias	no serious inconsistency	serious[2]	no serious imprecision	none	300/756 (39.7%)	273/662 (41.2%)	RR 0.99 (0.88 to 1.12)[3]	4 fewer per 1000 (from 49 fewer to 49 more)	MODERATE
Infection related death											
3	randomised trials[1]	no serious risk of bias	no serious inconsistency	no serious indirectness	serious[2]	none	8/137 (5.8%)	12/138 (8.7%)	RR 0.64 (0.28 to 1.49)[4]	31 fewer per 1000 (from 63 fewer to 43 more)	MODERATE

Quality assessment							No of patients		Effect		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Immuno-globulins	Placebo/no treatment/different preparation, schedule or dose	Relative (95% CI)	Absolute	Quality
Clinically documented infections											
5	randomised trials[1]	no serious risk of bias	no serious inconsistency	serious[2]	no serious imprecision	none	267/388 (68.8%)	181/300 (60.3%)	RR 1.00 (0.9 to 1.1)[5]	0 fewer per 1000 (from 60 fewer to 60 more)	MODERATE
Adverse Events											
5	randomised trials[1]	no serious risk of bias	no serious inconsistency	serious[2]	no serious imprecision	none	49/415 (11.8%)	2/313 (0.64%)	RR 8.12 (3.15 to 20.97)[6]	45 more per 1000 (from 14 more to 128 more)	MODERATE

1 *1 Raanani et al (2009); 2 Not all included patients were Myeloma patients*

2 **Table 83: GRADE Profile: What is the most effective prophylactic strategy for infection in patients with myeloma (G-CSF**
3 **(conventional dosing) versus delayed or reduced dose for patients undergoing autologous stem cell transplant)?**

Quality assessment							No of patients		Effect		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	G-CSF (conventional dosing)	Delayed or reduced dose	Relative (95% CI)	Absolute	Quality
Neutrophil engraftment (randomised trials) (Better indicated by lower values)											

Quality assessment							No of patients		Effect		
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**G-CSF (conventional dosing)**	**Delayed or reduced dose**	**Relative (95% CI)**	**Absolute**	**Quality**
1	randomised trials[1]	no serious risk of bias	no serious inconsistency	serious[2]	serious[3]	none	21	26	-	Median 18 days in both groups	LOW
Neutrophil engraftment (observational studies)											
1	observational studies[4]	serious[5]	no serious inconsistency	no serious indirectness	no serious imprecision	none	52	65	-	Mean 12 days with conventional versus 15 days with delayed dose	VERY LOW
Blood stream infections											
1[1]	randomised trials	no serious risk of bias[2]	no serious inconsistency	serious[2]	serious[3]	none	3/21 (14.3%)	5/26 (19.2%)	RR 0.74 (0.20 to 2.76)	50 fewer per 1000 (from 154 fewer to 338 more)	LOW
Hospitalisation (Better indicated by lower values)											
1	randomised trials[1]	no serious risk of bias	no serious inconsistency	serious[2]	serious[3]	none	21	26	-	MD 1.1 days shorter with conventional dose	LOW

1 *1 Ozkan (2013); 2 Mixed haematological malignancies including myeloma; 3 Small sample size; 4 Cox (2014); 5 Unbalanced baseline characteristics between groups*

Table 84: GRADE profile: What is the most effective prophylactic strategy for infection in patients with myeloma (miconazole mucoadhesive buccal tablets versus oral amphotericin-B suspension in patients receiving high dose melphalan and autologous stem cell transplant for haematological malignancy)?

Quality assessment							No of patients		Effect		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Miconazole mucoadhesive buccal tablets	Oral amphotericin-B suspension	Relative (95% CI	Absolute	Quality
Duration of hospital stay (Better indicated by lower values)											
1	observational studies[1]	serious[2]	no serious inconsistency	serious[3]	serious[4]	none	60	44	-	MD 1.1 lower with MBT	VERY LOW

1 Orvain (2015); 2 Not a randomised trial (prospective cohort compared with a historical cohort); 3 All haematological malignancies; 51/104 patients with myeloma; Small sample size

Table 85: GRADE profile: What is the most effective prophylactic strategy for infection in patients with myeloma (viral vaccines versus placebo, no vaccines, alternative dosing regimens or schedules in patients with haematological malignancies)?

Quality assessment							No of patients		Effect		
No of studies	Design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Viral vaccines	Placebo, no vaccines, alternative dosing regimens or schedules	Relative (95% CI)	Absolute	Quality
All cause mortality (Varicella zoster vaccine)											
2	randomised trials[1]	no serious risk of bias	no serious inconsistency	serious[2]	serious[3]	none	17/67 (25.4%)	19/72 (26.4%)	RR 0.96 (0.54 to 1.69)	11 fewer per 1000 (from 121 fewer to 182 more)	LOW

Quality assessment							No of patients		Effect		
No of studies	**Design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Viral vaccines**	**Placebo, no vaccines, alternative dosing regimens or schedules**	**Relative (95% CI)**	**Absolute**	**Quality**
Local adverse events (Varicella zoster vaccine)											
2	randomised trials[1]	no serious risk of bias	no serious inconsistency	serious[2]	serious[3]	none	20/97 (20.6%)	0/97 (0%)	RR 20.94 (2.88 to 152.36)	-	LOW
Systemic adverse events (Varicella zoster vaccine)											
2	randomised trials[1]	no serious risk of bias	no serious inconsistency	serious[2]	serious[3]	none	5/97 (5.2%)	0/97 (0%)	RR 5.94 (0.73 to 48.55)	-	LOW

1 *1 Cheuk (2011); 2 All haematological malignancies; 3 Low sample size*

2

1 Cost effectiveness evidence

2 A literature review of published cost-effectiveness analyses did not identify any relevant
3 papers for this topic. Whilst there were potential cost implications of making
4 recommendations in this area, other questions in the guideline were agreed as higher
5 priorities for economic evaluation. Consequently no further economic modelling was
6 undertaken for this question.

7

Recommendations	**Offer people with myeloma the seasonal influenza vaccination.** **Consider extending the pneumococcal vaccination to people with myeloma who are under 65.** **Consider intravenous immunoglobulin replacement therapy for people who have hypogammaglobulinaemia and/or recurrent infections.** **Consider continuing aciclovir[b] or equivalent antiviral prophylaxis after treatment with bortezomib or other proteasome inhibitors ends.** **Consider aciclovir[b] or equivalent antiviral prophylaxis for people who are taking both immunomodulatory drugs and high-dose steroids.** **Consider testing for hepatitis B, hepatitis C and HIV before starting myeloma treatment.**
Relative value placed on the outcomes considered	The Guideline Committee considered the outcomes sepsis, recorded infections, death related to infection, hospital admissions, adverse events, response to vaccination, and patient adherence and acceptability to be the most relevant in determining the most effective prophylactic strategy for infection in patients with myeloma. Of these, evidence was identified for all outcomes except patient adherence and acceptability. Sepsis was reported in some studies however not included in the evidence review as the Guideline Committee considered this to be covered by the NICE guidance on neutropenic sepsis. When drafting the recommendations the Guideline Committee considered recorded infections, hospital admissions, response to vaccination and patient acceptability to be the most important.
Quality of the evidence	The quality of the evidence was assessed using GRADE methodology and appropriate NICE checklists and ranged from moderate to very low quality. There was moderate quality evidence for overall survival, infection-related mortality, recorded infection and response to vaccination. Evidence for hospital admissions ranged from very low to moderate quality and for adverse events ranged from moderate to high quality.
Trade off between clinical benefits and harms	The Guideline Committee agreed to recommend using the seasonal influenza vaccination based on the clinical evidence which

b At the time of consultation (August 2015), aciclovir did not have a UK marketing authorisation for this indication. The prescriber should follow relevant professional guidance, taking full responsibility for the decision. Informed consent should be obtained and documented. See the General Medical Council's Prescribing guidance: prescribing unlicensed medicines for further information.

	demonstrated a decrease in respiratory tract infections and a reduction in hospital admissions. Timing of any vaccination is important as it is likely to be ineffective during periods of active disease and treatment. The Guideline Committee recommended using the pneumococcal vaccination based on their knowledge of evidence from other immunological deficiencies that it is of benefit in preventing infection in the post transplant setting. The Guideline Committee were aware that use of this vaccination in people over 65 is already recommended in Department of Health policy. However, the Guideline Committee noted that patients with myeloma, who could also benefit from this vaccine, would not all fall within this age group. They therefore recommended its use be extended to patients under 65. The Guideline Committee made a recommendation for regular intravenous immunoglobulin replacement therapy in myeloma patients with hypogammaglobulinaemia and/or recurrent infections based on the clinical evidence. It was recognised that the effectiveness of this strategy is likely to change with different phases of the disease but the optimal timing for immunoglobulin replacement remains unclear. The Guideline Committee noted that no evidence had been identified on the use of aciclovir for prophylaxis of Herpes zoster infection and that it is not licensed for this indication. However, the GDG were aware, based on their clinical experience of the evidence in other areas, that bortezomib is particularly likely to induce zoster re-activations in excess of 25% of patients receiving this drug. All clinical trials involving bortezomib therefore mandate the use of aciclovir or equivalent antiviral prophylaxis in the treatment protocol. Data from clinical trials involving the use of bortezomib with aciclovir prophylaxis have reported zoster re-activation rates of less than 5%. The GDG therefore agreed to recommend aciclovir or equivalent antiviral prophylaxis based on their clinical experience that it can reduce zoster re-activation whilst on therapy. The Guideline Committee recommended prolonged intervention beyond completion of treatment with bortezomib or other proteasome inhibitors because reactivations can occur for some months after stopping treatment. Based on their clinical experience, the Guideline Committee made a recommendation for pre-treatment screening for viruses to reduce uncontrolled viral re-activation, with its associated morbidity and mortality. The Guideline Committee chose to recommend screening for those viruses where management can be altered if they are known to exist before treatment starts. The Guideline Committee concluded that the potential benefits would be a reduction in infections and a subsequent reduction in hospital admissions and requirement for pain relief, which would result in an improvement in quality of life. Potential harms were the risk of adverse reactions to the immunoglobulin therapy and vaccinations and the potential for drug related side-effects. The Guideline Committee considered that the benefits outweighed the harms in this instance.
Trade-off between net health benefits and	The Guideline Committee noted that no relevant published economic evaluations had been identified and no additional economic analysis

resource use	had been undertaken in this area. The Guideline Committee considered that the recommendations made would result in an increase in costs for providing immunoglobulin therapy. However it was suggested that this cost would be offset by the resulting decrease in hospital admissions due to infection. The Guideline Committee agreed that there would be an increase in costs associated with recommending pre-treatment screening. In addition, if screening identified a virus there would be additional costs involved to treat this. However the Guideline Committee thought these additional costs were acceptable given the potential improvements in patient quality of life. For the other prophylactic strategies recommended, the Guideline Committee concluded that there would be minimal cost change as these treatments are already happening.
Other considerations	The Guideline Committee felt that there would be a change in practice as a result of the recommendations as viral screening is not currently standard practise. They noted that the facilities already exist to do this but there would now be an increase in the volume of work. The Guideline Committee considered that the recommendations would result in myeloma practise being brought in line with current practice in lymphoma.

9.2.1 Managing peripheral neuropathy

Neuropathy is the condition when nerves (including the spinal cord) are damaged or diseased. This can occur in myeloma due to nerve compression, amyloidosis, paraprotein related demyelination, herpes zoster virus and side effects of drug treatment (particularly from thalidomide and bortezomib). It remains important to avoid development of neuropathy whenever possible, for example by reducing or even stopping neuropathic drugs.

Neuropathy causes several unpleasant symptoms which can impair the patient's quality of life. The main symptoms are numbness, pins and needles (paraesthesia), pain, and in severe cases, it may cause muscle weakness and adversely affect proprioception. The feet, lower legs and hands are most commonly affected by drug-related neuropathy. Shingles may affect any part of the body, including the face.

Neuropathy and in particular the related painful symptoms can be managed pharmacologically. All of the drugs used carry potentially difficult or even dangerous side-effects. Recommendations on drug management of neuropathic pain are covered by NICE guidance on the pharmacological management of neuropathic pain in adults in non-specialist settings, and are not covered here.

