The relative importance of musculoskeletal research questions

Protocol version 161116

1   INVESTIGATORS

This project is being led by the Executive Committee of the Australian and New Zealand Musculoskeletal Clinical Trials network (ANZMUSC):

Associate Professor William Taylor (Project Lead), University of Otago, Wellington.

Professor Rachelle Buchbinder (Chair), Cabrini Hospital & Monash University, Melbourne.

Professor Ian Harris (Chair SAG), University of New South Wales, Sydney

Dr Sam Whittle, Queen Elizabeth Hospital, Adelaide.

Dr Bethan Richards, Royal Prince Alfred Hospital, Sydney.

Professor Chris Maher, The George Institute for Global Health, Sydney.

Professor Jane Latimer, The George Institute for Global Health, Sydney.

Ms Ornella Clavisi (Chair CAG), Muscle Bone & Joint Health (MOVE), Melbourne.

2   OBJECTIVES

1. To develop a scoring system that will allow for the explicit and transparent ranking of musculoskeletal research questions in terms of their intrinsic importance. In other words, development of a framework to determine what makes a research question important enough to receive criterion 1 endorsement by ANZMUSC.
2. Develop a non-exhaustive list of MSK research questions that are ranked by importance to provide some guidance to researchers.

3   BACKGROUND

3.1   RESEARCH PRIORITY SETTING

Research priority setting can be considered at a number of levels, depending on precisely what is being prioritised. "Prioritisation" means placing the research in some kind of order, with the best research first. This ordering (as to what 'best' means) can be done in terms of funding decisions, scientific rigor, feasibility, etc. The particular value to which this project refers to (and is what is being ordered), is the intrinsic **importance of a research question** and arises directly from the first criteria that ANZMUSC has formulated for endorsement of a research project:

1. *Must satisfy an ANZMUSC research priority (reflecting important disease burden and an important evidence- or evidence-practice gap)*

This criterion reflects the need to investigate the 'big' questions – those research questions which truly have the potential to make a major difference for people with MSK conditions. Issues such as feasibility, the ability of the research team, and scientific rigor are examined by other endorsement criteria. Criterion 1 refers to the intrinsic importance of the research question itself: is it even worth trying to answer this question?

For the purpose of this project description and protocol, prioritisation of research questions refers explicitly to the intrinsic importance of the question.

A scoping review of previous research priority setting in musculoskeletal (MSK) disorders is being undertaken separately. There are many alternate ways of developing a priority list. In this section, the approach taken by the James Lind Alliance (JLA) is briefly reviewed and critiqued as an example of the standard approach. The ultimate product of a JLA exercise is a list of the top 10 topics that require addressing by research. For MSK conditions, only 2 top 10 sets developed by the JLA have been identified (Table 1). It is interesting to note that this approach does not necessarily lead to answerable research questions or testable hypotheses, despite the specificity of the topics. For example, it is not clear how the question 'How can we ensure the patients see the right doctors and clinicians promptly and correctly, and does this lead to better outcomes (results)?' can be translated into an answerable research question.

The JLA approach involves a Priority Setting Partnership (coordinator, patients, clinicians, stakeholders) who gather uncertainties within the area of interest often by survey, verify that these actually are uncertainties (by literature review), turn these into a list of research questions, shorten the list by surveying stakeholders for their top 10 research questions, and then finalize the top 10 in a workshop setting.

Table 1. The top 10 MSK research questions from JLA in UK	
Shoulder surgery (2015)	Hip and knee joint replacement for OA (2014)
For the main shoulder conditions of arthritis, frozen shoulder, impingement, rotator cuff tears and instability, can you predict which patients will do well with surgery to help them decide on whether to have surgery or not?	What are the most important patient and clinical outcomes in hip and knee replacement surgery, for people with osteoarthritis, and what is the best way to measure them?
In patients with 3 and 4 part proximal humeral fractures what is the long term outcome of reverse total shoulder replacement compared to hemiarthoplasty (half shoulder replacement)?	What is the optimal timing for hip and knee replacement surgery, in people with osteoarthritis, for best post-operative outcomes?
Does arthroscopic (keyhole) subacromial decompression surgery in patients with degenerative rotator cuff tendon problems improve outcome and prevent further tendon degeneration and tears compared to patients with no surgical intervention?	In people with osteoarthritis, what are the pre-operative predictors of post-operative success (and risk factors of poor outcomes)?

