Project plan:

Treatments for relapsing, refractory multiple myeloma: A health technology assessment

Project number	ID2019_072
Plan prepared :	6 January 2021

Short description and summary

Multiple myeloma (MM), the second most common type of blood cancer, reduces the body's ability to fight infection and can lead to anemia, bone damage, nerve damage, kidney damage, etc. Approximately 450 Norwegians are diagnosed with myeloma annually. Because there is currently no cure for myeloma, it is important to evaluate the clinical and cost-effectiveness of the increasing number of treatments available for patients who experience relapsed or refractory multiple myeloma (RRMM).

Short title

RR Multiple Myeloma

Norsk tittel

Behandlinger av tilbakefall eller refraktært myelomatose: En fullstendig metodevurdering

Norsk sammendrag

Myelomatose (multippel myelom, MM), som er den nest vanligste formen for blodkreft, reduserer kroppens evne til å bekjempe infeksjon og kan føre til anemi, beinskade, nerveskader, nyresvikt, osv. Omtrent 450 nordmenn får myelomatose årlig. Fordi det ikke finnes en kur mot myelomatose er det viktig å evaluere klinisk effekt og kostnadseffektivitet av det økende antallet behandlingsalternativ for pasienter som opplever tilbakefall eller refraktært myelomatose (RRMM).

Project category and commissioner	
Product (program area)	Health Technology Assessment
Thematic areas	
Commissioner:	Commissioning Forum (Bestillerforum RHF), consisting of four medical directors representing each of the Regional Health Authorities, and two

	delegates from the Norwegian Directorate of Health. The Forum's mandate is to prioritize HTA topics based on submitted proposals and horizon scanning reports.
Project management and participants	
Project manager	Arna Desser (AD)
Responsible for the project	Atle Fretheim (AF)
Internal project participants	Ingrid Kristine Ohm (IKO) Liv Giske (LG) Elisabet Hafstad (EH) Gunn Eva Næss (GEN) Christopher Rose (CR) Ulrikke Lund (UL) Anna Espeland (AE) [replacing UL in 01-2021]
External project participants	Einar Haukås, Stavanger University Hospital Olav Ljøsne, Blodkreftforeningen Mats Irgen Olsen, University Hospital Northern Norway-Tromsø Fredrik Schjesvold, Oslo University Hospital
Plan for replacement by project participants' absence	The person responsible for the project will replace the project participants when needed
Internal reviewer	Kjetil Bruberg
External reviewers	Øyvind Hjertner, St. Olav's University Hospital One health economist (to be determined)

Mandate

On March 30, 2020, the commissioning forum in the New Methods system ("Bestillerforum RHF") requested that the Norwegian Institute of Public Health perform a health technology assessment of treatments for patients with relapsed, refractory multiple myeloma (RRMM) focusing on medications or combinations of medications relevant for use in Norway. This commission is a revision of a commission issued on May 27, 2019 for of a health technology assessment of multiple myeloma. The current commision reflects a more precise specification of the relevant patient population following an evidence mapping [1] performed by the Norwegian Institute of Public Health at the request of "Bestillerforum".

Objective

To determine the clinical effectiveness, safety, and cost-effectiveness of treatments for patients with relapsed, refractory multiple myeloma in a Norwegian context.

Background

Myeloma (multiple myeloma, MM), the second most common type of blood cancer, affects plasma cells in bone marrow, most often in the bones of the spine, skull, pelvis, rib cage, shoulder, and hips. Plasma cells are a type of white blood cell that produce immunoglobulins, which are complex proteins known as antibodies. Myeloma cells (malignant plasma cells) produce abnormal M-proteins (monoclonal proteins) rather than normal, functioning antibodies. In addition to reducing the body's ability to fight infection, M-proteins can cause other serious problems, e.g. kidney damage, anemia, bone damage, nerve damage, etc. Myeloma is often called 'multiple myeloma' to indicate that it normally occurs simultaneously in multiple sites in the body. Age and previous monoclonal gammopati of undetermined significance (MGUS) are the most important risk factors for the disease. [2] Approximately 450 new cases of myeloma are diagnosed annually in Norway. The median age at diagnosis is approximately 70 years, and incidence is rare among individuals under age 30. [3]

Myeloma is often first suspected when patients experience skeletal pain, anemia, frequent respiratory or other infections, poor kidney function, or elevated calcium levels in the blood (hypercalcemia). Diagnosing multiple myeloma that will require treatment[1] involves a bone marrow or tumor biopsy to confirm the presence of malignant plasma cells and tests to confirm incidence of one or more CRAB criteria: (C) hypercalsemia, (R) kidney (renal) damage, (A) anemia, and (B) bone damage involving bone lesions or low bone density. [4]

Because there is currently no cure for myeloma, the goal of both initial treatment and subsequent treatments is to achieve as strong a response as possible without unacceptable side effects, and as long a period of progression-free survival as possible. Norwegain guidelines recommend the most effective available treatment, given at the dose recommended in