There are a variety of non-pharmacological strategies used to manage neuropathy. These include lowering the dose of the drug thought to be responsible, or stopping it for a period of time, complementary therapies such as reflexology and acupuncture, TENS (trans-cutaneous nerve stimulation), and vitamin supplements such as vitamin B complex, folic acid, magnesium and alphalipoic acid. Stopping treatment can be very difficult to accept if it is effective at treating myeloma as it may mean sub-optimal disease management and associated affect on survival.

Clinical question: What is the most effective way to manage neuropathy in patients with myeloma (excluding pharmacological management of neuropathic pain)?

Clinical evidence (see also Appendix G)

See table 86

Study quality

The evidence base consisted of one non-randomised, comparative study (Cho et al, 2014) and five single arm, non-comparative studies all of very low quality (Bao et al, 2014; Garcia et al, 2014; Mack et al, 2010; Richardson et al, 2009; Truni et al, 2011) as assessed by GRADE and NICE checklists. Evidence was not available for all interventions or outcomes of interest, with no evidence found to report on use of nutritional supplements, active monitoring or TENS. None of the included studies reported overall survival as an outcome, primarily because follow-up in the studies was restricted to only a short period of time following treatment. In reporting and assessing the effect of interventions on neuropathy, all studies relied on self reporting of outcomes by included patients through the use of standard questionnaires, leaving them at high risk of bias.

All included studies had very small sample sizes, while one study included participants other than those with myeloma. Given these considerations therefore, the evidence presented should be considered with caution.

Myeloma treatment modifications

In one cohort study (Richardson et al, 2009), 72/91 patients had chemotherapy dose modification per guidelines and 49/72 (68%) experienced improvement or resolution of peripheral neuropathy in a median of 110 days (range: 4-376) [Very low quality evidence].

41 patients had dose modifications but did not discontinue bortezomib; 71% (n=29) had resolution of peripheral neuropathy in a median of 78 days (range 9-376) and in the patients who discontinued treatment, 65% (n=20) experienced improvement (n=8) or resolution (n=12) in a median of 122 days (range 4-296) [Very low quality evidence].

From one cohort study (Richardson et al, 2009), the occurrence of peripheral neuropathy did not adversely affect response rate, median time to progression or median overall survival and no effect of dose reductions or modification was observed for response rate, median time to progression or median overall survival [Very low quality evidence].

From one study which evaluated the impact of dose-modification on treatment compliance (Cho et al, 2014) patients who received dose modifications according to guidelines were more likely to complete bortezomib treatment (OR=1.4, 95% CI, 0.31-6.32, p=0.66) though the difference was not statistically significant [Very low quality evidence].

Acupuncture/Electroacupuncture

From two studies (Boa et al, 2014; Garcia et al, 2014) no significant adverse events (no excessive bruising, local persistent pain or evidence of excessive bleeding at point of needle placement) associated with acupuncture treatment were reported in a total of 46 patients [Very low quality evidence].

From two studies (Boa et al, 2014; Garcia et al, 2014), mean scores, as assessed using FACT/GOG-NTx were significantly improved from baseline indicating a benefit of acupuncture [Very low quality evidence]

Nutritional supplements

One prospective case series study (n=30) evaluated the therapeutic potential of palmitoylethanolamide (PEA) on pain and nerve function (Truni et al, 2011) and reported a

reduction in mean pain scores following 2 months of treatment (4.5±2.4 versus 3.4±1.0, p<0.002) [Very Low quality evidence].

Other interventions

Mack et al (2010) conducted a single arm, cohort study including 20 patients of whom 16 were myeloma patients evaluating Viv-Arte training program including whole body vibration with Galileo training device (SKMT) for chemotherapy induced peripheral neuropathy and found that treatment was well tolerated in all patients [Very Low].

A large difference was observed with regard to locomotoric and sensoric multi dimensional tests pre and post treatment with pre-treatment paraesthesia of the feet measured on a scale of 1-10 showing the greatest change from pre-treatment to post treatment (median 8 (range: 1-10) versus median 2 (range: 0-7))

Cost effectiveness evidence

A literature review of published cost-effectiveness analyses did not identify any relevant papers for this topic. Whilst there were potential cost implications of making recommendations in this area, other questions in the guideline were agreed as higher priorities for economic evaluation. Consequently no further economic modelling was undertaken for this question.

Table 86: GRADE profile: What is the most effective way to manage neuropathy in patients with myeloma (graded dose reduction/anti-myeloma drug withdrawal/use of nutritional supplements/complementary therapies/TENS/active monitoring versus each other/standard care)?

Quality assessment							
No of studies	**Design**	**Limitations**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**	**Quality**
Resolution or improvement of symptoms							
6	observational studies	very serious[1]	no serious inconsistency	no serious indirectness[2]	no serious imprecision	none[3]	VERY LOW
Adverse Events							
2	observational studies	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	VERY LOW
Reduction/discontinuation of myeloma treatment							
1	observational studies	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none[3]	VERY LOW
Overall Survival							
1	observational studies	very serious[1,4]	no serious inconsistency	no serious indirectness	no serious imprecision	none	VERY LOW
Physical and Social Functioning							
5	observational studies	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	VERY LOW

1 All studies were single arm, no comparative studies with small sample sizes; 2 One study included non-myeloma patients however it was 4/20 patients who were not myeloma patients; 3 Dose-response is an outcome that is relevant to this topic however the sample sizes in the individual studies were too small to accurately assess the size of the effect; 4 Follow-up time does not appear to be long enough to make accurate assessments of overall survival

Recommendations	**Explain the symptoms of neuropathy to people with myeloma, and encourage them to tell their clinical team about any new, different or worsening neuropathic symptoms immediately.** **If people who are receiving bortezomib develop neuropathic symptoms, consider immediately:** • **switching to subcutaneous injections and/or** • **reducing to weekly doses and/or** • **reducing the dose.** **Consider reducing the dose if people are taking a drug other than bortezomib and develop neuropathic symptoms.** **Temporarily stop neuropathy-inducing myeloma treatments if people develop either of the following:** • **grade 2 neuropathy with pain** • **grade 3 or 4 neuropathy.**
Relative value placed on the outcomes considered	The Guideline Committee considered the outcomes of improvement or resolution of symptoms, quantitative sensory testing, overall survival, health-related quality of life, physical and social functioning, adverse events and reduction or early discontinuation of myeloma treatment to be the most relevant in identifying the most effective way to manage neuropathy in patients with myeloma. Of these, evidence was found for all outcomes except for overall survival, health-related quality of life and quantitative sensory testing. No additional outcomes were reported. When drafting the recommendations the Guideline Committee considered improvement/resolution of symptoms, adverse events and reduction or early discontinuation of treatment to be the most important as these outcomes are considered to be the most important to patients.
Quality of the evidence	The quality of the evidence was assessed by GRADE methodology and appropriate NICE checklists to be of very low quality for all outcomes. The Guideline Committee noted a number of issues with the evidence: the included studies were of small sample size and were non-comparative observational studies from single centres. Furthermore all the studies relied on self-reporting of outcomes through the use of questionnaires, leaving the studies at high risk of bias. As a result of these limitations and the low quality evidence the Guideline Committee made 'consider' recommendations rather than 'offer'.
Trade off between clinical benefits and harms	Based on the clinical evidence the Guideline Committee made a recommendation for dose modification of bortezomib to manage neuropathy. Based on their clinical experience, the Guideline Committee recommended that dose modification should be prompt to avoid irreversible damage, caused by further administration of bortezomib at the existing dosage. The Guideline Committee noted that the evidence for dose modification was limited to bortezomib treatment. However, based on their clinical experience, they agreed it was appropriate to make a similar recommendation for dose reduction of any other treatment related neuropathy, to avoid irreversible damage. Based on clinical experience, the Guideline Committee considered that the most significant clinical consequence of treatment-induced neuropathy is neuropathy interfering with function and

	activities of daily living. They therefore recommended treatment interruption if this develops. The Guideline Committee noted that people can develop neuropathy both whilst on treatment and between treatments and it is important that this is identified before the administration of the next dose of myeloma treatment. Based on their clinical experience, they therefore also recommended that the symptoms of neuropathy should be explained to the patient, to improve their awareness of these symptoms and encourage patients to report these symptoms to their clinical team. The Guideline Committee concluded that the benefits would include a reduction in incidence, duration and severity of treatment emergent neuropathy, an improvement in quality of life, increased availability of a safe and non-pharmacological treatment for neuropathy and a reduction in early discontinuation of therapy (resulting from the development of neuropathy), potentially leading to improved outcomes. The Guideline Committee agreed that a potential harm could be patients receiving a sub-optimal treatment dose (as a result of the dose modification to manage neuropathy), potentially leading to inferior outcomes. They agreed that the recommendations made balanced trying to maximise treatment efficacy whilst maintaining quality of life.
Trade-off between net health benefits and resource use	The Guideline Committee noted that no relevant published economic evaluations had been identified and no additional economic analysis had been undertaken in this area. On discussing the potential costs and savings as a result of the recommendations made the Guideline Committee agreed that dose modification was already being done in most centres and so this was unlikely to incur any additional costs. With regards to patient awareness this was agreed to be cost neutral. It was agreed that there may be increased costs resulting from the recommendation on acupuncture, but that any increase in cost was likely to be offset by a reduction in analgesic drug costs and improvement in functioning. The Guideline Committee agreed that the net effect of these recommendations would probably be an increase in costs but there was uncertainty over the size of this increase as not all patients get neuropathy and therefore it was uncertain what the uptake of these recommendations would be.
Other considerations	The Guideline Committee noted that acupuncture is not currently available everywhere and where it is available; it is variably supplied (by the NHS, charities/hospices and the private sector). The Guideline Committee acknowledged that the recommendation to consider acupuncture could represent a change in practice, however the variability of current provision meant the Guideline Committee were not clear what the extent of a change in practice would be.

9.3 Preventing thrombosis

Venous thromboembolism (VTE) is a recognised complication of most cancers. This is particularly the case in myeloma because of the frequent combined occurrence of multiple thrombotic risk factors including age, immobility, fractures and infection in addition to the myeloma diagnosis. Newer treatment approaches involving immunomodulatory drugs are well recognised to increase the risk of both venous and arterial thrombotic events. The risk of VTE is greatest during the first few months of treatment, particularly using combination

chemotherapy that involves immunomodulatory drugs. VTE remains a significant cause of morbidity and mortality.

A range of preventative strategies have been used to reduce the risk of thrombotic events including anti platelet agents, low molecular weight heparin, vitamin K antagonists such as warfarin and the novel oral anticoagulants (non of which are licensed for primary prophylaxis). All of these treatments carry with them practical advantages and disadvantages including differing routes of administration, monitoring requirements and side effect profile. All will increase the risk of haemorrhage.

Clinical practice varies across the country and therefore there is a need to establish standard practice for prevention of thrombosis. Also there is little evidence on safety issues or adherence to treatment.

Clinical question: What is the most effective method for prevention of thrombosis in patients with myeloma?

Clinical evidence (see also Appendix G)

See Tables 87 - 92.

Thrombosis

For the outcome of thrombosis there was very low to low quality evidence from mostly observational studies. From these studies it is clear that prophylaxis with aspirin, LMWH or VKA is effective in preventing thrombosis in myeloma patients as fewer thrombotic events occurred in patients receiving any of these interventions compared to patients that did not receive any prophylaxis. However it is unclear from these studies which intervention is most effective at preventing thrombosis. Most of these studies were not randomised as they were not designed to answer the question of thrombosis prophylaxis.

There was moderate quality evidence from two large RCTs studies (from the same research group) of thromboprophylaxis in myeloma. The first studied thromboprophylaxis with LMWH, aspirin or VKA in 667 newly diagnosed myeloma patients (Palumbo et al., 2011). Patients treated with thalidomide-containing regimens were randomly assigned in a 1:1:1 ratio to receive LMWH (enoxaparin 40 mg/d), aspirin (100 mg/d), or VKA (warfarin 1.25 mg/d). The investigators concluded that LMWH was better than VKA in reducing the incidence of thrombosis events but was no different from aspirin. In another study of newly diagnosed myeloma patients treated with lenalidomide (Larocca et al 2012), 342 patients were randomized to aspirin (100 mg/d) or LMWH (enoxaparin 40 mg/d). The data replicated the results from Palumbo et al in that there was no significant difference in the incidence of thrombosis events between aspirin and LMWH. These RCTs are limited as the participants are not representative of the entire myeloma population as high risk individuals (patients at high risk of thromboembolic events such as patients with a previous history of thromboembolism, cardiac disease, infections, immobilization or surgery) were excluded.