Does early mobilisation and physiotherapy after shoulder surgery improve patient outcome compared to standard immobilisation and physiotherapy?	What (health service) pre-operative, intra-operative, and post-operative factors can be modified to influence outcome following hip and knee replacement?
In patients with shoulder arthritis is a hemiarthroplasty (half shoulder replacement) or a total shoulder replacement or a reverse (ball on shoulder socket and socket on arm bone) replacement most effective?	What is the best pain control regime pre-operatively, peri-operatively and immediately post-operatively for hip and knee joint replacement surgery for people with osteoarthritis?
Are patients (including older age groups) with rotator cuff tendon tears in their shoulder best treated with surgery or physiotherapy?	What are the best techniques to control longer term chronic pain and improve long term function following hip and knee replacement?
How can we ensure the patients see the right doctors and clinicians promptly and correctly, and does this lead to better outcomes (results)?	What are the long-term outcomes of non-surgical treatments compared with operative treatment for patients with advanced knee/hip osteoarthritis?
In patients with Frozen Shoulder, does early surgery improve outcome compared to non-surgery treatments such as injection and dilatation?	What is the most effective pre and post-operative patient education support and advice for improving outcomes and satisfaction for people with osteoarthritis following hip/knee replacement?
In patients with newly diagnosed calcific tendinitis (calcium in a shoulder tendon), is early surgical intervention more clinically effective than non-operative treatments?	What is an ideal post-operative follow up period and the best long term care model for people with osteoarthritis that have had hip/knee replacement?
Do patients with partial thickness rotator cuff tendon tears benefit more from a surgical repair compared to a decompression and debridement (cleaning up operation) alone?	What is the best way to protect patients from the risk of thrombotic (blood clots, bleeding) events associated with hip/knee replacement?

There are several problems with this approach. Firstly, the level of specificity and granularity of the subject under consideration needs to be matched to the range of relevant stakeholders. For a topic such as MSK disorders, the range of stakeholders is very (impractically) large. Secondly, the number of uncertainties generated within a broad topic area will also be impractically large. Furthermore, examining the literature for evidence that each uncertainty really is uncertain is a mammoth undertaking for a very large number of uncertainties. Thirdly, the process takes no account of new knowledge and technology arising beyond the time-point that the uncertainty generation exercise took place. Thus the prioritisation exercise is immediately out of date. In addition, the identification of 10 important questions appears arbitrary – is the 11th ranked research question really unimportant enough not to be supported? Finally, the cognitive task required to select top 10 research questions from a list of possibly 100's of questions is substantial and the process may therefore not necessarily reflect participant views. Moreover, the process provides no insight into why certain questions are prioritised more than others, and may be excessively dependent upon the kinds of participants who took part in the exercise.

3.2 PRIORITISATION IN OTHER APPLICATIONS

Priority setting is a common requirement in a range of situations where there are limited resources. Individuals, institutions and governments must constantly make decisions about how to spend their money or use their resources. The process for all priority setting is conceptually similar – either all possible alternatives are considered and ranked by consideration of each alternative in relation to all the others without really considering why (the JLA approach), or the underlying structure of what makes one alternative more important than another is uncovered, and this structure is used to inform priorities.