[1] In addition to myeloma requiring treatment, there are two categories of myeloma ('smoldering multiple myeloma' and MGUS) in which no CRAB criteria are present, but where there is evidence of monoclonal proteins in bone marrow. MGUS (monoclonal gammopati of undetermined significance) is a non-cancerous condition that is considered a precursor of myeloma. There is a 1% annual risk that MGUS will progress to multiple myeloma. Smoldering multiple myeloma (SMM) is diagnosed when levels of monoclonal protein in bone marrow exceed the MGUS cut-off level but is below the level at which multiple myeloma is diagnosed. SMM patients have an annual 10% risk of developing active MM within the first five years, and an annual risk that declines to 3% over the next five years, and to 1%-2% over the 10-year period after that. [2] In Norway, regular follow-up is recommended for MGUS and SMM patients.

supporting clinical studies, based on a patient's age, overall health, and response to earlier rounds of treatment. Dosages can be adjusted downwards in response to side effects. Patients generally will require constant treatment over their remaing lifetime, possibly with short periods without treatment. [4]

The initial choice of treatment for multiple myeloma is based on patient age. Individuals under age 70 are usually offered high-dose chemotherapy with an autologous stem cell transplant (ASCT, transplant with patient's own stem cells). The treatment consists of five phases: 1) induction, i.e. treatment with myeloma directed drugs intended to achieve maximum response without negatively effecting stem cell harvesting, 2) stem cell harvesting, 3) ASCT, 4) consolidation, generally a repeat of the induction treatment, and 5) maintenance. Treatment involving stem cell transplantation may also be appropriate for very healthy and motivated individuals over age 70. [4]

Patients over age 70 or individuals who cannot tolerate or do not wish to undergo the ASCT process are treated with chemotherapy and combinations of medications that inhibit plasma cell division in bone marrow. Pasients under active treatment can often be treated at home with regular outpatient follow-up.

Both groups of newly diagnosed patients can experience disease remission, symptom relief, and increased survival. Patients experiencing side effects from treatments receive medications to relieve symptoms and pain. Radiation therapy can also be used either therapeutically or to control pain.

A large and growing number of potential treatment options, involving either a single drug or combinations of multiple drugs, can result in potentially complex decisions about treatment paths when patients experience relapse following a period of remission, or become resistant to the current treatment (refractory disease). As many of the new treatment options are aimed at patients who have suffered a relapse or are refractory to a particular treatment, it is important to have a good understanding of the clinical and cost-effectiveness of all available treatments.

Methods

Based on our objectives, we will produce a health technology assessment in accordance with the National Institute of Public Health's handbook, "Slik oppsummerer vi forskning" [5].

Search strategy

We will use the systematic reviews included in our published mapping review of March 2020 [1] as the basis to identify randomized, controlled trials (RCT) that are relevant for our health technology assessment. The search for systematic reviews in the mapping review was initially performed in February 2020 in the Epistemonikos database and consisted only of the search term "myeloma".

Additionally, we will search for eligible RCTs that: 1) are ongoing, and 2) have been published after the last literature search as described in the included systematic reviews. We will search the following databases/registries: Cochrane Central Register of Controlled Trials (Wiley); MEDLINE (Ovid); Embase (Ovid); Web of Science (Clarivate Analytics); ClinicalTrials.gov; International Clinical Trials Registry Platform; EU Clinical Trials Register. We will either update the last search as described in the included systematic reviews or search the databases from inception.

A librarian will define and process the search terms in collaboration with the researchers of the team and put terms together in a search strategy. The search strategies will use a combination of controlled terms, i.e., Medical Subject Headings (MeSH), Emtree terms, as well as free-text terms with various synonyms that reflect the concepts of the population, "multiple myeloma", and the generic names of relevant pharmaceuticals – see inclusion criteria. Where appropriate, we will use a filter for study design (RCT). The librarian will then adapt the strategies to each database, run and document the searches, and prepare the retrieved records for screening. In addition, the researchers will manually search the bibliographies of included RCTs for any additional relevant trials not captured by the database searches. Following the completion of the search we will contact relevant firms to confirm that there are no additional publications that meet our inclusion criteria.

To facilitate the use of GRADE and standardized reporting statements in assessing the quality of evidence, we will attempt to identify literature, through a literature search or based on expert advice, to help establish minimally important relative or "absolute" treatment effects.

Eligibility criteria

Our framework for searching for and selecting relevant literature for our health technology assessment is outlined in the PICO (Population, Intervention, Comparator, Outcome). The list of relevant drugs for the intervention is based on 1) a proposal for a commission sent to Nye Metoder from The Norwegian Medicine Agency, Sykehusinnkjøp RHF and The Norwegian Institute of Public Health (13th May 2019) [6], and 2) the new national guidelines for treating relapsed/refractory multiple myeloma (RRMM) [4].