Only 1 study (including 542 myeloma patients) stratified results according to risk for thrombosis (Leleu et al., 2013). They found the lowest incidence of thrombosis in the patients at highest risk (incidence of thrombosis 3% in high risk individuals, 6% in those at intermediate risk and 7% in those at low risk) because these patients received better and optimized prophylaxis with LMWH and VKA compared to low risk patients who mostly received aspirin.

Bleeding events

There was very low to low quality evidence from 2 observational studies and moderate quality evidence from 2 RCTs for incidence of bleeding events.

The data from the observational studies indicates that bleeding events are more likely in patients receiving prophylaxis with VKA, LMWH and aspirin compared to patients not receiving prophylaxis. The data also shows that VKA results in fewer bleeding events than aspirin and LMWH.

The data from the RCTs replicated this and also demonstrated a lower incidence of bleeding in patients receiving VKA compared to those receiving aspirin or LWMH. Patients receiving aspirin had the greatest risk of bleeding.

Death/mortality

Sudden death presumed to be a result of PE, MI or stroke was reported in 1 observational study and 1 RCT. There was no difference in the number of deaths between the different prophylactic interventions. However death was a rare event with too few events to make valid conclusions with regards to this outcome.

Adverse events

No evidence was found for this outcome.

Health-related quality of life

No evidence was found for this outcome.

Compliance/adherence and patient acceptability

No evidence was found for this outcome.

Table 87: GRADE profile: What is the most effective method for prevention of thrombosis in patients with myeloma (no prophylaxis versus aspirin)?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	aspirin	no prophylaxis	Relative (95% CI)	Absolute	Quality
Incidence of thromboembolic events											
4	observational studies	no serious limitations	Serious[1]	no serious indirectness	no serious imprecision	none	587	861	-	-0.2% to 39% fewer patients receiving aspirin suffered a thromboembolic event compared to those receiving no prophylaxis.	VERY LOW
Incidence of bleeding											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none	307	81	-	4.9% fewer patients receiving no prophylaxis suffered a bleeding event compared to those receiving aspirin.	LOW

1 heterogeneity between populations

Table 88: GRADE profile: What is the most effective method for prevention of thrombosis in patients with myeloma (no prophylaxis versus vitamin K antagonists)?

Quality assessment	Summary of findings		
	No of patients	Effect	Quality

No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	no prophylaxis	VKA	Relative (95% CI)	Absolute	
Incidence of thromboembolic events											
4	observational studies	no serious limitations	Serious[1]	no serious indirectness	no serious imprecision	none	934	412	-	-1.2% to 15.7% fewer patients receiving VKA suffered a thromboembolic event compared to those receiving no prophylaxis.	VERY LOW
Incidence of bleeding											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none	81	48	-	1.7% fewer patients receiving no prophylaxis suffered a bleeding event compared to those receiving VKA.	LOW
Incidence of death											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	serious imprecision[2]	none	19	246	-	0.8% fewer patients receiving no prophylaxis died compared to those receiving LMWH.	VERY LOW

1 *1 heterogeneity between populations; 2 very low number of events*

Table 89: GRADE profile: What is the most effective method for prevention of thrombosis in patients with myeloma (no prophylaxis versus low molecular weight heparin)?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	no prophylaxis	LMWH	Relative (95% CI)	Absolute	Quality
Incidence of thromboembolic events											
3	observational studies	no serious limitations	Serious[1]	no serious indirectness	no serious imprecision	none	308	274	-	5% to 9% fewer patients receiving LMWH suffered a thromboembolic event compared to those receiving no prophylaxis.	VERY LOW
Incidence of bleeding											
2	observational studies	no serious limitations	Serious[1]	no serious indirectness	no serious imprecision	none	221	206	-	-4.7% to 0.6% fewer patients receiving LMWH suffered a bleeding event compared to those receiving no prophylaxis.	VERY LOW

1 heterogeneity between populations

Table 90: GRADE profile: What is the most effective method for prevention of thrombosis in patients with myeloma (aspirin versus vitamin K antagonists)?

Quality assessment	Summary of findings		
	No of patients	Effect	Quality

No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	aspirin	VKA	Relative (95% CI)	Absolute	
Incidence of thromboembolic events											
3	observational studies	no serious limitations	Serious[1]	no serious indirectness	no serious imprecision	none	679	146	-	-1% to 7% fewer patients receiving VKA suffered a thromboembolic event compared to those receiving aspirin.	VERY LOW
Incidence of thromboembolic event											
1	randomized trials	Serious [2,3,4]	no serious inconsistency	no serious indirectness	no serious imprecision	none	220	220	-	2.3% fewer patients receiving aspirin suffered a thromboembolic event compared to those receiving VKA.	MODERATE
Incidence of bleeding											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none	307	48	-	3.2% fewer patients receiving VKA suffered a bleeding event compared to those receiving aspirin.	LOW
Incidence of bleeding											
1	randomized trials	Serious [2,3,4]	no serious inconsistency	no serious indirectness	no serious imprecision	none	220	220	-	3.5% fewer patients receiving VKA suffered a bleeding event compared to those receiving aspirin.	MODERATE
Incidence of death											

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	aspirin	VKA	Relative (95% CI)	Absolute	Quality
1	randomized trials	Serious [2,3,4]	no serious inconsistency	no serious indirectness	serious imprecision [5]	none	220	220	-	0.4% fewer patients receiving aspirin died compared to those receiving VKA.	LOW

1 *1 heterogeneity between populations; 2 Open-label trial (not blinded).; 3 selection bias - high risk individuals excluded; 4 No placebo. However it would not be ethical to*
2 *include a placebo with the high risk of thrombosis; 5 very low number of events*

3 **Table 91: GRADE profile: What is the most effective method for prevention of thrombosis in patients with myeloma (aspirin versus**
4 **low molecular weight heparin)?**

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	aspirin	LMWH	Relative (95% CI)	Absolute	Quality
Incidence of thromboembolic events											
2	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none	472	108	-	4% to 7% fewer patients receiving LMWH suffered a thromboembolic event compared to those receiving aspirin.	LOW
Incidence of thromboembolic events											
2	randomized	Serious	no serious	no serious	no serious	none	396	385	-	1.1% to 2.7% fewer patients	MODERATE

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	aspirin	LMWH	Relative (95% CI)	Absolute	Quality
	trials	[1,2,3]	inconsistency	indirectness	imprecision					receiving LMWH suffered a thromboembolic event compared to those receiving aspirin.	
Incidence of bleeding											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none	307	88	-	0.2% fewer patients receiving LMWH suffered a bleeding event compared to those receiving aspirin.	LOW
Incidence of bleeding											
2	randomized trials	Serious [1,2,3]	no serious inconsistency	no serious indirectness	no serious imprecision	none	396	385	-	-0.6%to 2.6% fewer patients receiving LMWH suffered a bleeding event compared to those receiving aspirin.	MODERATE
Incidence of death											

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	aspirin	LMWH	Relative (95% CI)	Absolute	Quality
1	randomized trials	Serious [1,2,3]	no serious inconsistency	no serious indirectness	serious imprecision[4]	none	220	219	-	There was no difference in the numbers of sudden deaths between patients receiving aspirin and those receiving LMWH.	LOW

1 *1 Open-label trial (not blinded); 2 Selection bias - high risk individuals excluded; 3 No placebo. However it would not be ethical to include a placebo with the high risk of*
2 *thrombosis; 4 very low number of events*

3 **Table 92: GRADE profile: What is the most effective method for prevention of thrombosis in patients with myeloma (vitamin K**
4 **antagonists versus low molecular weight heparin)?**

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	VKA	LMWH	Relative (95% CI)	Absolute	Quality
Incidence of thromboembolic events											
2	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none	679	146	-	-3% to 16.7% fewer patients receiving LMWH suffered a thromboembolic event compared to those	LOW

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	VKA	LMWH	Relative (95% CI)	Absolute	Quality
										receiving VKA.	
Incidence of thromboembolic events											
1	randomized trials	Serious [1,2,3]	no serious inconsistency	no serious indirectness	no serious imprecision	none	220	219	-	5% fewer patients receiving LMWH suffered a thromboembolic event compared to those receiving VKA.	MODERATE
Incidence of bleeding											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none	48	88	-	3% fewer patients receiving VKA suffered a bleeding event compared to those receiving LMWH.	LOW
Incidence of bleeding											
1	randomized trials	Serious [1,2,3]	no serious inconsistency	no serious indirectness	no serious imprecision	none	220	219	-	0.9% fewer patients receiving VKA suffered a bleeding event compared to those receiving	MODERATE

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	VKA	LMWH	Relative (95% CI)	Absolute	Quality
										LMWH.	
Incidence of death											
1	randomized trials	Serious [1,2,3]	no serious inconsistency	no serious indirectness	serious imprecision [4]	none	220	219	-	0.4% fewer patients receiving LMWH died compared to those receiving VKA.	LOW

1 *1 Open-label trial (not blinded); 2 Selection bias - high risk individuals excluded; 3 No placebo. However it would not be ethical to include a placebo with the high risk of*
2 *thrombosis; 4 very low number of events*

3

1 Cost effectiveness evidence

2 A literature review of published cost-effectiveness analyses did not identify any relevant
3 papers for this topic. Whilst there were potential cost implications of making
4 recommendations in this area, other questions in the guideline were agreed as higher
5 priorities for economic evaluation. Consequently no further economic modelling was
6 undertaken for this question.

7

Recommendations	**For people with myeloma who are starting immunomodulatory drugs, offer thromboprophylaxis with either:** • **low molecular weight heparin (LMWH) at a prophylactic dose, or** • **vitamin K antagonists at a therapeutic dose, to maintain an international normalised ratio (INR) of 2–3.** **If LMWH or vitamin K antagonists are unsuitable, consider low-dose aspirin[c].** **When starting thromboprophylaxis, assess the risk factors, contraindications and practicalities of each prophylactic strategy.** **Do not offer fixed low-dose vitamin K antagonists for thromboprophylaxis to people with myeloma who are starting immunomodulatory drugs.** **Consider switching thromboprophylaxis to low-dose aspirin for people who:** • **are taking immunomodulatory drugs and** • **have achieved maximum response and** • **have no high risk factors.**
Relative value placed on the outcomes considered	The Guideline Committee considered the outcomes of venous thromboembolisim (VTE) rate, arterial thrombosis rate, bleeding events, adverse events, death/mortality, health-related quality of life, compliance/adherence and patient acceptability to be the most relevant in identifying the most effective method for the prevention of thrombosis in patients with myeloma. Of these, evidence was only found for VTE rate, bleeding events, and death/mortality. No other outcomes were reported. When drafting the recommendations the Guideline Committee considered VTE rate and bleeding events to be the most important as these outcomes have serious consequences for the patient and therefore it is hoped that the recommendations would prevent them. Although no evidence was found for the outcome of patient acceptability, the Guideline Committee also considered this to be important as the recommendations will result in practical issues surrounding the need to have daily injections and/or more frequent visits to the clinic.

c At the time of consultation (August 2015), aspirin did not have a UK marketing authorisation for this indication. The prescriber should follow relevant professional guidance, taking full responsibility for the decision. Informed consent should be obtained and documented. See the General Medical Council's Prescribing guidance: prescribing unlicensed medicines for further information.