There are many example of the latter approach in marketing research, where it is useful to understand the factors that go into a person's preference for one product over another. In general, these are examples of multi-attribute decision-making, where more than one issue informs the decision. In healthcare, a clearly analogous example is priority setting for elective surgery. In countries where surgery cannot be offered to all those who might benefit (because of resource constraints) the traditional method of prioritisation was the waiting list. However, access to treatment determined by the time arrived on the waiting list is neither fair nor maximises the outcomes for a given resource. In New Zealand, patients are prioritised for different types of surgery based upon a transparent system that explicitly weights priority based upon factors that determine the extent of benefit and need for each type of surgery (1). A similar process was undertaken in Victoria, resulting in the Multi-attribute Arthritis Priority Tool (MAPT) (2, 3).

This process comprises 3 steps: (1) identification of factors that determined surgeons' views about how they prioritised patients for elective surgery and how these factors could be turned into categories indicating less to more priority; (2) using a discrete choice experiment, the relative weighting of the identified factors/categories could be determined; (3) finally, the threshold at which a derived prioritisation score should be set to allow access to elective surgery, given a particular set of resource constraints.

A similar process can be used to uncover the relative importance of MSK research questions:

(1) Identify the factors that determine stakeholders' views of what makes a research question important or not important, and how these factors can be turned into categories indicating less to more importance;

(2) Perform a discrete choice experiment to identify the relative value of the identified factors/categories; and finally,

(3) Determine the threshold at which a particular question can be said to be sufficiently important for ANZMUSC endorsement.

3.3 WHAT ARE THE IMPORTANT MSK RESEARCH QUESTIONS?

At the same time as developing an explicit and transparent framework for judging what makes a research question important, we plan to generate a list of potentially high priority research questions. Research questions can arise from multiple sources and it is likely that no single source would produce all the research questions of importance (4).

We propose to generate a list of MSK research questions using the following sources:

- Web-Survey of ANZMUSC members
- Workshop format of clinicians and consumers to identify contemporary clinical issues that participants would like addressed by research
- Review of Therapeutic Guidelines: Rheumatology, 3rd edition to identify recommendations that were made based upon either low quality evidence or consensus
- Review of all Cochrane reviews relating to MSK conditions (mostly published by Cochrane Musculoskeletal and Cochrane Back) to identify where implications for research included the need for further high quality evidence
- Review of MSK questions produced by previous priority setting projects internationally.
- Review of Choosing Wisely and like initiatives such as EVOLVE for MSK conditions to identify evidence-practice gaps

3.4   IMPLICATIONS OF THE ANZMUSC PRIORITY LIST OF RESEARCH QUESTIONS

Based upon the development of a transparent method of ranking the importance of research questions that we can apply to a large list of MSK research questions identified from a variety of sources, we will generate an ANZMUSC Priority list of research questions.

It is important to note that this list will not preclude researchers submitting proposals that address different, and also potentially important questions, that might rank higher than questions generated in this exercise. It is also possible that ANZMUSC itself, or the CAG might subsequently identify new research questions that are also ranked high in importance. Therefore, this list of potentially important MSK research questions should be considered dynamic and evolving.

4   METHODOLOGY

This project consists of a series of opinion surveys in which stakeholder opinion is elicited in various ways: web-based Delphi survey, consensus workshop, discrete choice experiment, literature review. An overview of the project is shown below (each element is subsequently described in more detail in Figure 1).