Inclusion and exclusion criteria

PICOS	Inclusion	Exclusion
Population	Relapsed/refractory multiple myeloma (RRMM), i.e., individuals over 18 years, diagnosed with multiple myeloma 1) who are refractory to treatment, or 2) who have experienced one or more relapses	Newly diagnosed multiple myeloma, smouldering myeloma, monoclonal gammopathy of undetermined significance (MGUS),
Intervention	Treatment with any of the following drugs, alone or in combination with each other and/or with glucocorticosteroids (e.g. dexamethasone, or prednisone): • Bortezomib (Velcade) • Carfilzomib (Kyprolis) • Daratumumab (Darzalex) • Elotuzumab (Empliciti) • Ixazomib (Ninlaro) • Isatuximab (Sarclisa) • Lenalidomide (Revlimid) • Panobinostat (Farydak) • Pomalidomide (Imnovid)	Any other agents used in the treatment of multiple myeloma, doses or administration forms of the listed drugs that are off label,
Comparison	• All intervention-drugs alone or in combination with each other, or in combination with other drugs • Placebo • Standard treatment • Glucocorticosteroids, e.g. dexamethasone, and prednisone	
Outcome	Primary • Overall survival • Quality of life • Serious adverse events Secondary • Progression-free survival • Adverse events • Withdrawal from study due to adverse events	Reports on cellular and molecular mechanisms
Study design	• Systematic reviews based on RCTs • RCTs, phase 2 or 3	Systematic reviews based on non-RCTs, studies with other design, e.g. non-controlled trials, non-RCTs using registry data, animal studies, in vitro studies, etc.

RRMM: relapsing/refractory multiple myeloma, RCT: randomised, controlled trials

Selection of studies

We will select studies found in the literature search in a two-step selection strategy:

1. Screening: two researchers will independently screen titles and abstracts (where available) using Rayyan QCRI software [7], to include or exclude articles based on their relevance to our research question. When in doubt, full-text version will be retrieved.
2. Full-text assessment: two researchers will read the full-text articles to assess which will be included in our HTA.

Both steps will adhere to the eligibility criteria listed above. Disagreements in either of the two steps will be resolved through discussion, or by consultation with a third researcher or other members of the project team.

Risk of bias

For RCTs identified through the systematic reviews from our mapping review [1], we will refer to the risk of bias assessment made by the authors of the systematic reviews. For RCTs that we include based on our own literature search, we will assess quality by using the Cochrane Risk of Bias Tool [8]

Two researchers will perform the assessment independently, and any potential differences will be resolved through discussion between the researchers, or by consultation with a third researcher or other members of the project team. We will not perform our own risk of bias assessment of RCTs that already have been assessed by a systematic review, given that we find their assessment reasonable.

Data extraction

One researcher will extract relevant data from full-text articles to Covidence [9]. This will then be verified by a second researcher. Any potential disagreements will be resolved through discussion, or by consultation with a third researcher or other members of the project team. If necessary (e.g. if data are unintelligible, etc.), we will contact the authors for them to provide us with sufficient information to use in our HTA. Data will also be checked by a statistician prior to analysis. If it is necessary for the statistician to manually convert the extracted data from one format to another, the converted data will also be checked by another researcher (e.g., to protect against transcription errors).

For systematic reviews, we will extract information regarding the following:

- Publication details (e.g., authors, journal, publication year, etc.)
- Literature search (e.g., date, databases, search terms, etc.)
- Selection criteria
- Included RCTs

For RCTs, we will extract the following information:

About	Information to be extracted
The study	Authors, publication year, study design, country, clinical identification number, eligibility criteria, follow-up time, funding source (industry or non-industry)
The participants	For each trial arm and each outcome: numbers of participants randomized; numbers of participants included in the analysis; average age; percentage of participants who were female; percentage of participants who were caucasian; diagnosis; disease severity at baseline; percentage of participants who had received previous treatment (including stem cell treatment); average number of relapses.
The intervention and comparators	For each trial arm: name of intervention or comparator (including combinations); posology (incl. dose level, frequency, duration, and route of administration)
The outcome	For each pairwise comparison and each outcome: name of relative treatment effect estimate (e.g., HR, RR, OR); point estimate; name of measure of precision (e.g., 95% CI, SE, SD); precision (e.g., limits of the 95% CI). (See PICO and "Measures of relative treatment effect.")
The analysis	For each pairwise comparison and each outcome: analysis method (e.g., Cox regression, GLM); for cross-over and cluster studies, whether a unit of analysis error was made.

If any of the included studies report results for the same participants (e.g., extension studies), we will only extract data for the study or arms with longest follow-up, to avoid "double counting."

If any of the included studies use a cross-over design and have unit of analysis errors (i.e., failure to model treatment-sequence) or have not accounted for possible carry-over effects, we will extract data for the first period only.