	The Guideline Committee agreed that the evidence for the outcomes of arterial thrombosis and death/mortality should not be used when drafting recommendations as there were too few events to be able to make valid conclusions.
Quality of the evidence	The quality of the evidence was assessed using GRADE methodology and appropriate NICE Checklists to be of very low to moderate quality for the outcomes of VTE rate and bleeding events and very low to low for the outcome of death. The Guideline Committee noted that 3 studies (including the 2 RCTs) had excluded high risk individuals. The Guideline Committee therefore applied more weight to Leleu at al 2013 which had reported on all risk groups when making recommendations according to risk category. The Guideline Committee noted that in Leleu et al (2013) high-risk patients had a higher intensity of anticoagulation and the lowest risk of VTE (of all risk groups) indicating that the therapy they received was the most effective. Since Leleu et al. (2013) included data on all risk groups, the Guideline Committee agreed to use this data to inform the recommendations. The Guideline Committee noted that Palumbo et al (2011) was not powered to detect differences between aspirin and low molecular weight heparin in VTE rate. Consequently they put more weight on data from the observational studies which showed low molecular weight heparin and warfarin were better than aspirin. The Guideline Committee noted that many of the studies were European and used a dose of aspirin not available in the UK. Therefore when making recommendations the dose recommended was the standard UK prophylactic dose. The Guideline Committee made a research recommendation to investigate the effectiveness of novel oral anticoagulants due to the absence of evidence on these. These are drugs licensed and used to prevent thrombosis for other conditions but not in myeloma.
Trade off between clinical benefits and harms	The Guideline Committee were concerned about the cumulative risk of bleeding for those patients on long term anticoagulation therapy. Although the evidence appeared to show that bleeding risk was lower for vitamin K antagonists compared with aspirin and low molecular weight heparin, the Guideline Committee noted that the studies were not powered to detect this difference and the differences shown were not statistically significant. Based on their clinical experience, the Guideline Committee were aware that aspirin has been shown to have a lower bleeding risk in large studies in the general medical population. Furthermore, switching from low molecular weight heparin or vitamin K antagonists in suitable patients, offers an improvement in quality of life. They therefore recommended switching thromboprophylaxis to low-dose aspirin in people who have achieved maximum response with their immunomodulatory drug and have no high-risk factors. The Guideline Committee concluded that the benefits would be a reduction in VTE rate in patients taking immunomodulatory drugs. Also by recommending thromboprophylaxis, the Guideline Committee considered that discontinuation of anti-myeloma therapy would be avoided, resulting in better outcomes for the patient.

	The Guideline Committee acknowledged that there is a risk of bleeding associated with the prophylaxis recommended. They balanced this risk by recommending aspirin for stable disease, which has a reduced risk of bleeding. The Guideline Committee acknowledged that the potential harms of the use of low molecular weight heparin were the pain of injection and inconvenience from daily administration of this treatment. Use of vitamin K antagonists requires an increased frequency of blood monitoring for International Normalised Ratio. The Guideline Committee balanced these harms against the benefits by recommending that the risk factors, contraindications and practicalities of each treatment strategy be assessed to determine the most appropriate treatment.
Trade-off between net health benefits and resource use	The Guideline Committee noted that no relevant published economic evaluations had been identified and no additional economic analysis had been undertaken in this area. The Guideline Committee agreed that the recommendations made were unlikely to result in any additional costs as thromboprophylaxis is already commonly used in patients with myeloma.
Other	The Guideline Committee were aware of existing NICE guidance on the prevention of venous thromboembolism (CG92). They noted that patients with myeloma are normally out-patients and therefore would not be covered by CG92 as this only made recommendations for cancer in-patients. However, the Guideline Committee acknowledged that there would be a minority of patients with myeloma who may be receiving in-patient treatment. The Guideline Committee noted that the evidence base for the recommendations for cancer in-patients in CG92 were based on studies of general medical patients which included some cancer patients. They considered it unlikely that patients with myeloma formed a significant proportion of these cancer patients. In addition, myeloma patients receiving immunomodulatory drugs are at a particularly high risk of developing venous thrombosis, regardless of ambulatory status. The Guideline Committee also noted that the evidence review for this question had identified studies directly relevant to the myeloma population that showed strong evidence of benefit for the interventions recommended. For these reasons it was agreed not to cross-reference the recommendation for prevention of venous thromboembolism in cancer in-patients in CG92.

1

Research recommendation	**A randomised controlled trial should be carried out comparing new oral anticoagulants compared with low molecular weight heparin in people with myeloma who are starting treatment with immunomodulatory drugs. Outcomes of interest are venous thromboembolism rate, arterial thrombosis rate, bleeding rate, progression-free survival, overall survival, patient compliance.**
Why is this important?	New oral anticoagulants are licensed and used to prevent thrombosis in conditions other than myeloma. There is a lack of evidence of their effectiveness in preventing drug induced thrombosis in patients with myeloma. There is evidence to support the use of low molecular weight heparin for the prevention of thrombosis. However this can only be given subcutaneously and so it is a further patient discomfort. Further research comparing these interventions is therefore desirable.

9.41 Managing fatigue

Cancer-related fatigue is a persistent tiredness or lethargy which affects the ability to complete activities of daily living. Almost all people with myeloma will experience fatigue at varying degrees at some point, either caused by the myeloma itself or by its treatment. It is recognised that cancer-related fatigue is different to and more severe than normal fatigue as it tends to last longer and be exhausting and debilitating.

Adopting strategies to manage fatigue can help improve quality of life. Some causes of fatigue are easily correctable, e.g. anaemia. A variety of interventions such as psychostimulants, over-the-counter stimulants and 'energy drinks', exercise programmes, complementary therapies, dietary intervention, rest and sleep hygiene education have been postulated to improve cancer-related fatigue. However, there is uncertainty over their effectiveness and the optimal way of using them.

There is also considerable variation between centres on the use and availability of treatments for cancer-related fatigue. Geographical variation also affects when patients are referred to other specialists.

Clinical question: Which interventions are most effective in reducing fatigue in patients having treatment for myeloma?

Clinical evidence (see also Appendix G)

See Tables 93 – 95.

Reduction of fatigue

Moderate quality evidence from a randomized trial (Coleman et al, 2012) suggests that an individualized exercise program is not effective for reducing fatigue in myeloma patients. There was very little difference in the fatigues scores (FACT and POMS) between patients undertaking a home-based individualized exercise program (HBIEP), combining aerobic and strength resistance training, and the control group receiving the current best practice recommendation to walk 20 minutes three times a week (usual care).

Moderate quality evidence from a randomized trial (Berenson et al, 2015) including 42 patients, suggests that moderately fatigued patients with myeloma treated with placebo for 28 days show similar improvements in self-reported fatigue to those treated with armodafinil.

Performance (aerobic capacity)

Moderate quality evidence from a randomized trial (Coleman et al, 2012) suggests that an individualized exercise program is not effective for improving aerobic capacity (measured by distance walked in 6 minutes) when compared to usual care (Coleman et al, 2012). Patients in the exercise program group walked on average an additional 50 feet compared to the usual care group but the difference was not statistically significant.

ECOG performance score

Moderate quality evidence from a randomized trial (Dammacco et al., 2001) suggests that that epoetin alfa can improve ECOG performance score in myeloma patients when compared to placebo. 20% of patients receiving epoetin alfa showed a one-point improvement in ECOG performance score compared to 6% of those receiving placebo.

Daytime and night-time sleep (ActiGraph)

Moderate quality evidence from a randomized trial (Coleman et al, 2012) suggests that an individualized exercise program is not effective for improving sleep in myeloma patient. There was very little difference in minutes of daytime and nighttime sleep between patients undertaking the HBIEP, coming aerobic and strength resistance training, and the control group receiving the current best practice recommendation to walk 20 minutes three times a week (usual care).

Quality of life

Moderate quality evidence from a randomized trial (Dammacco et al., 2001) suggests that that epoetin alfa can improve QOL in myeloma patients when compared to placebo. Within-group changes from baseline to week 12 revealed statistically significant improvement in emotional reactions, social interaction, energy and ability to do daily activities in patients treated with epoetin alfa. Placebo-treated patients, in contrast, showed no significant improvement except in sleep. Between-group differences in effect on QOL were not detected.

Moderate quality evidence from a randomized trial (Berenson et al, 2015) including 42 patients, suggests that moderately fatigued patients with myeloma treated with placebo for 28 days show similar improvements in self-reported quality of life to those treated with armodafinil.

Adverse events

High quality evidence from a randomized trial (Dammacco et al., 2001) suggests that adverse events are similar in myeloma patients receiving epoetin alfa and myeloma patients receiving placebo. No differences were found for overall incidence of adverse events (72.5% epoetin alfa-treated; 75.0% placebo-treated). Type and frequency of individual adverse events were similar throughout the study. The most commonly reported adverse events in either treatment group were fever, pain and leucopenia.

Exercise tolerance, Muscle function, Mobility – physical and social functioning, Dependency for activities of daily living

The literature searches did not find evidence for these outcomes.

Table 93: GRADE profile: Which interventions are most effective in reducing fatigue in patients having treatment for myeloma (individualised exercise program versus usual care)?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	an individualized exercise program	usual care	Relative (95% CI)	Absolute	Quality
Fatigue (POMS and FACT-F)											
1[2]	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	91	75	-	The effect of exercise was minimal on decreasing fatigue: At the end of the 15 week experimental period patients in the intervention group had a mean FACT fatigue score of 31.34 (scores range from 0-52 with higher scores indicating less fatigue) and a mean POMS fatigue score of 10.63	MODERATE

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	an individualized exercise program	usual care	Relative (95% CI)	Absolute	Quality
										(scores range from 0-28 with higher scores indicating less fatigue). Patients in the control group had a mean FACT fatigue score of 31.71 a mean POMS fatigue score of 10.92.	
Daytime and night-time sleep (actigraph)											
1[2]	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	91	75	-	The effect of exercise was minimal on improving sleep: At the end of the 15 week experimental period patients in the intervention group had a mean of 411.7	MODERATE

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	an individualized exercise program	usual care	Relative (95% CI)	Absolute	Quality
										minutes nighttime and 113.17 daytime sleep, whilst patients in the control group had a mean 414.33 minutes nighttime and 114 daytime sleep.	
Performance (aerobic capacity) – measured by distance walked in 6 minutes											
1[2]	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	91	75	-	The effect of exercise was minimal on improving performance: At the end of the 15 week experimental period patients in the intervention group walked 1594.69 feet in 6 minutes compared to those in the	MODERATE

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	an individualized exercise program	usual care	Relative (95% CI)	Absolute	Quality
										control group who walked 1545.07 feet in 6 minutes.	

1 The patients self-reported their compliance with the exercise program. Observation of the exercise and activity was not possible because this was a home-based program.
Exercise was individualized for each patient so no consistent pattern of exercise across the population. 2 Coleman et al., 2012.

Table 94: GRADE profile: Which interventions are most effective in reducing fatigue in patients having treatment for myeloma (epoetin alfa versus placebo)?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	epoetin alfa	placebo	Relative (95% CI)	Absolute	Quality
Quality of life											
1[2]	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	66	72	-	Improvement in more QOL measures with epoetin than with placebo. No Absolute data reported.	MODERATE
ECOG performance score											
1[2]	randomised trials	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	66	66	-	13.6% more patients in the intervention group had a 1-point improvement	MODERATE

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	epoetin alfa	placebo	Relative (95% CI)	Absolute	Quality
										in performance score compared to the placebo group.	
Adverse events											
1[2]	randomised trials	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none	50/69 (72.5 %)	57/76 (75%)	-	2.5% fewer patients in the intervention group experienced an adverse event, compared to the placebo group.	HIGH

1 *1 Changes in functional status and QOL in the study reported here were secondary efficacy assessments, and the study was not powered to measure absolute change, but*
2 *rather statistical trends; 2 Dammacco et al., 2001*

3 **Table 95: GRADE profile: Which interventions are most effective in reducing fatigue in patients having treatment for myeloma**
4 **(armodafinil versus placebo)?**

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	armodafinil	placebo-first	Relative (95% CI)	Absolute	Quality
Quality of life (FACIT-G; higher scores better; measured after 28 days of treatment)											
1[2]	randomised trials	no serious limitations	no serious inconsistency	no serious indirectness	serious imprecision	none	19	23	-	FACIT-G was 75.8 (12.9) in placebo-	MODERATE

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	armodafinil	placebo-first	Relative (95% CI)	Absolute	Quality
										first group and 68.5 (20.5) in the treatment only group (P=0.377)	
Fatigue (BFI; lower scores better; measured after 28 days of treatment)											
1[2]	randomised trials	no serious limitations	no serious inconsistency	no serious indirectness	serious imprecision	none	19	23	-	BFI was 41.5 (18.4) in placebo-first group and 48.8 (22.4) in the treatment only group (P=0.289)	MODERATE
Serious adverse events (during 28 days of treatment)											
1[2]	randomised trials	no serious limitations	no serious inconsistency	no serious indirectness	serious imprecision	none	2/19	0/23	-	Overall toxicities were similar between the two groups. 4% of adverse events were deemed to be drug	MODERATE

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	armodafinil	placebo-first	Relative (95% CI)	Absolute	Quality
										related.	