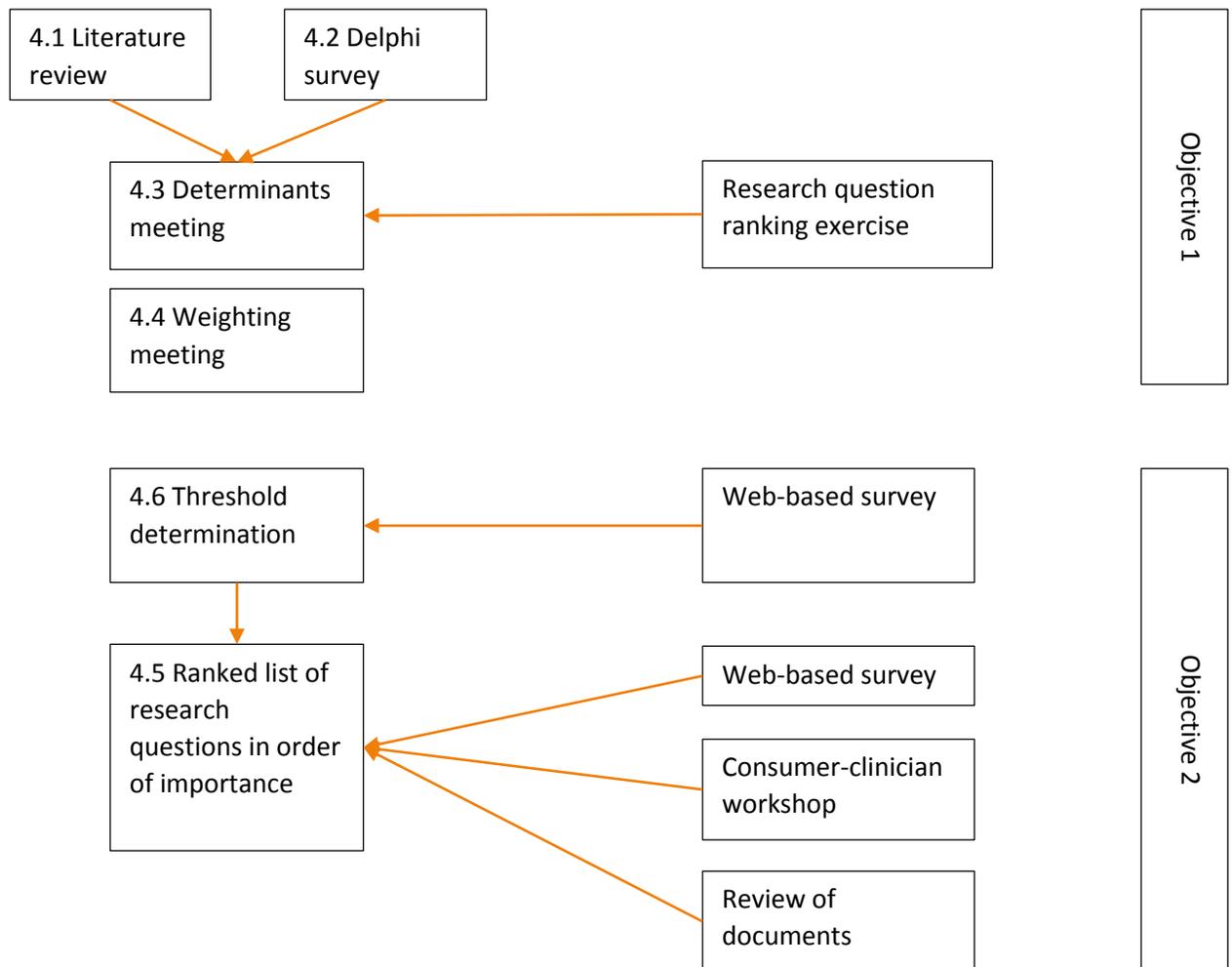The text is divided into 2 parts, supporting objective 1 and objective 2.

Figure 1: Overview of the project

Methodology to support objective 1

4.1   LITERATURE REVIEW

A systematic literature review of MSK research priority setting. While the focus of this review is the methods employed to elicit high priority research questions, it will also identify what those research questions are and the factors that might underlie the decision to make them high priority.

For the purposes of this project, the main outcome of interest from the literature review is a list of possible determinants that contribute to the importance of a research question.

4.2   DELPHI SURVEY

Purpose – this web-based survey is to elicit the main determinants of what makes a research question important, in the opinion of ANZMUSC members[1].