We will follow the intention-to-treat (ITT) principle and extract relative treatment effect estimates corresponding to the trial arms to which patients were randomized. We will include "modified intention-to-treat analyses" if we judge that the method used would provide similar estimates to an ITT analysis. We will exclude "per protocol" results.

Measures of relative treatment effect

The following table outlines, for each outcome, the measures of relative treatment effect we anticipate the included studies will report, along with the scale on which we plan to perform meta-analysis. Where possible, we will extract and meta-analyze published relative treatment effect estimates directly. Studies may report results in other ways (e.g., as numbers of events or arm-wise means). In such cases we will extract the available data and, where possible, use standard methods to impute relative treatment effect estimates that can be used in meta-analysis.

Outcome	Anticipated treatment effect estimate(s)	Scale for meta-analysis
Primary outcomes		
Overall survival	Hazard ratio (HR)	log HR
Quality of life	Mean difference (MD); arm-wise means	MD
Serious adverse events	Risk ratio (RR); odds ratio (OR)†	log RR
Secondary outcomes		
Progression-free survival	Hazard ratio (HR)	log HR
Adverse events	Risk ratio (RR); odds ratio (OR)†	log RR
Withdrawal from study due to adverse events	Risk ratio (RR); odds ratio (OR)†	log RR

† Where odds ratios are reported, we will impute corresponding risk ratios and report the method used.

We will use alternative scales of measurement if necessary. For example, if a continuous outcome has been reported in the included RCTs using different instruments or scales we may perform meta-analysis on a standardized mean difference (SMD) scale.

Treatment definition

We will define a treatment (i.e., intervention or comparator) to be a unique combination of a nonproprietary active drug name (e.g., bortezomib) and its posology (dose level, frequency, and route of administration), or, for combinations of active drugs, a unique combination of such. We anticipate that patients participating in studies will normally be given a treatment (as defined above) with other "nonactive" interventions (e.g., glucocorticosteroids) and that these may differ between studies of the same active drug. We will not consider such unique combinations to constitute distinct treatments (i.e., we will focus on the drugs named in the inclusion and

exclusion table above). We anticipate that distinct nonactive interventions may be associated with the relative differences in safety and efficacy, which we will account for using random effects models (see Statistical analysis section, below).

Missing data and unit of analysis errors

We will not impute outcome data for studies with missing data. If fewer than 5% of the randomized participants were not included in a given, published analysis, and those participants can plausibly be assumed to be missing completely at random (MCAR), we will assume that the corresponding treatment effect estimate is unlikely to be biased [10]. If more than 5% of the randomized participants were not included in a published analysis, and that study is a source of inconsistency in the network meta-analysis (see below), we may exclude the study from the meta-analysis. We will report and justify any exclusion of studies. We will account for the possible effect of missing data on meta-analysis results via our risk of bias and certainty of evidence (GRADE) assessments.

If any of the included studies use a cluster design and have unit of analysis errors (e.g., they have not modelled possible intra-cluster correlation), we will impute cluster-adjusted standard errors (e.g., using an assumed intra-cluster correlation coefficient taken from relevant literature).

Statistical analyses

We will only include studies in statistical analyses if they have directly compared at least two of the included treatments or comparators.

If necessary, we will adjust standard errors of estimates from studies with more than two arms [11]. We will judge the possibility of publication bias for each primary outcome using funnel plots: we will plot relative treatment effect against standard error, generating one funnel plot for each direct pairwise comparison. Because formal statistical tests of funnel plot asymmetry generally lack power, we will not use such tests to exclude the possibility of publication bias [12]. Instead, we will briefly comment on any apparent asymmetries and their possible implications in the text of the report.

For each outcome we will conduct a pairwise meta-analysis for each comparison supported by direct evidence. We will then perform frequentist inverse variance-weighted network meta-analysis (NMA) for each outcome. We will use a random effects model for all meta-analyses because we anticipate heterogeneity. For example, two studies comparing the same active drug to placebo may use different combinations of “nonactive” interventions (e.g., glucocorticosteroids). We will use contrast-wise NMA except in the case that the included studies do not form a single network of evidence and there are interventions of interest that are disconnected from the “main” network. In this case, arm-wise network meta-analysis will be performed.

For each outcome, we will consider the appropriateness of the transitivity assumption that underpins the approach by plotting distributions of the potential treatment effect modifiers (the variables defined above in “Data extraction”) among the included studies. If we judge that there are important differences in the distributions of these variables across the studies such that the transitivity assumption is threatened, we will either adjust for the variables using network meta-regression (NMR) or omit the studies that threaten transitivity from the analysis. We will report

and justify the omission of any studies. If there are too few studies to support NMA, we will perform pairwise meta-analyses instead.