1 Changes in functional status and QOL in the study reported here were secondary efficacy assessments, and the study was not powered to measure absolute change, but rather statistical trends; 2 Berenson et al (2015)

Cost effectiveness evidence

A literature review of published cost-effectiveness analyses did not identify any relevant papers for this topic. Whilst there were potential cost implications of making recommendations in this area, other questions in the guideline were agreed as higher priorities for economic evaluation. Consequently no further economic modelling was undertaken for this question.

Recommendations	**If other treatable causes of anaemia have been excluded, consider erythropoietin analogues (adjusted to maintain a steady state of haemoglobin at 110–120 g/litre) to improve fatigue in people with myeloma who have symptomatic anaemia.**
Relative value placed on the outcomes considered	The Guideline Committee considered the outcomes of reduction of fatigue, performance status, daytime sleepiness, quality of life, exercise tolerance, actimetry, muscle function, physical and social functioning, dependency for activities of daily living, adverse events and patient reported outcomes to be the most relevant in identifying the most effective intervention for reducing fatigue in patients being treated for myeloma. Of these, evidence was only found for reduction in fatigue, performance (aerobic capacity), daytime sleepiness, performance status, quality of life and adverse events. No other outcomes were reported. The Guideline Committee considered performance status and quality of life to be the most important outcomes when determining their recommendations for erythropoietin analogues as they have the biggest impact on patients and are good surrogates for the outcome of reduction in fatigue – which was not reported by the evidence for this intervention. Although there was evidence on adverse events, such as venous thrombosis, the Guideline Committee did not consider them significant because by maintaining haemoglobin below 120 g/litre these are rare. The Guideline Committee noted that reduction in fatigue, performance (aerobic capacity) and daytime sleepiness were reported for the intervention of exercise. However the Guideline Committee agreed not to base any recommendations on these outcomes, the reasons for which are reported in the 'Quality of the evidence' section.
Quality of the evidence	The evidence for performance status and quality of life was moderate quality as assessed by GRADE. The Guideline Committee noted that in Dammacco et al. 2001, fatigue was not measured specifically – surrogates were used instead. However the Guideline Committee agreed that these surrogates were reasonable. In addition, the measurement of quality of life was only a secondary efficacy assessment and the study was not powered to measure absolute change but rather statistical trends. The Guideline Committee also noted that the study used outcome measures for quality of life which have now been superseded. Consequently the Guideline Committee only made a "consider" recommendation. The Guideline Committee noted that there were several limitations with the evidence on exercise programmes (Coleman et al, 2012). Firstly, patients self-reported their compliance with the exercise program as observation of the exercise and activity was not possible

	because this was a home-based program. Secondly the exercise was individualised for each patient and there wasn't a consistent pattern of exercise across the population in the study. In addition the Guideline Committee thought that since this was a single centre study it was possible the control arm may have been influenced by the overall practice in the centre to promote exercise. The Guideline Committee acknowledged that evidence had only been identified for two of the interventions of interest. They debated whether or not to recommend further research into these interventions but agreed that this research was unlikely to be practical. The Guideline Committee noted that research on exercise for fatigue in myeloma patients is currently ongoing.
Trade off between clinical benefits and harms	The Guideline Committee noted that studies in patients with solid tumours have previously shown a beneficial effect of exercise programmes on fatigue. However, the Guideline Committee acknowledged that Dammacco et al. 2001 had not shown the same result. Because of the limitations with the evidence, the Guideline Committee decided not to make any recommendations on the use of exercise programmes to manage fatigue in patients with myeloma. The Guideline Committee considered that the potential benefits would be reduced fatigue and better quality of life for patients with myeloma. There was also the potential to avoid transfusions. However the Guideline Committee noted that people taking erythropoietin analogues (especially those taking immunomodulatory drugs to manage their myeloma) are at increased risk of thrombotic events if their haemoglobin levels go above 120g/litre. They balanced this potential harm against the benefits by recommending that administration of erythropoietin analogues should be adjusted to maintain a steady state haemoglobin between 110 and 120g/L. The Guideline Committee acknowledged that the study which informed these recommendations had looked specifically at erythropoietin alpha. However, the Guideline Committee noted that there were some local variations in access to erythropoietin alpha. In addition, based on their clinical experience, the Guideline Committee considered that all erythropoietin analogues had equivalent effectiveness. Therefore they agreed to recommend the use of erythropoietin analogues to allow more flexible prescribing.
Trade-off between net health benefits and resource use	The Guideline Committee noted that no relevant, published economic evaluations had been identified and no additional economic analysis had been undertaken in this area. The Guideline Committee agreed that as a result of the recommendations made, more erythropoietin analogues would be prescribed which would be associated with an increase in costs. However there were likely to be some cost savings resulting from a reduction in transfusions. The net effect was uncertain but likely to be small. The Guideline Committee agreed that fatigue has a significant impact on quality of life for patients with myeloma. As such the benefits from the improvement in fatigue would outweigh any potential small increase in costs.

References

Berenson JR (2015). A phase 3 trial of armodafinil for the treatment of cancer-related fatigue for patients with multiple myeloma. Supportive Care in Cancer. 23: 1503-1512.

Blijlevens N (2013) In a high-dose melphalan setting, palifermin compared with placebo had no effect on oral mucositis or related patient's burden. Bone Marrow Transplantation 48(7): 966-971.

Boa T, Goloubeva O, Pelser C,et al (2014) A pilot study of acupuncture in treating bortezomib induced peripheral neuropathy in patients with multiple myeloma. Integrative Cancer Therapies 13;5:396-404

Cheuk-Daniel KL, Chiang-Alan KS, Lee Tsz Leung, et al (2011) Vaccines for prophylaxis of viral infections in patients with hematological malignancies. Cochrane.Database.of Systematic.Reviews. [3].

Cho J, Kang D, Lee JY, et al (2014) Impact of dose modification on intravenous bortezomib-induced peripheral neuropathy in multiple myeloma patients. Support Cancer Care. 22(10): 2669-75

Coleman EA, Goodwin JA, Kennedy R, et al. (2012) Effects of exercise on fatigue, sleep, and performance: a randomized trial. Oncology Nursing Forum, 39: 468-477.

Cox JE, Campos S, Wu J et al. (2014) Efficacy of deferred dosing of granulocyte colony-stimulating factor in autologous hematopoietic transplantation for multiple myeloma. Bone Marrow Transplantation 49[2], 219-222.

Dammacco F, Castoldi G & Rödjer S. (2001) Efficacy of epoetin alfa in the treatment of anaemia of multiple myeloma. Br J Haematol. 113(1), 172-179.

Garcia MK, Cohen L, Guo Y, et al (2014) Electroacupuncture for thalidomide/bortezomib-induced peripheral neuropathy in multiple myeloma: a feasibility study. Journal of Haematology and Oncology 7;41

Larocca A, Cavallo F, Bringhen S, et al. (2012) Aspirin or enoxaparin thromboprophylaxis for patients with newly diagnosed multiple myeloma treated with lenalidomide. Blood, 119: 933-939.

Leleu X, Rodon P, Hulin C, et al. (2013) MELISSE, a large multicentric observational study to determine risk factors of venous thromboembolism in patients with multiple myeloma treated with immunomodulatory drugs. Thrombosis & Haemostasis, 110: 844-851.

Mack S et al (2010) The Viv-Arte training program supplemented by whole-body vibration training in the treatment of chemotherapy induced sensorimotor polyneuropathy. Onkologie 3;suppl 6:1-294 (abstract)

Oken MM, Pomeroy C, Weisdorf D, et al (1996) Prophylactic antibiotics for the prevention of early infection in multiple myeloma. American.journal of medicine 100(6): 624-628.

Ozkan HA, Ozer UG, Bal C, et al (2013) Daily vs every other day administration of G-CSF following autologous peripheral stem cell transplantation: A prospective randomized study. Transfusion and apheresis science 49(2): 163-167.

Palumbo A, Cavo M, Bringhen S, et al. (2011) Aspirin, warfarin, or enoxaparin thromboprophylaxis in patients with multiple myeloma treated with thalidomide: a phase III, open-label, randomized trial. Journal of Clinical Oncology. 29: 986-993.

Raanani P, Gafter-Gvili A, Paul M, et al (2009) Immunoglobulin Prophylaxis in Hematopoietic Stem Cell Transplantation: Systematic Review and Meta-Analysis. Journal of Clinical Oncology 27(5): 770-781..

Raanani P, Gafter-Gvili A, Paul M, et al. (2009b) Immunoglobulin prophylaxis in chronic lymphocytic leukemia and multiple myeloma: systematic review and meta-analysis. [Review] [20 refs]. Leukemia & lymphoma 50(5): 764-772.

Richardson PG, Sonneveld P, Schuster MW, et al (2009) Reversibility of symptomatic peripheral neuropathy with bortezomib in the phase III APEX trial in relapsed multiple myeloma: impact of a dose-modification guideline. British Journal of Haematology 144;895-903

Truini A, Biasiotta A, Di Stefano G, et al (2011) Palmitoylethanolamide restores myelinated fibre function in patients with chemotherapy induced painful neuropathy. CNS and Neurological Disorders- Drug Targets 10(8):916-920

Vesole DH, Oken MM, Heckler C, et al. (2012) Oral antibiotic prophylaxis of early infection in multiple myeloma: a URCC/ECOG randomized phase III study. Leukemia 26(12), 2517-2520

10 Monitoring

Myeloma is characterised by a remitting and relapsing clinical course. This means that most patients are not cured and will need continuing follow up as relapse can occur gradually or suddenly, and is unpredictable. Furthermore, many patients who are diagnosed with myeloma may not have symptoms and therefore do not need immediate treatment. Appropriate monitoring of these patients with smouldering myeloma is crucial to insure early detection of disease progression before the development of potentially irreversible complications such as spinal cord compression, bone fracture or renal failure.

Disease monitoring is performed by regular clinical assessment when patients attend for their out-patient clinics and by checking various laboratory tests performed on blood and/or urine. In addition, a number of radiological imaging techniques may be used to investigate skeletal related symptoms and disease activity. The frequency of monitoring patients who are on active treatment is often dictated by the nature of their chemotherapy protocols. However, there is variation in practice in the modality and frequency of monitoring patients who are not on active anti-myeloma therapy.

Clinical question: What is the optimal follow-up protocol for patients with myeloma (including duration, frequency, investigations and onward referral)?

Clinical evidence (see also Appendix G)

See Table 96.

No studies were identified that investigated follow-up protocols for patients with myeloma. One observational study was identified that reported on patient monitoring/follow up after first line autologous stem cell transplant (ASCT) and ten studies were identified that investigated individual follow-up tests and their accuracy in detecting disease in the follow-up setting. Diagnostic accuracy is not listed in our review protocol or PICO but on discussion with the sub-group for this topic it was agreed that this evidence was of interest and clinical relevance to determine how accurate these tests are in follow up setting.

Study quality

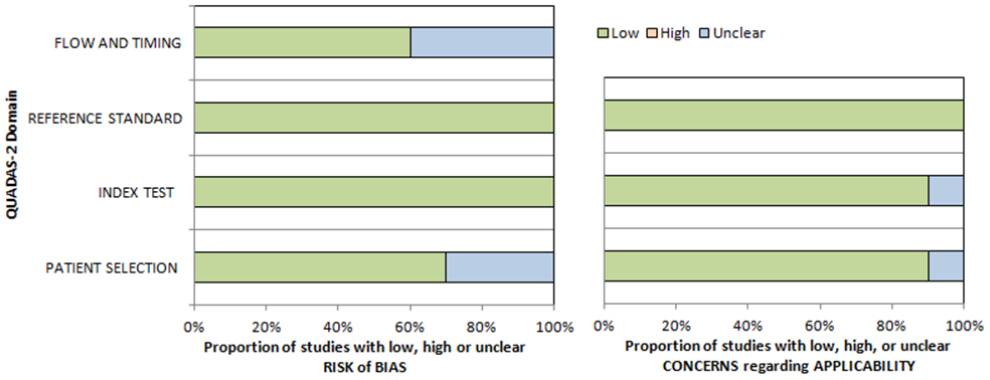
The QUADAS-2 assessment tool was used to evaluate risk of bias in these studies (Figure 11). Generally there was a low risk of bias across the studies and the studies were found to be applicable to the review question. For some of the studies the risk of bias is unclear as there was under-reporting in some studies with regards to the timing of the index and reference tests. Also some studies did not report the patient selection methods and so it was unclear whether a consecutive or random sample of patients had been recruited and if inappropriate exclusions had been avoided.