Participants – ANZMUSC members from the email membership list

Process – using SurveyMonkey or similar web-based survey tool, participants will engage in a series of surveys. The first survey will consist of simple open ended questions to elicit as many underlying determinants of what makes a research question important. Together with the findings from 4.1 Literature review, the determinants will be organized into categories and similar items merged to produce a list of determinants for rating. Second and subsequent iterations of the survey will ask participants to rate each determinant for how relevant that particular item is for determining the importance of a research question (scale of 1 to 9, low to high relevance). Where there is significant disagreement (defined by the UCLA/RAND disagreement index (5)), items will be re-rated in subsequent surveys. The group opinion about each of these items (median rating with interpercentile range) will also be displayed so that respondents have an opportunity to change their rating in light of the group's opinion. Items for which there is uncertainty (median rating 4 to 6) will also be re-rated in subsequent iterations. Survey iterations will cease when there is no more disagreement, no items with median rating 4 to 6 or there are no changes in how participants rate remaining items.

Outcome – the results of this Delphi survey will be a list of the determinants of what makes a research question important (according to ANZMUSC members) ranked in order of median relevance rating.

4.3   RESEARCH QUESTION RANKING EXERCISE

Purpose – to describe 30 'paper' MSK research questions in terms of the most relevant determinants (median rating of at least 7 from the Delphi survey) and to describe the variation in how these research questions are ranked in importance

Participants – ANZMUSC members (including affiliated funders and consumers) selected to participate in the "determinants meeting" workshop (see below) in which an evaluation framework for deciding on the importance of a research question is constructed

Process – All ANZMUSC members will be asked to submit 3 to 5 research questions together with a description of the research question in terms of the most relevant determinants of importance (defined as a median rating of at least 7 from the 4.2 Delphi survey). These determinants will be listed in a pro-forma for members to complete when they submit each research question. Members will be asked to submit questions that are both high and low in importance. From these research questions, 30 will be selected that provide a range of research questions from very high importance to very low importance.

---

[1] A range of stakeholders might be considered relevant here – consumers, clinicians and funders, but since the ultimate objective is to figure out how ANZMUSC decides what is an important research question, arguably the only relevant stakeholders are ANZMUSC members. This is a potentially controversial point and needs to be discussed further.

These 30 research questions will then be assessed and ranked in order of importance, independently, by each member selected to participate in the "Determinants meeting" workshop described below (n=20). These data will be summarized by a plot showing the distribution of rankings for each research question, in order of median ranking.

Outcome – the variation in how the relative importance of research questions are determined, using simple judgement, will be presented graphically. Research questions in which there was very high and very low variation can be used to help guide discussion about what really determines the importance of a research question in the next phase of the project.

4.4   DETERMINANTS MEETING WORKSHOP

Purpose – to arrive at a manageable framework of the determinants of the importance of a research question, and is constructed so that each determinant can be reliably assessed.

Participants – 20 selected members of ANZMUSC (including associate members). This will comprise 4 consumers, 4 funders/policy makers, 6 clinician-researchers and 6 research-only members.

Process – a 1.5 day consensus workshop meeting where the results of the preceding studies are presented and facilitated discussion leads to an agreed list of determinants and categories of each determinant. A structured discussion using the previously ranked research questions that display very high and very low variability will help guide the process. The wording of each determinant and category of each determinant will be carefully discussed. All participants will need to agree to the framework to the extent there is no serious disagreement.

The inter-rater reliability of a preliminary framework will be tested by asking each participant to categorize a random selection of 10 research questions (from the existing 30 research questions) according to the framework. The inter-rater reliability of categorizing each determinant will be assessed using Gwet's AC1 (which is less affected by the distribution of category prevalence than is Cohen's κ) using AgreeStat2015.6 software.

For determinant categories with less than 70% chance-corrected agreement further discussion on the inclusion, measurement or wording of the determinant and its categories will occur before re-categorizing a second randomly selected set of 10 research questions. Where significant disagreement still occurs serious consideration will need to be given to discarding that determinant, unless further adjustment can be made to plausibly enhance its objectivity.