We will preferentially use frequentist NMA methods described by Rücker [13] and Schwarzer et al. [14] using the netmeta (version 1.2-1 or later) or metafor (version 2.4-0 or later) R packages. Briefly, these methods pose NMA as a generalized linear model (GLM). However, these methods may yield biased results if the normal approximation assumption is violated (particularly if there are a substantial number of studies where few or no events were observed). [15]

If there are primary outcomes for which a substantial number of studies have few events, we will use a Bayesian NMA approach, such as the GLM-based framework described by Dias et al. [16], for example, using the gemtc R package (version 0.8-6 or later) or by coding models in Stan [17]. Any Bayesian analyses will use binomial likelihoods for dichotomous outcomes and normal likelihoods for continuous outcomes; we will use vague priors on all parameters [18]. We will use four MCMC chains with different initializations and conservative numbers of burn-in and posterior samples. We will evaluate the quality of the samples drawn to approximate posterior distributions quantitatively using the potential scale reduction factor, and qualitatively by inspecting diagnostic plots. We will briefly summarize our conclusions about the quality of the samples. Due to the potentially large number of summary values and plots that may result, we will present these in an electronic appendix or omit them, unless their inclusion is deemed particularly important. If we judge posterior samples to be untrustworthy (e.g., there is poor mixing of the chains) and the problem cannot be solved straightforwardly by changing readily available options, the Bayesian NMA will be abandoned.

We will check for inconsistency between direct, indirect, and network evidence. We will consider estimates to be inconsistent if their confidence or credible intervals do not overlap. For frequentist NMAs, we will present net heat plots to aid the assessment of inconsistency and will assess network and within- and between-design homogeneity and consistency via decompositions of Cochrane's Q statistic. For frequentist and Bayesian NMAs, we will also assess inconsistency by "node-splitting" [19].

We will re-express relative treatment effect estimates as "anticipated" (i.e., "absolute") values [20, 21] to aid interpretation. For example, we may re-express a relative risk as the number of patients who would be expected to experience the event if they were given standard care along with the corresponding number of patients expected to experience the event if they were treated with a given intervention. Uncertainty on the "corresponding" values will be expressed as 95% CIs or CrIs.

We will use P-scores [22] or SUCRA values [23] to quantify the extent of evidence that each treatment is superior to all other treatments and will use these values to rank the treatments. We will interpret rankings cautiously, considering the quality of evidence.

Where possible, we will use items from the NMA extension to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses) checklist [18] to ensure analyses are performed and reported according to accepted practice. We will report any limitations of the analyses and justify any substantial deviations from this statistical analysis plan.

If we judge that the planned NMA or NMR analyses are inappropriate because it is difficult to define treatments as described above, we will consider alternative methods such a component network meta-analysis [24], in which each study arm is modelled as a combination of treatments that act in combination.

Outcomes and results which cannot be combined in meta-analysis will be presented narratively.

Presentation of results

The table below summarizes the display elements we will present:

Report section	Display elements
Executive summary	Radar plots
Main body of report	Network geometry plots; forest plots; matrix plots; summary of findings tables
Electronic appendix	Summaries of included studies, study arms, participants, and follow-up; distributions of potential treatment effect modifiers; network geometry plots; funnel plots; forest plots of meta-analytical estimates of relative and "absolute" treatment effect; net heat plots; partial SoF tables; matrix plots of relative treatment effects; tables and forest plots comparing direct, indirect, and network evidence; tables of treatment rankings; radar plots; tables to aid in GRADEing; model diagnostic tables and figures

We will present radar plots in the executive summary that summarize the estimated superiority of each treatment with respect to the primary and secondary outcomes. We will present a radar plot for each intervention, where each axis shows the P-score for one of the six outcomes.

In the main body of the report, we will present the following for each outcome: a figure showing the geometry of the network and the direct evidence available (i.e., all pairwise comparisons), and its influence (weight) in the NMA; a forest plot showing anticipated (i.e., "absolute") treatment effect estimates for each intervention; a matrix plot showing the network meta-analytical relative treatment effect estimate for each pair of interventions; and a summary of findings (SoF) table. The SoF table will have columns for: intervention name; meta-analytical relative treatment effect estimates; anticipated effects (with a reference treatment, with the intervention, and the difference); certainty of evidence (GRADE); and treatment rank with P-score or SUCRA value. Precisions will be presented as 95% confidence intervals (CI; for frequentist analyses) or 95% credible intervals (CrI; for Bayesian analyses). For Bayesian analyses, we will present point estimates as posterior means.

Because network meta-analyses are complex, the complete results will be presented in an electronic appendix (see table above).

Assessment of quality of evidence

We will assess the quality of evidence for each selected outcome using the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) system [25]. In brief, the

GRADE system evaluates the certainty of evidence through assessment of several criteria, either downgrading or upgrading the certainty of evidence [25]:

- For downgrading, the following is considered: a) study limitations (risk of bias), b) inconsistency of results, c) indirectness of evidence, d) imprecision, and e) publication bias.
- For upgrading, the following is considered: a) large magnitude of effect, b) dose-response gradient, and c) all plausible confounders and bias that would reduce a demonstrated effect or increase the effect if no effect was observed.