Other limitations of the included studies are that they are mostly single centre studies and many have small sample sizes. Furthermore, the patient populations studied are heterogeneous in that the patients included have undergone different treatments. However the studies aim to evaluate the performance of the diagnostic test for re-evaluation of myeloma post treatment rather than efficacy of a specific treatment approach, and these differences in prior treatment may well reflect clinical reality.

When comparing the results of the different diagnostic accuracy studies it is important to note that there is variability in the reference standards used in the different studies. Although a majority studies use the European group for blood and marrow transplantation criteria modified by the international uniform response criteria for multiple myeloma (panel of haematological and immunological parameters and bone marrow aspiration or biopsy where

appropriate) there are some studies which use different criteria to establish the presence of disease.

Figure 11: Risk of bias and applicability across studies



Observational data from 1 study

Evidence was identified from a retrospective study (Zamarin et al., 2013) examining the patterns of relapse or progression of disease (R/POD) in 273 patients treated with induction therapy followed by ASCT. The authors made several observations the most relevant ones being:

- The overwhelming majority of R/POD was associated with concurrent serological R/POD, with only a small percentage of patients (2%) presenting with symptomatic clinical disease in the absence of serological R/POD.
- A total of 85% had asymptomatic R/POD, first detected by serological testing, whereas 15% had symptomatic R/POD with aggressive disease, early R/POD and short survival, with poor cytogenetics and younger age identified as risk factors
- Although occult skeletal lesions were found in 40% of asymptomatic patients tested following serological R/POD, yearly skeletal surveys and urine testing were poor at heralding R/POD.

Diagnostic accuracy

10 diagnostic accuracy studies (with 22 - 168 patients) were identified and included in the evidence review (Bannas et al., 2012; Cascini et al., 2013; Derlin et al., 2012; Derline et al., 2013; Elliott et al., 2011; Fallahi et al., 2005; Harrington et al., 2009; Horger et al., 2007; Mele et al., 2007; Villa et al., 2005). They investigated lab tests, CD56 immunohistochemistry, and imaging methods including WB-MRI, WBLD-MDCT, FDG PET-CT and TC99MIBI. The results for diagnostic accuracy including sensitivity, specificity, positive predictive value and negative predictive value can be seen in table 1. The data indicate that lab tests and WBLD-MDCT are the most effective tests for detecting disease in follow up with the highest sensitivity, specificity and accuracy, whilst TC99MIBI and FDG PET-CT appear to be least effective.

Table 96: Diagnostic accuracy of various follow-up tests for detecting disease/remission following treatment (Note: variability in reference standard used in different studies)

Index tests	study	TP	FN	FP	TN	sensitivity	specificity	PPV	NPV	accuracy
Whole body MRI	Bannas et al., 2012	7	4	3	19	64%	86%	70%	83%	79%
	Cascini et al., 2013	9	0	8	12	100%	60%	33%	100%	72%
	Derlin et al., 2013	8	2	13	8	80%	38%	38%	80%	52%
FDG PET/CT	Elliott et al., 2011	12	6	2	17	67%	89%	86%	74%	78%
	Cascini et al., 2013	7	2	4	16	78%	80%	64%	9%	79%
	Derlin et al., 2012	NR	NR	NR	NR	55%	82%	82%	54%	66%
	Derlin et al., 2013	5	5	3	18	50%	86%	63%	78%	74%
WBLD-MDCT	Horger et al., 2007	411	2	1	25	99.5%	96.2%	99.8%	92.6%	99.3%
TC99MIBI bone scan	Fallahi et al., 2005	NR	NR	NR	NR	69%	100%	100%	61%	79%
	Villa et al., 2005	10	1	3	4	91%	57%	77%	80%	78%
	Mele et al., 2007	62	77	4	25	45%	86%	94%	25%	52%
Lab tests	Elliott et al., 2011	16	2	4	15	89%	79%	80%	88%	84%
	Horger et al., 2007	413	0	0	26	100%	100%	100%	100%	100%
Lab tests + PET/CT	Elliott et al., 2011	12	2	0	13	86%	100%	100%	87%	93%
CD56 immunohistochemistry	Harrington et al., 2009	59	15	3	50	80%	94%	95%	77%	86%

TP: true positive, FN: false negative, FP: false positive, TN: true negative, PPV: positive predictive value, NPV: negative predictive value, NR: not reported

1 Cost effectiveness evidence

2 A literature review of published cost-effectiveness analyses did not identify any relevant
3 papers for this topic. Whilst there were potential cost implications of making
4 recommendations in this area, other questions in the guideline were agreed as higher
5 priorities for economic evaluation. Consequently no further economic modelling was
6 undertaken for this question.

7

Recommendations	**Monitor people with smouldering myeloma every 3 months for the first 5 years, and then decide the frequency of further monitoring based on the long-term stability of the disease.**
	
	Monitor people who have completed myeloma treatment and recovered at least every 3 months. Take into account any risk factors for progression, such as:
	• **high-risk fluorescence in-situ hybridisation (FISH)**
	• **impaired renal function**
	• **disease presentation.**
	
	Monitoring for myeloma and smouldering myeloma should include:
	• **assessment of symptoms related to myeloma and myeloma treatment and**
	• **the following laboratory tests:**
	  ○ **full blood count**
	  ○ **renal function**
	  ○ **bone profile**
	  ○ **serum immunoglobulins and serum protein electrophoresis**
	  ○ **serum-free light-chain assay, if appropriate**
	
	Do not offer people with myeloma or smouldering myeloma routine skeletal surveys for disease monitoring.
	
	Consider symptom-directed imaging for people with myeloma or smouldering myeloma if any new bone symptoms develop.
	
	Consider whole-body MRI, spinal MRI or fluorodeoxyglucose positron emission tomography CT (FDG PET-CT) for people with myeloma or smouldering myeloma if there is serological relapse or disease progression.

Relative value placed on the outcomes considered	The Guideline Committee considered the outcomes of health related quality of life, progression free survival, overall survival, adverse events, patient reported outcome measures and patient experience to be the most relevant to identify the optimal follow-up protocol for patients with myeloma. However, no studies were identified that investigated follow-up protocols for patients with myeloma so no evidence was found on any of these outcomes. One observational study was identified that reported on patient monitoring/follow up after first line autologous stem cell transplant and ten studies were identified that investigated individual follow-up tests and their accuracy in detecting disease in the follow-up setting. The Guideline Committee agreed that whilst diagnostic accuracy was not listed in the review question this evidence was clinically relevant to determine how accurate these tests are in the follow-up setting. So

	these data and the outcomes of detection of relapse, detection of progression of disease and diagnostic accuracy were used to draft the recommendations.
Quality of the evidence	The quality of the evidence was assessed using QUADAS-2. Generally there was a low risk of bias across the studies and the studies were found to be applicable to the review question, but owing to the study design and that they were mostly single centre studies with small sample sizes, the quality of the evidence was assessed as being low to moderate. Owing to the fact that there was no direct evidence to answer this review question and the limitations with the available evidence, the Guideline Committee also relied on their clinical experience to draft recommendations. There was no evidence for the timing of follow-up tests but the Guideline Committee discussed optimal timings and agreed to recommend that the timing of monitoring should be 3 monthly to reflect the biology of the disease. It is known that some subgroups behave more aggressively so it was recommended that risk factors should be taken into account. The Guideline Committee also recommended that asymptomatic patients should be re-evaluated after 5 years of monitoring. This was based on the Guideline Committee's clinical knowledge of the long-term follow up of asymptomatic patients which is that if there is no progression after 5 years, the risk of progression drops dramatically. The Guideline Committee made recommendations on what tests the monitoring should include based both on the evidence of diagnostic accuracy and their clinical experience. Based solely on their clinical experience the Guideline Committee agreed that clinical assessment of symptoms related to myeloma and its treatment should be included as part of the monitoring to define the need for further imaging and prevent unnecessary imaging. Based on the evidence the Guideline Committee noted that annual skeletal surveys were poor at predicting relapse or progression of myeloma and therefore are of little benefit. The Guideline Committee also acknowledged based on the evidence and their clinical experience that there is cumulative radiation exposure from skeletal survey. As a result they agreed to recommend that skeletal surveys were not performed routinely as part of follow-up. The evidence suggested that urine testing was not useful and also the Guideline Committee's experience was that the number of patients in whom this is helpful is very small and it is no longer common practice. Most light chain progressions are more easily detected by the serum free-lite tests. However it was noted that the evidence was not strong enough to make a recommendation not to use these tests and the Guideline Committee agreed that they could be useful in certain circumstances. Therefore the Guideline Committee did not make any recommendations on urine testing.
Trade off between clinical benefits and harms	The Guideline Committee concluded that the recommendations made would result in standardisation as well as clarity of follow-up. Disease progression would be detected earlier – thereby avoiding new symptoms and unnecessary tests. The majority of progression is detected from blood tests rather than presenting with new symptoms. Patients may be made anxious by the frequency of testing but this is balanced against patients who desire more frequent testing.

Trade-off between net health benefits and resource use	The Guideline Committee noted that no relevant published economic evaluations had been identified and no additional economic analysis had been undertaken in this area. The Guideline Committee agreed that there would be cost savings from the recommendation not to do annual skeletal survey. The Guideline Committee also agreed that there were unlikely to be any additional costs associated with the other recommendations as these are standard practice for the follow-up of people with myeloma.
Other considerations	Athough skeletal surveys are still sometimes used in follow-up, the Guideline Committee felt that this was not common practice, so there is unlikely to be a significant change in practice as a result of the recommendations made not do do skeletal surveys.

References

Bannas P, Hentschel HB, Bley TA, et al. (2012) Diagnostic performance of whole-body MRI for the detection of persistent or relapsing disease in multiple myeloma after stem cell transplantation. European Radiology, 22: 2007-2012.

Cascini GL, Falcone C, Console D, et al. (2013) Whole-body MRI and PET/CT in multiple myeloma patients during staging and after treatment: personal experience in a longitudinal study. Radiologia Medica, 118: 930-948.

Derlin T, Weber C, Habermann CR, et al. (2012) 18F-FDG PET/CT for detection and localization of residual or recurrent disease in patients with multiple myeloma after stem cell transplantation. Eur J Nucl Med Mol Imaging 39(3), 493-500.

Derlin T, Peldschus K, Munster S, et al. (2013) Comparative diagnostic performance of 8F-FDG PET/CT versus whole-body MRI for determination of remission status in multiple myeloma after stem cell transplantation. European Radiology, 23: 570-578.

Elliott BM, Peti S, Osman K, et al. (2011) Combining FDG-PET/CT with laboratory data yields superior results for prediction of relapse in multiple myeloma. European Journal of Haematology, 86: 289-298.

Fallahi B, Saghari M, Fard Esfahani A, et al. (2005) The value of 99mTc-MIBI whole body scintigraphy in active and in remission multiple myeloma. Hellenic Journal of Nuclear Medicine, 8: 165-168.

Harrington AM, Hari P, Kroft SH (2009) Utility of CD56 immunohistochemical studies in follow-up of plasma cell myeloma. American Journal of Clinical Pathology, 132: 60-66.

Horger M, Kanz L, Denecke B, et al. (2007) The benefit of using whole-body, low-dose, nonenhanced, multidetector computed tomography for follow-up and therapy response monitoring in patients with multiple myeloma. Cancer, 109: 1617-1626.