Outcome – a framework of determinants (attributes) with associated categories that can be reliably used to describe MSK research questions.

4.5   DEVELOPMENT OF A SCORING SYSTEM

Purpose – since it is implausible that each determinant is equally relevant or important in determining the importance of a MSK research question, the purpose of this section is to assign meaningful numerical weights to each category so that a summated score can be generated that indicates the overall importance of a MSK research question.

Participants – 20 invited members of ANZMUSC (including consumers and funders) who may be different from the determinant workshop meeting (4.4).

Process – a 1-day workshop in which a discrete choice experiment is conducted using hypothetical research questions which are described only in terms of the evaluative framework identified in 4.4. Each decision will consist of choosing which is the MOST important research question when only 2 research questions are present and described only in terms of 2 attributes at a time. The decision will be made by group consensus. 1000Minds© software will be used to implement and analyse this process. The number of choice decisions that need to be made will depend upon the number of determinants and categories but will typically be more than 60.

A typical choice decision as presented by 1000Minds© might look like (Figure 2):

Figure 2. Format of a typical choice decision

In addition, the validity of this scoring system can be tested by regression to the rank order of the 30 test MSK research questions that were ranked prior to the determinants workshop meeting (4.4).

Outcome – a scoring system for the evaluative framework. 1000Minds© displays this in such a way that the highest scoring categories in each determinant will sum to 100, but it is quite possible to simplify the category scores by rounding and linear transformation to give a computationally easier scoring system.

Methodology to support objective 2

4.6   INITIAL RESEARCH QUESTIONS

Purpose – to identify a bank of research questions that are ranked in order of importance (using the evaluative framework in 4.5), thus providing a resource for researchers and funders.

Participants – for the first part, all ANZMUSC members using the email list; for the second part, invited consumers and clinicians; the third part is a document review

Process – MSK research questions will be generated by ANZMUSC members using two approaches. Firstly, a web-based survey of ANZMUSC members will elicit research questions that the member thinks he/she and/or ANZMUSC would like to pursue. Secondly, a workshop for consumers (n=5) and clinicians (n=5) where a Nominal Group Technique (6) will be used to anonymously generate important issues that require research, followed by group discussion to stimulate other ideas and research questions.

In addition, research questions will be identified by structured review of the following documents:

- Review of Therapeutic Guidelines: Rheumatology, 3rd edition to identify recommendations that were made based upon either low quality evidence or consensus
- Review of all Cochrane reviews relating to MSK conditions (mostly published by Cochrane Musculoskeletal and Cochrane Back) to identify where implications for research included the need for further high quality evidence
- Review of MSK questions produced by previous priority setting projects internationally.
- Review of Choosing Wisely and like initiatives such as EVOLVE for MSK conditions to identify evidence-practice gaps

These research questions or ideas will be organized into similar themes and individual PICOT-structured research questions will be constructed from these themes.

Each research question will be scored and rank ordered using the evaluative framework from 4.5.

Outcome – a list of ranked (in order of importance) MSK research questions.

4.7   THRESHOLD SCORE

Purpose – to identify the scoring threshold at which ANZMUSC members perceive a research question to be sufficiently important to meet the first criterion of the endorsement criteria

Participants – ANZMUSC members by email list.

Process – ANZMUSC members will be asked to examine 30 randomly selected MSK research questions from the question bank in 4.6, and asked to indicate whether each research question is sufficiently important in their view to meet the requirements of Criterion 1 of the ANZMUSC endorsement criteria.

The results can be expressed graphically as shown in Fig 3 below (shown for an example situation where diagnosis of gout was the issue of interest). In this plot, the evaluation score is shown on the *x*-axis and the number of respondents who believe that the case/item/question is indeed a case (is important enough for criterion 1) is on the *y*-axis. The relationship between probability of meeting Criterion 1 and the evaluation score will usually be logistic. The most appropriate threshold that minimizes false negatives and false positives can be easily identified directly from the plot, or a logistic function can be fitted to the data and the value of the score when the slope starts to flatten can be determined mathematically.