The certainty of evidence is classified as follows:

Grade	Definition
High certainty ⨁⨁⨁⨁	Further research is very unlikely to change our confidence in the estimate of effect
Moderate certainty ⨁⨁⨁◯	Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate
Low certainty ⨁⨁◯◯	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 certainty ⨁◯◯◯	Any estimate of effect is very uncertain

Two review authors will independently assess the certainty of the evidence for the primary outcomes; if the total number of outcomes (i.e., network meta-analyses) is fewer than five, we will assess the certainty of the evidence for all outcomes. If the total number of comparisons in a network meta-analysis is 10 or fewer, we will assess the certainty of the evidence for all comparisons, otherwise we will assess the certainty of the evidence for comparisons with a reference treatment only.

To facilitate assessment of the certainty of evidence using the GRADE methodology for network meta-analysis [26], which requires identifying a "dominant [indirect] path" for each comparison, we will define such a path as the one with minimal total sampling variance. We will report dominant paths in an appendix.

Standardised statements for the reporting of effects

We plan to present textual descriptions of effect estimates using standardised statements for the reporting of effects in the summary of findings tables [27]. Given a judgement about whether an effect estimate corresponds to an important, less important, or no benefit or harm, and a GRADE assessment of the certainty of evidence, a standardized statement can be chosen and adapted to communicate the magnitude, direction, and the certainty of evidence supporting an effect estimate in "plain language".

GRADE	Important benefit/harm	Less important benefit/harm	No important benefit/harm
High	[Intervention] improves/reduces [outcome] (high certainty evidence)	[Intervention] slightly improves/reduces [outcome] (high certainty evidence)	[Intervention] makes little or no difference to [outcome] (high certainty evidence) Or [Intervention] does not have an important effect on [outcome] Or

			[Intervention] has little or no effect on [outcome]
Moderate	[Intervention] probably improves/reduces [outcome] (moderate certainty evidence)	[Intervention] probably slightly improves/reduces [outcome] (moderate certainty evidence) Or [Intervention] probably leads to slightly better/worse/less/more [outcome] (moderate certainty evidence)	Intervention] probably makes little or no difference to [outcome] (moderate certainty evidence)
Low	[Intervention] may improve/reduce [outcome] (low certainty evidence)	[Intervention] may slightly improve/reduce [outcome] (low certainty evidence	[Intervention] may make little or no difference to [outcome] (low certainty evidence
Very low	We do not know if/It is uncertain whether [intervention] improves/reduces [outcome] because the certainty of this evidence is very low		

Health economic evaluation

We will perform an assessment of the cost-effectiveness and budgetary consequences of those treatments for relapsed, refractory multiple myeloma, considered by clinical experts to be relevant for use in Norway. We will perform a cost-utility analysis in which clinical outcomes are expressed in terms of gains in both quality-adjusted life years (QALYs) and life-years. We will calculate costs in Norwegian kroner (NOK), based on an expanded health sector perspective and Norwegian treatment guidelines. Model results will be expressed as incremental cost effectiveness ratios (ICERs), where the numerator is the difference in costs between an intervention and the comparator, and the denominator is the difference in effect between an intervention and the comparator. The analysis will be based on a probabilistic model in order to capture the effect of uncertainty in important clinical and cost parameters on predicted results.

The exact form of the model will be contingent upon the results of the network meta-analysis. Although it is most likely that we will perform a cost-utility analysis based on a Markov model approach or a partitioned survival analysis approach, it might be possible to implement a microsimulation model if sufficiently detailed clinical effect data are available. Our choice of model will comply with the NICE Technical support document concerning the use of partitioned survival analysis for health economic modeling. [28] We will use Norwegian epidemiological data, when available, or best transferable European data, in the absence of Norwegian data. All cost data will reflect treatment costs in Norway and will be collected from Norwegain sources. We will perform a separate search for the utility weights needed in the model to calculate quality adjusted life-years. We will collect utility weights to reflect quality of life at various stages of the disease, if such information exists; and to capture the effect on quality of life of serious, treatment-related, adverse events. Although our first preference would be to rely on utility weights gathered from Norwegian studies that use EQ5D methodology, we will use the best available information if our preferred data do not exist.

In addition to the cost-effectiveness analysis, we will conduct a budget impact analysis to estimate costs to the health care sector over the next five years of the included treatment options.

Model results will provide decision-makers with information that can inform consideration of treatments based on the three principle priority criteria: (1) health benefit, (2) resource use, and (3) severity.