Mele A, Offidani M, Visani G, et al. (2007) Technetium-99m sestamibi scintigraphy is sensitive and specific for the staging and the follow-up of patients with multiple myeloma: a multicentre study on 397 scans. British Journal of Haematology, 136: 729-735.

Villa G, Balleari E, Carletto M, et al. (2005) Staging and therapy monitoring of multiple myeloma by 99mTc- sestamibi scintigraphy: A five year single center experience. Journal of Experimental and Clinical Cancer Research, 24: 355-361.

Zamarin D, Giralt S, Landau H, et al. (2013) Patterns of relapse and progression in multiple myeloma patients after auto-SCT: implications for patients' monitoring after transplantation. Bone Marrow Transplantation 48, 419–424.

11 Managing relapsed myeloma

11.1 First relapse

NICE has developed a suite of technology appraisal guidance on myeloma. It has not been possible to develop recommendations on primary disease treatment, salvage therapy for relapsed myeloma and consolidation/maintenance therapy after primary management in this guideline due to published technology appraisals or those in development.

There is no significant new evidence that would lead to a change in the existing recommendations in the published appraisals, and following consultation with relevant stakeholders, it was decided that these appraisals should be moved to the static list, thus preserving the funding direction associated with any positive recommendations. It is therefore possible for these recommendations to be incorporated into any future clinical guideline, but they cannot be updated and replaced at this time.

Recommendations in this guideline will complement the existing technology appraisals.

For more information on the relationship between the technology appraisal and clinical guidelines programmes please see Updating technology appraisals in the context of clinical guidelines

Recommendations	**Bortezomib monotherapy is recommended as an option for the treatment of progressive multiple myeloma in people who are at first relapse having received one prior therapy and who have undergone, or are unsuitable for, bone marrow transplantation, under the following circumstances:** • **the response to bortezomib is measured using serum M protein after a maximum of four cycles of treatment, and treatment is continued only in people who have a complete or partial response (that is, reduction in serum M protein of 50% or more or, where serum M protein is not measurable, an appropriate alternative biochemical measure of response), and** • **the manufacturer rebates the full cost of bortezomib for people who, after a maximum of four cycles of treatment, have less than a partial response (as defined above). [This recommendation is from Bortezomib monotherapy for relapsed multiple myeloma (NICE technology appraisal guidance 129).]** **People currently receiving bortezomib monotherapy who do not meet the criteria in the recommendation above should have the option to continue therapy until they and their clinicians consider it appropriate to stop. [This recommendation is from Bortezomib monotherapy for relapsed multiple myeloma (NICE technology appraisal guidance 129).]**
	These recommendations are from Bortezomib monotherapy for relapsed multiple myeloma (NICE technology appraisal guidance 129). They were formulated by the technology appraisal and not by the guideline developers. They have been incorporated into this guideline in line with NICE procedures for developing clinical guidelines, and the evidence to support these recommendations can be found at www.nice.org.uk/TA129.

11.2 Second autologous stem cell transplant

For many years, some UK centres have advocated a second autologous stem cell transplant in those patients who relapse following a first transplant, whereas in other centres it is not a standard approach. Second autologous stem cell transplantation carries a risk of both treatment related morbidity and mortality but has been shown to prolong progression-free survival. It also involves a potentially lengthy inpatient admission and post-transplant recovery period that can impact on quality of life.

New therapies have resulted in improved outcomes for patients with relapsed disease meaning that more patients are likely to be suitable for a second autologous stem cell transplant. However there is increasing uncertainty over the benefit of a second autologus stem cell transplant compared to these newer drug therapies, in terms of improved outcomes. Factors of likely importance in determining the potential benefit of a second ASCT include depth and duration of response to first autologous stem cell transplant, age and performance status, co-morbidities and the cytogenetic profile of the patient.

Clinical question: In which patients with relapsed or refractory myeloma is a second autologous stem cell transplant more effective than other therapy?

Clinical evidence (see also Appendix G)

See Tables 97 – 107.

Comparative studies

From the literature search one RCT was identified (Cook et al., 2014). The study was a multicentre, randomised, open-label, phase 3 study comparing high-dose melphalan plus salvage autologous stem cell transplant (ASCT) (n=89) with weekly cyclophosphamide (n=85) in patients with relapsed multiple myeloma who had previously undergone ASCT and provides moderate quality evidence that time to progression is longer following treatment with salvage ASCT. Results of the predefined subgroup analysis of time to progression in Cook et al (2014) suggest that salvage ASCT is more effective than cyclophosphamide, irrespective of the quality of response to PAD re-induction and the concentration of β2-microglobulin at registration. Furthermore, ASCT was more effective than cyclophosphamide irrespective of the response duration to the initial ASCT, although time to progression was longer (TTP 24 months) in patients with a response lasting longer than 24 months after their first ASCT than in those with a response of 24 months or less (TTP 13 months). The relative effectiveness of salvage ASCT and cyclophosphamide in patients with adverse cytogenetics was uncertain due to the small number of patients with an adverse cytogenetic risk profile (n=13). Follow up in this study was not long enough (median 34 months) to confidently assess the effect of salvage therapy on survival.

Very low to low quality evidence from 4 retrospective comparative studies including 1134 patients suggests that outcomes are better (OS and/or PFS are longer) following treatment with a second ASCT compared to salvage systematic chemotherapy or alternative treatments in patients with relapsed myeloma who had previously undergone ASCT and belonging to the following subgroups: patients who respond well following ASCT1, (Cook et al., 2011), patients with longer time to progression after ASCT1 (Alvares et al., 2006; Cook et al., 2011), patients with a younger age (Cook et al., 2011), patients with a poor prognosis (as determined by time to progression after ASCT1 and ISS) (Yhim et al., 2013). Grovdal et al (2015) reported that both overall survival and time to next treatment were longer with a second ASCT than with either conventional cytotoxic chemotherapy or novel drugs (proteosome inhibitors or immunomodulatory drugs). There is the potential for selection bias in these retrospective comparative studies as the choice of therapy after relapse is often governed by a complex list of unmeasured factors that can potentially affect outcomes and

not all patients will be suitable for salvage ASCT. Two studies (Cook et al., 2011 and Yhim et al., 2013) matched patients in the intervention and comparator groups for a number of potential risk factors in an attempt to overcome selection bias. However, only a randomised trial can exclude such bias completely.

No evidence was identified for the outcomes treatment related morbidity and mortality, health related quality of life, adverse events, patient/carer/family acceptability and PROMs.

Prognostic studies

Moderate quality evidence from multivariate analysis in non-comparative retrospective studies that reported predictive factors (high quality prognostic factor studies but downgraded as comparative studies are better for answering the review question) suggest that in relapsed myeloma patients time to progression following an initial ASCT is an important predictor of survival following salvage ASCT. All 11 studies reported that a longer TTP after first ASCT was associated with longer PFS and/or OS after salvage ASCT. However the studies were inconsistent with regard to the length of remission that predicted improved survival outcomes, with reports of increased PFS and/or OS if TTP was more than 12 months (Olin et al., 2009; Fenk et al., 2011; Wirk et al., 2013), 18 months (Chow et al., 2013; Sellner et., 2013), 21.5 months (Auner et al., 2013) and 24 months (Jimenez-Zepeda et al., 2012; Lemieux et al., 2013; Michaelis et al., 2013).

Evidence also indicated a lack of response to initial ASCT (Olin et al., 2009), higher number of treatment regimens before second ASCT (Olin et al., 2009; Shah et al., 2012; Gonsalves et al., 2013), higher plasma cell labelling index at second ASCT (Gonsalves et al., 2013), elevated LDH at second ASCT (Sellner et al., 2013), adverse cytogenetics (Shah et al., 2012; Sellner et., 2013) age >60 (Lemieux et al., 2013) or age >65 (Olin et al., 2009), and being of african-american ethnicity (Shah et al., 2012) was predictive of worse survival outcomes. Whilst disease status (> PR) at salvage ASCT (Auner et al., 2013) and ISS stage I before salvage ASCT (Sellner et al., 2013) was predictive of better survival outcomes.

Myeloma subtype was also found to be an important predictor of survival. However it is unclear which subtype is associated with better or worse outcomes as one study reported an association between the IgG subtype and worse outcomes (Shah et al., 2012) whilst another study demonstrated that patients with non IgG subtype had worse outcomes (Sellner et., 2013).

All the evidence was in relation to survival outcomes and no evidence was identified for the outcomes treatment related morbidity and mortality, health related quality of life, adverse events, patient/carer/family acceptability and PROMs.

Table 97: GRADE profile: In which patients with relapsed or refractory myeloma is a second autologous stem cell transplant more effective than other therapy (ASCT2 versus alternative treatment in patients with a relapse-free survival > 18 months from ASCT1)?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	ASCT2	alternative treatment	Relative (95% CI)	Absolute	Quality
median OS											
1	observational studies	serious[1]	no serious inconsistency	no serious indirectness	no serious imprecision	none	63	43	-	Median OS was 1.7 years longer in patients that underwent salvage ASCT compared to patients that underwent other salvage treatments.	VERY LOW

1 published as letter: limited study details and not peer-reviewed (Alvares et al., 2006)

Table 98: GRADE profile: In which patients with relapsed or refractory myeloma is a second autologous stem cell transplant more effective than other therapy (ASCT2 versus salvage systematic chemotherapy in patients < 54 years at ASCT1)?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	ASCT2	salvage systematic chemotherapy	Relative (95% CI)	Absolute	Quality
median OS from relapse											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none	60	60	-	Median OS was 1.75 years longer in	LOW

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	ASCT2	salvage systematic chemotherapy	Relative (95% CI)	Absolute	Quality
										patients that underwent salvage ASCT compared to patients that underwent salvage chemotherapy.	

Table 99: GRADE profile: In which patients with relapsed or refractory myeloma is a second autologous stem cell transplant more effective than other therapy (ASCT2 versus salvage systematic chemotherapy in patients 55 - 65 years at ASCT1)?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	ASCT2	salvage systematic chemotherapy	Relative (95% CI)	Absolute	Quality
median OS from relapse											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	serious[1]	none	?	?	-	Median OS was 1.7 years longer in patients that underwent salvage ASCT compared to patients that underwent salvage chemotherapy.	VERY LOW

1 number of patients in subgroup unclear (maximum 46)

Table 100: GRADE profile: In which patients with relapsed or refractory myeloma is a second autologous stem cell transplant more effective than other therapy (ASCT2 versus salvage systematic chemotherapy in patients > 65 years at ASCT1)?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	ASCT2	salvage systematic chemotherapy	Relative (95% CI)	Absolute	Quality
median OS from relapse											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	serious1	none	?	?	-	Median OS was not significantly different in patients that underwent salvage ASCT and patients that underwent salvage chemotherapy.	VERY LOW

1 number of patients in subgroup unclear (maximum 46)

Table 101:GRADE profile: In which patients with relapsed or refractory myeloma is a second autologous stem cell transplant more effective than other therapy (ASCT2 versus salvage systematic chemotherapy in patients with a duration of response greater than 18 months post ASCT1)?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	ASCT2	salvage systematic chemotherapy	Relative (95% CI)	Absolute	Quality
median OS from relapse											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none	52	59	-	Median OS was 2.1 years longer in patients that	LOW

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	ASCT2	salvage systematic chemotherapy	Relative (95% CI)	Absolute	Quality
										underwent salvage ASCT compared to patients that underwent salvage chemotherapy.	