Outcome – a ranked (by importance) list of MSK research questions that also displays the threshold below which questions are not perceived as important enough to be endorsed by ANZMUSC. The list of questions above the threshold should be seen as dynamic, since new questions that meet the threshold score can be added to the list over time. Also, it is possible that the score for some questions will change over time as new information about the topic becomes available. Therefore, each research question in this list should be rescored each year. It may be useful to have the endorsable and regularly updated list of research questions available for general display on the ANZMUSC website.
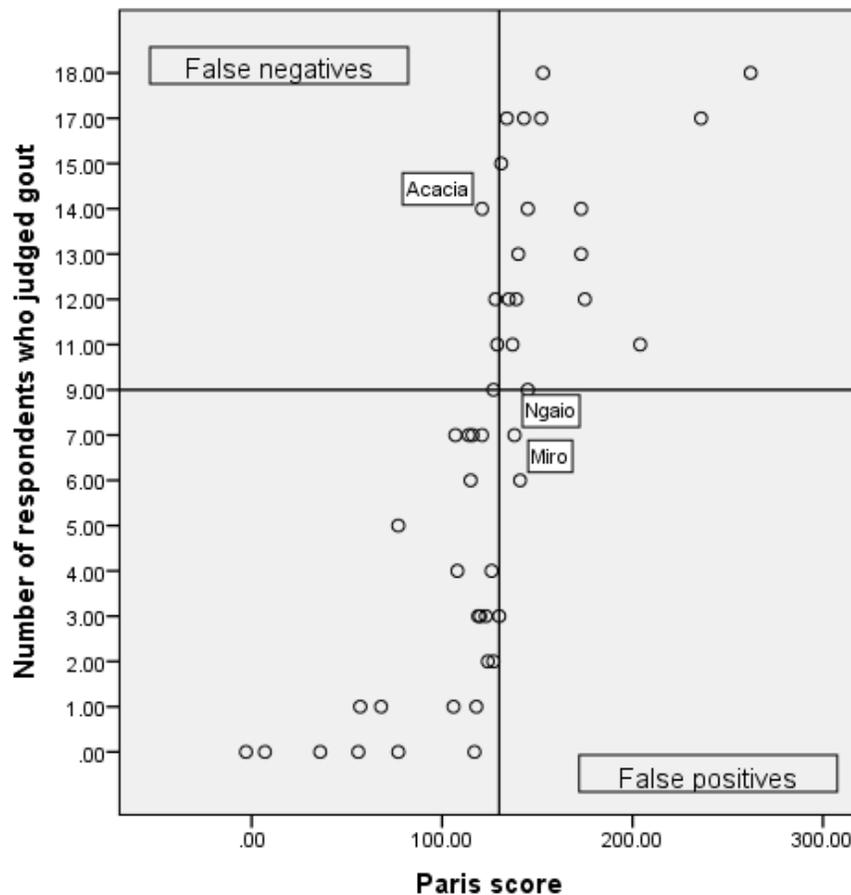

Figure 3: the relationship between the classification score ("Paris score") and the number of participants who felt that the case (open circles, each case was labelled with the name of a tree for identification purposes only) should be classified as gout for the purpose of enrolment into a clinical trial of gout treatment.

5   ANTICIPATED OUTCOMES

The culmination of the work is two products:

1.   A transparent and robust method for determining the importance of a MSK research question.

2. A list of contemporary MSK research question ranked by importance, which meet the ANZMUSC Criterion 1 for endorsement.