Patient perspectives

The patient representative for the project has participated with the clinical experts in our initial meeting to discuss the PICO for the project and will provide information about patient experiences to inform the background and discussion chapters of the report. Understanding patient perspectives can provide a broader context for patients' health related quality of life during treatment. We will rely on discussions with the patient representative in conjunction with a form adapted from EUnetHTA for gathering information on patient perspectives, as well as information from a recent survey of members of Blodkreftforeningen, the Norwegian association for hematologic cancers.

Review process for the project plan and final report

We will follow the established review processes in the the Norwegian Institute of Public Health the for approval of project plans and final reports in the Reviews and Health Technology Assessments (HTA) unit in the Heath Services division. Approval of the project plan requires an internal review by a current or former leader in the Health Services division. The final report will be reviewed by two internal research directors at NIPH and at least two external experts. The final draft will be approved by the management group in the HTA-unit before submission to the New Methods system. The completed report will be published by the Norwegian Institute of Public Health and by the New Methods system (https://nyemetoder.no/metoder), generally, within 10 days of submission to New Methods Commissioning forum.

Activities and schedule

As this commission is a revision of an earlier commission following the submission of a mapping report to facilitate the choice of a specific patient population, so we had already conducted many of the preliminary steps in the project process when the revised commission was received. Unfortunately, constraints resulting from Covid-19, summer holidays, availability of key personnel and need for further clarification of included drugs have delayed our progress. The following information reflects the actual starting points for the work that is currently in process, and the expected starting points for aspects of the project that cannot be done until other work is completed, as well as expected ending dates for work that is not yet completed.

- Find and include external reviewers
- Discuss and revise project plan with project experts and internal reviewer
- Approval of project plan
- Select studies according to inclusion/exclusion criteria
- Evaluate methodological quality
- Extract data on efficacy and safety and conduct statistical analyses
- Collect data required for the health economic analysis, design model, and conduct analyses
- GRADE evaluation for each outcome
- Write and review the draft report
- Approve and submit the report

Date for commission

30.03.2020

Start date (for FHI.no): 30.03.2020

End date: planned 30.06.2021

Table 4. Activities and schedule

Activities	Responsible	Start	Finish
Start to write project plan	AD/IKO/CR	20.04.2020	05.10.2020
Project plan accepted by internal project members	Internal project group	06.10.2020	30.11.2020
Send project plan to external experts	AD	04.12.2020	11.12.2020
Revise and send project plan to internal reviewer (KB)	AD	14.12.2020	18.12.2020
Revise and send project plan to Klyngeledelsen	AD	04.01.2021	05.01.2021
Approval of project plan by Klyngeledelsen	Klyngeledelsen	21.12.2020	29.12.2020
Selection of articles (screening)	IKO/LG	25.09.2020	15.10.2020
Selection of full text articles	IKO/LG	15.10.2020	30.10.2020
Assess risk of bias in included articles	IKO/LG	02.01.2021	31.01.2021
Data extraction (table characteristics of included articles)	IKO/LG	02.11.2020	31.01.2021
Clinical data analysis (network meta-analysis)	CR	01.02.2021	02.04.2021
GRADE	IKO/LG	05.04.2021	30.04.2021
Collect data for health economic analysis	AD/UL/AE	10.08.2020	19.02.2021
Plan and build economic model	AD/AE	22.02.2021	029.04.2021
Conduct health economic analyses	AD/UL	09.04.2021	07.05.2021
Write and finalize the HTA	AD/Internal project group	ongoing	28.05.2021
Send draft to external experts/peer reviewers for comments and approval	AD/internal project group	28.05.2021	11.06.2021
Revise, send draft to internal peer review (KB)	AD (IKO/CR if needed)	11.06.2021	14.06.2021
Final revisions after internal peer review	AD	18.06.2021	21.06.2021
Send and approval of report by Klyngeledelsen	AD/Klyngeledelsen	21.06.2021	29.06.2021
Send HTA to Bestillerforum and publish at FHIs webpage	AD	29.06.2021	30.06.2021

Publication / dissemination

The final report will be published by the Division of Health Services, Norwegian Institute of Public Health and accessible on the NIPH website (https://www.fhi.no/) and that of the Nye metoder system (https://nyemetoder.no). We may also submit a scientific article.

Indexing for web page

Internal Myeloma related projects/publications

There is one previous project published by the Norwegian Institute of Public Health on this topic (for more information, follow the links below).