Table 102: GRADE profile: In which patients with relapsed or refractory myeloma is a second autologous stem cell transplant more effective than other therapy (ASCT2 versus salvage systematic chemotherapy in patients with achievement of at least a PR (CR/PR) following ASCT1)?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	ASCT2	salvage systematic chemotherapy	Relative (95% CI)	Absolute	Quality
median OS from relapse											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	no serious imprecision	none	91	91	-	Median OS was 2 years longer in patients that underwent salvage ASCT compared to patients that underwent salvage chemotherapy.	LOW

Table 103: GRADE profile: In which patients with relapsed or refractory myeloma is a second autologous stem cell transplant more effective than other therapy (ASCT2 versus salvage systematic chemotherapy in patients with poor responding disease to ASCT1 (no response, minimal disease or progressive disease))?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	ASCT2	salvage systematic chemotherapy	Relative (95% CI)	Absolute	Quality
median OS from relapse											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	serious[1]	none	15	15	-	Median OS was 1 year longer in patients that underwent salvage ASCT compared to patients that underwent salvage chemotherapy.	VERY LOW

1 small sample size

Table 104: GRADE profile: In which patients with relapsed or refractory myeloma is a second autologous stem cell transplant more effective than other therapy (ASCT2 versus salvage systematic chemotherapy in patients with a good prognosis (TTP >18 months after ASCT1 and ISS 1 or II))?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	ASCT2	salvage systematic chemotherapy	Relative (95% CI)	Absolute	Quality
median PFS											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	serious[1]	none	13	34	-	Median OS was no different in	VERY LOW

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	ASCT2	salvage systematic chemotherapy	Relative (95% CI)	Absolute	Quality
										patients that underwent salvage chemotherapy and patients that salvage ASCT.	
median OS											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	serious[1]	none	13	34	-	Median PFS was 23.7 months longer in patients that underwent salvage ASCT compared to patients that underwent salvage chemotherapy.	VERY LOW

1 *1 small number of patients in the intervention group (ASCT2)*

Table 105: GRADE profile: In which patients with relapsed or refractory myeloma is a second autologous stem cell transplant more effective than other therapy (ASCT2 versus salvage systematic chemotherapy in patients with a poor prognosis (TTP <18 months after ASCT1 and/or ISS III))?

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	ASCT2	salvage systematic chemotherapy	Relative (95% CI)	Absolute	Quality
median OS											

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	ASCT2	salvage systematic chemotherapy	Relative (95% CI)	Absolute	Quality
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	serious[1]	none	35	110	-	Median OS was 32.7 months longer in patients that underwent salvage ASCT compared to patients that underwent salvage chemotherapy.	VERY LOW
median PFS											
1	observational studies	no serious limitations	no serious inconsistency	no serious indirectness	serious[1]	none	35	110	-	Median PFS was 6.6 months longer in patients that underwent salvage ASCT compared to patients that underwent salvage chemotherapy.	VERY LOW

1 *1 small number of patients in the intervention group (ASCT2)*

2 **Table 106: GRADE profile: In which patients with relapsed or refractory myeloma is a second autologous stem cell transplant more**
3 **effective than other therapy (ASCT2 versus cyclophosphamide in patients with a first response to ASCT1 longer than 24**
4 **months)?**

Quality assessment	Summary of findings		
	No of patients	Effect	Quality

No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	ASCT2	Cyclophosph-amide	Relative (95% CI)	Absolute	
median time to progression											
1	randomised trials	no serious limitations	no serious inconsistency	serious[1]	no serious imprecision	none	64	64	-	Median TTP was 13 months longer in patients that underwent salvage ASCT compared to patients that underwent cyclophosamide.	MODERATE

1 *1 choice of cyclophosphamide might be questioned in current treatment landscape.*

2 **Table 107: GRADE profile: In which patients with relapsed or refractory myeloma is a second autologous stem cell transplant more**
3 **effective than other therapy (ASCT2 versus cyclophosphamide in patients with a first response to ASCT1 of 24 months or**
4 **less)?**

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	ASCT2	Cyclophosphamide	Relative (95% CI)	Absolute	Quality
median time to progression											
1	randomised trials	no serious limitations	no serious inconsistency	serious[1]	no serious imprecision	none	25	21	-	Median TTP was 4 months longer in patients that underwent salvage ASCT compared to patients that underwent cyclophosami	MODERATE

Quality assessment							Summary of findings				
							No of patients		Effect		
No of studies	Design	Limitations	Inconsistency	Indirectness	Imprecision	Other considerations	ASCT2	Cyclophosphamide	Relative (95% CI)	Absolute	Quality
										de.	

1 *1 choice of cyclophosphamide might be questioned in current treatment landscape.*

1 Cost effectiveness evidence

2 A literature review of published cost-effectiveness analyses did not identify any relevant
3 papers for this topic. Whilst there were potential cost implications of making
4 recommendations in this area, other questions in the guideline were agreed as higher
5 priorities for economic evaluation. Consequently no further economic modelling was
6 undertaken for this question.

7

Recommendations	**Offer a second autologous stem cell transplant to people with relapsed myeloma are suitable and who have:** • **completed re-induction therapy and** • **had a response duration of more than 24 months after their first autologous stem cell transplant.** **Consider a second autologous stem cell transplant for people with relapsed myeloma who are suitable and who have:** • **completed reinduction therapy and** • **had a response duration of between 12 and 24 months after their first autologous stem cell transplant.** **When assessing whether people with relapsed myeloma are suitable for a second autologous stem cell transplant, take into account:** • **response to the first autologous stem cell transplant** • **International Staging System (ISS) stage** • **number of prior treatments** • **age, frailty and comorbidities** • **adverse fluorescence in-situ hybridisation (FISH) results.**
Relative value placed on the outcomes considered	The Guideline Committee considered the outcomes of overall survival, progression-free survival, health-related quality of life, treatment related mortality and morbidity, patient/carer/family acceptability, adverse events and patient reported outcome measures to be the most relevant in determining whether second autologous stem cell transplant was effective in specific subgroups of patients with relapsed/refractory myeloma. Of these, evidence was identified for overall survival and progression-free survival. Evidence was also reported for time to progression in one study. When drafting the recommendations the Guideline Committee considered overall survival and progression-free survival to be the most important as these are most clinically meaningful.
Quality of the evidence	The quality of the evidence was assessed using GRADE methodology and appropriate NICE Checklists. There was moderate quality evidence for time to progression and very low to moderate quality evidence for overall survival and progression free survival.
Trade off between clinical benefits and harms	The Guideline Committee noted that there was moderate quality data from 1 study of improved outcomes when patients had a response duration longer than 24 months following their first autologous stem cell transplant. They therefore recommended a second autologous stem cell transplant be offered to these patients. However the Guideline Committee discussed that it is possible that there may be patients who meet the criteria of having a response duration greater than 24 months but who may still not be fit for transplant. They therefore included ‘if suitable’ in the recommendation.

	A consider recommendation was made for patients with a response duration of 12-24 months as there was weaker quality evidence from a number of different studies demonstrating the effectiveness of second transplant following response durations in this timeframe. The Guideline Committee also made a recommendation regarding which factors needed to be taken into account when deciding on whether to proceed with a second transplant. The factors were based on evidence for their prognostic value. The Guideline Committee's clinical experience was used to exclude those factors that were only reported by individual studies or where the effect that was reported between different studies was inconsistent. The Guideline Committee also used their clinical experience in assessing the short term toxicities of autologous stem cell transplant. The Guideline Committee concluded that the benefits would be improved progression-free survival and potentially improved overall survival and long term quality of life. It was discussed that there could be potential treatment related mortality, but this was thought to only be a small percentage. Also it was suggested that there would be an increase in treatment related morbidity (both short and long term) as well as potentially worse short-term quality of life. The Guideline Committee balanced low mortality and low long-term morbidity against improvements in progression-free survival. It was discussed that short term acute toxicity will happen in most patients but patients will have already experienced this in their first transplant and are likely to regard it as acceptable and tolerable to achieve an improvement in long term quantity and quality of life.
Trade-off between net health benefits and resource use	The Guideline Committee noted that no relevant published economic evaluations had been identified and no additional economic analysis had been undertaken in this area. The Guideline Committee thought the recommendations would result in an increase in the number of patients transplanted and so an increase in costs associated with this as well as increased demand on transplant units and in-patient bed capacity. It was thought that there would be increased costs of either storing second harvests or re-harvesting. The Guideline Committee agreed that there may be savings as a result of delayed use of drugs for relapse (which are expensive). Furthermore delaying progression would result in saving costs of treating disease-related morbidity. The Guideline Committee were unclear of the net effect. But concluded that it was likely to be cost effective on cost/QALY grounds as there is a clear benefit in terms of progression-free survival.
Other considerations	The Guideline Committee felt that there may be a change in practice as a result of the recommendations with increased number of transplants meaning increased demand on capacity of transplant units, in-patient beds, stem cell harvesting and storage. The Guideline Committee discussed that capacity can potentially be addressed by introducing ambulatory care and shared-care pathways.

11.3 Subsequent therapy

For guidance on the use of lenalidomide in people who have received at least 1 prior therapy, see Lenalidomide for the treatment of multiple myeloma in people who have received at least one prior therapy (NICE technology appraisal guidance 171).

For guidance on the use of pomalidomide in people who have relapsed and refractory disease, see Pomalidomide for relapsed and refractory multiple myeloma previously treated with lenalidomide and bortezomib (NICE technology appraisal guidance 338).

References

Alvares CL, Davies FE, Horton C, et al. (2006) The role of second autografts in the management of myeloma at first relapse. Haematologica. 91(1), 141-142.

Auner HW, Szydlo R, Rone A, et al. (2013) Salvage autologous stem cell transplantation for multiple myeloma relapsing or progressing after up-front autologous transplantation. Leuk Lymphoma 54(10), 2200-2204.

Chow AW, Lee CH, Hiwase DK, et al. (2013) Relapsed multiple myeloma: who benefits from salvage autografts? Internal Medicine Journal, 43: 156-161.

Cook G, Liakopoulou E, Pearce R, et al. (2011) Factors influencing the outcome of a second autologous stem cell transplant (ASCT) in relapsed multiple myeloma: a study from the British Society of Blood and Marrow Transplantation Registry. Biology of Blood & Marrow Transplantation, 17: 1638-1645.

Cook G, Williams C, Brown JM, et al. (2014) High-dose chemotherapy plus autologous stem-cell transplantation as consolidation therapy in patients with relapsed multiple myeloma after previous autologous stem-cell transplantation (NCRI Myeloma X Relapse [Intensive trial]): A randomised, open-label, phase 3 trial. Lancet oncology, 15: 874-885.

Fenk R, Liese V, Neubauer F, et al. (2011) Predictive factors for successful salvage high-dose therapy in patients with multiple myeloma relapsing after autologous blood stem cell transplantation. Leukemia & Lymphoma, 52: 1455-1462.

Gonsalves WI, Gertz MA, Lacy MQ, et al. (2013) Second auto-SCT for treatment of relapsed multiple myeloma. Bone Marrow Transplantation, 48: 568-573.

Grovdal M, Nahi H, Gahrton G, et al. (2015). Autologous stem cell transplantation versus novel drugs or conventional chemotherapy for patients with relapsed multiple myeloma after previous ASCT. Bone Marrow Transplantation, 50, 808-812.

Jimenez-Zepeda VH, Mikhael J, Winter A, et al. (2012) Second autologous stem cell transplantation as salvage therapy for multiple myeloma: impact on progression-free and overall survival. Biology of Blood & Marrow Transplantation, 18: 773-779.

Lemieux E, Hulin C, Caillot D, et al. (2013) Autologous stem cell transplantation: an effective salvage therapy in multiple myeloma. Biology of Blood & Marrow Transplantation, 19: 445-449.

Michaelis LC, Saad A, Zhong X, et al. (2013) Salvage second hematopoietic cell transplantation in myeloma. Biology of Blood & Marrow Transplantation, 19: 760-766.

Olin RL, Vogl DT, Porter DL, et al. (2009) Second auto-SCT is safe and effective salvage therapy for relapsed multiple myeloma. Bone Marrow Transplantation, 43: 417-422.

Sellner L, Heiss C, Benner A, et al. (2013) Autologous retransplantation for patients with recurrent multiple myeloma: a single-center experience with 200 patients. Cancer, 119: 2438-2446.

Shah N, Ahmed F, Bashir Q, et al. (2012) Durable remission with salvage second autotransplants in patients with multiple myeloma. Cancer, 118: 3549-3555.

Wirk B, Byrne M, Dai Y, et al. (2013) Outcomes of salvage autologous versus allogeneic hematopoietic cell transplantation for relapsed multiple myeloma after initial autologous hematopoietic cell transplantation. Journal of Clinical Medicine Research, 5: 174-184.

Yhim HY, Kim K, Kim JS, et al. (2013) Matched-pair analysis to compare the outcomes of a second salvage auto-SCT to systemic chemotherapy alone in patients with multiple myeloma who relapsed after front-line auto-SCT. Bone Marrow Transplant 48(3):425-32.