6 TIMEFRAME

Activity	Timing	Milestone outcome
4.1 Literature review		
4.2 Delphi web survey	Dec16 – Feb17	29 Feb 2017 List of potential determinants (attributes)
Question ranking exercise by web survey	Mar17 – Apr17	30 April 2017 Distribution of rankings for 30 research questions
4.3 Determinants workshop	May17	11-13 May, workshops to run sequentially following the annual ANZMUSC meeting
4.4 Weighting workshop	May17	
4.5a Research question generation web survey	Mar17 – May17	9 May 2017 List of research questions for evaluation
4.5b Research question generation workshop	Mar17	Date TBD
4.5c Organising and scoring research questions	Apr17 – Jul17	30 July 2017 A ranked list of MSK research questions
4.6 Threshold web survey	Aug17 – Oct17	31 October 2017 A ranked list of MSK research questions that meet ANZMUSC Criterion 1 for endorsement

7 BUDGET

		2016-17
Research Priority Setting Protocol		
Indicative Budget		
		2016-17
	Salary Costs	
	Research project manager, PRP4 0.5 FTE	44,111
	Methodological advisor/qualitative researcher, PSP5 0.1 FTE	9,498
	Recruitment Costs (2 positions)	500
	Computer expenses	

	(1 x laptop computer; 1,000 minds software funded from academic award)	2,500
	Workshop expenses	
	(Workshop 4.3 & 4.4=2.5 days)	
	Venue hire @ $470/day (Darlington Centre 20pax)	1,175
	Catering (20 pax, 2.5 days @ $39/day Darlington Centre or similar full day catering light package)	1,950
	Travel	
	Accomodation (Sydney/Melbourne: Novotel Central or similar @ $215/night x 2 nights: 20 pax)	8,600
	Flights (20 pax, Interstate/NZ: $600)	12,000
	Per Diem for dinner (interstate guests;10 pax @ $60 x 2nights)	1,200
	Facilitator travel (US)	3,000
	(Workshop 4.5=1 day)	
	Venue hire @ $470/day (Darlington Centre 20pax)	470
	Catering (10 pax, 1 day @ $39/day Darlington Centre or similar full day catering light package)	390
	Travel	
	Accomodation (Sydney/Melbourne: Novotel Central or similar @ $215/night x 10 pax for 1 night)	2,150
	Flights (10 pax, Interstate/NZ: $600)	6,000
	Printing & Stationery	250
	Telephone & Internet costs	300
	Printing & Courier costs for workshops	200
	Publication Costs (SR, Delphi Survey, Priorities paper)	2,500
	TOTAL GRANT REQUESTED	96,794

8 REFERENCES

1.   Hadorn DC, Holmes AC**.** The New Zealand priority criteria project. Part 1: Overview. BMJ. 1997;314(7074):131-4.

2.   Osborne R, Bucknill A, De Steiger R, Brand C, Graves S**.** VALIDATION OF THE MULTI-ATTRIBUTE PRIORITISATION TOOL (MAPT) FOR HIP AND KNEE OSTEOARTHRITIS AGAINST WOMAC, OXFORD HIP/KNEE SCORE. Journal of Bone & Joint Surgery, British Volume. 2012;94-B(SUPP XXIII):37-.

3.   Witt J, Scott A, Osborne RH**.** Designing choice experiments with many attributes: an application to setting priorities for orthopaedic waiting lists. Melbourne: Melbourne Institute of Applied Economic and Social Research; 2006.

4.   Clavisi O, Bragge P, Tavender E, Turner T, Gruen RL**.** Effective stakeholder participation in setting research priorities using a Global Evidence Mapping approach. J Clin Epidemiol. 2013;66(5):496-502 e2.

5.   Fitch K, Bernstein SJ, Aguilar MD, Burnand B, LaCalle JR, Lazaro P, et al. The RAND/UCLA Appropriateness Method User's Manual. Santa Monica, CA: RAND; 2001.

6.   Van de Ven A, Delbecq A**.** Nominal versus interacting group processes for committee decision-making effectiveness. Acad Manage J. 1971;14(2):203-12.