References

1. Ohm, I.R., C; Hafstad, E; Zinöcker, S; Lund, UH; Desser, AS, *Legemidler for behandling av myelomatose*. 2020, Folkehelseinstituttet: Oslo, Norway.
2. Durie, B., *Pasient-håndbok: Myelomatose*. 2017, International Myeloma Foundation.
3. Cancer Registry of Norway, *Cancer in Norway 2018 - Cancer incidence, mortality, survival and prevalence in Norway*. 2019: Oslo.
4. Helsedirektoratet, *Nasjonalt handlingsprogram med retningslinjer for diagnostikk, behandling og oppfølging av maligne blodsykdommer*. 2020, Helsedirektoratet: Oslo, Norway.
5. Folkehelseinstituttet, *Slik oppsummerer vi forskning. Håndbok for Nasjonalt kunnskapssenter for helsetjenesten*. 2015, Folkehelseinstituttet: Oslo. p. 71.
6. Folkehelseinstituttet, Statens Legemiddelverk, and Sykehusinnkjøp RHF, *Notat til Bestillerforum RHF - Forslag om fullstendige metodevurderinger*. 2019, Folkehelseinstituttet: Oslo, Norge.
7. Ouzzani, M.H., Hossam; Fedorowicz, Zbys; Elmagarmid, Ahmed *Rayyan — a web and mobile app for systematic reviews*. Systematic Reviews, 2016. **5:201**.
8. Julian PT Higgins, et al., *Chapter 8: Assessing risk of bias in a randomized trial*, in *Cochrane Handbook for Systematic Reviews of Interventions* J. Higgins, et al., Editors. 2020, Cochrane.
9. Veritas Health Innovation, *Covidence systematic review software*. Melbourne, Australia.
10. Jakobsen, J.C., et al., *When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts*. BMC Med Res Methodol, 2017. **17**(1): p. 162.
11. Rücker, G., C.J. Cates, and G. Schwarzer, *Methods for including information from multi-arm trials in pairwise meta-analysis*. Res Synth Methods, 2017. **8**(4): p. 392-403.
12. Page, M.J., J.P. Higgins, and J.A. Sterne, *Assessing risk of bias due to missing results in a synthesis*, in *Cochrane Handbook for Systematic Reviews of Interventions*. 2020, Cochrane.
13. Rücker, G., *Network meta-analysis, electrical networks and graph theory*. Res Synth Methods, 2012. **3**(4): p. 312-24.
14. Schwarzer, G., J.R. Carpenter, and G. Rücker, *Meta-Analysis with R* 1ed. Use R! 2015: Springer International Publishing.
15. Dias, S., et al., *NICE Decision Support Unit Technical Support Documents*, in *NICE DSU Technical Support Document 2: A Generalised Linear Modelling Framework for Pairwise and Network Meta-Analysis of Randomised Controlled Trials*. 2014, National Institute for Health and Care Excellence (NICE)

Copyright © 2014 National Institute for Health and Clinical Excellence, unless otherwise stated. All rights reserved.: London.

16. Dias, S.W., NJ; Sutton, AJ; Ades, AE, *Evidence Synthesis for Decision Making 2: A Generalized Linear Modeling Framework for Pairwise and Network Meta-analysis of Randomized Controlled Trials*. Medical Decision Making, 2013. **33**(5): p. 607-17.
17. Carpenter, B., et al., *Stan: A Probabilistic Programming Language*. Journal of Statistical Software, 2017. **76**(1).
18. Hutton, B., et al., *The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: checklist and explanations*. Ann Intern Med, 2015. **162**(11): p. 777-84.
19. Dias, S., et al., *Checking consistency in mixed treatment comparison meta-analysis*. Stat Med, 2010. **29**(7-8): p. 932-44.
20. Guyatt, G.H., et al., *GRADE guidelines: 13. Preparing summary of findings tables and evidence profiles-continuous outcomes*. J Clin Epidemiol, 2013. **66**(2): p. 173-83.
21. Higgins, J., et al., *Standards for the conduct and reporting of new Cochrane Intervention Reviews, reporting of protocols and the planning, conduct and reporting of updates*, in *Methodological Expectations of Cochrane Intervention Reviews (MECIR)*. 2020, Cochrane.

22. Rücker, G. and G. Schwarzer, *Ranking treatments in frequentist network meta-analysis works without resampling methods.* BMC Med Res Methodol, 2015. **15**: p. 58.
23. Salanti, G., A.E. Ades, and J.P. Ioannidis, *Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial.* J Clin Epidemiol, 2011. **64**(2): p. 163-71.
24. Welton, N.J., et al., *Mixed Treatment Comparison Meta-Analysis of Complex Interventions: Psychological Interventions in Coronary Heart Disease.* American Journal of Epidemiology, 2009. **169**(9): p. 1158-1165.
25. The GRADE Working Group, *GRADE Handbook*, in *Handbook for grading the quality of evidence and the strength of recommendations using the GRADE approach*, Holger Schünemann, et al., Editors. 2013, Cochrane.
26. Brignardello-Petersen, R., et al., *Advances in the GRADE approach to rate the certainty in estimates from a network meta-analysis.* J Clin Epidemiol, 2018. **93**: p. 36-44.
27. Cochrane Effective Practice and Organisation of Care Group, *Reporting the effects of an intervention in EPOC reviews*. 2018, Cochrane.
28. Woods, B., et al., *Partitioned survival analysis for decision modelling in health care: a critical review.* NICE DSU Technical Support Document, 2017. **19**.