# Statistical Analysis Plan (SAP)

Trial: Better Knee, Better Me Study: Effectiveness of two scalable health care interventions supporting self-management for knee osteoarthritis – a randomised controlled trial.

## Contents

# Section 1. Administrative Information

## 1. Title

Better Knee, Better Me Study: Effectiveness of two scalable health care interventions supporting self-management for knee osteoarthritis – a randomised controlled trial.

## 2. Trial registration

Prospectively registered (Australian New Zealand Clinical Trials Registry Trial Id: ACTRN12618000930280 (01/06/2018))

## 3. SAP version

Version: 1.0 Date: 4/03/2021

## 4. Protocol Version

This document has been written based on information contained in the BKBM study protocol version 1.5 dated 31/03/2018. The protocol was published as follows:

Bennell, K. L., C. Keating, B. J. Lawford, A. J. Kimp, T. Egerton, C. Brown, J. Kasza, L. Spiers, J. Proietto, P. Sumithran, Quicke JG, Hinman RS (2020). "Better knee, Better me™: Effectiveness of two scalable health care interventions supporting self-management for knee osteoarthritis–protocol for a randomized controlled trial." *BMC Musculoskeletal Disorders* **21**(1): 1-19.

## 5. SAP Revisions

Not applicable

## 6. Names and affiliations

Document prepared by Professor Kim Bennell, Professor Rana Hinman, Dr Belinda Lawford, University of Melbourne and Dr Jessica Kasza, Monash University.

Emails: KL Bennell k.bennell@unimelb.edu.au; RS Hinman ranash@unimelb.edu.au; BJ Lawford belinda.lawford@unimelb.edu.au; J Kasza jessica.kasza@monash.edu;

Signatures:

*Signature of senior statistician responsible (Dr Jessica Kasza):*

*Date: 04/03/21*

*Signature of chief investigator (Prof Kim Bennell/Prof Rana Hinman):*

*Date: 04/03/21*

# Section 2: Introduction

## 7. Background and rationale

Knee osteoarthritis (OA) is a major public health problem (1) and incurs enormous indirect and direct healthcare costs (2). All current clinical guidelines recommend non-surgical non-drug treatments for first-line management of OA, including education/advice, exercise, and, if appropriate, weight loss (3, 4). However, there is evidence that a substantial proportion of people with OA in Australia are not receiving recommended care (5). For example, although surgery is only recommended for those with severe symptomatic OA who do not respond to conservative management approaches, data from an Australian hospital found that 33% of people referred for orthopaedic management had not engaged in any non-drug conservative management methods (6). From 2005-06 to 2015-16, Australia has seen a 38% rise in the rate of total knee replacements for OA (7), yet 20% of people undergoing knee replacement surgery report unsatisfactory outcomes (8). Given that the prevalence of knee OA is projected to rise in the coming decades (2, 9), there is an urgent need to increase uptake of recommended conservative management options through implementation of effective, accessible and scalable models of service delivery.

The benefits of therapeutic exercise for people with knee OA are well-established (10), with similar efficacy to analgesics and non-steroidal anti-inflammatory drugs, but with fewer side-effects (11, 12), and with fewer risks than joint replacement surgery. Given that muscle weakness is common amongst people with OA (13), muscular strengthening exercises are important and can help reduce pain, and improve physical function (14). In addition, as most people with OA do not meet physical activity recommendations (15), advice to increase physical activity is also important. For example, research has demonstrated that people with knee OA who are less sedentary have better physical function, independent of time spent doing moderate or vigorous physical activity (16, 17), and that a threshold of 6,000 steps per day discriminates between people with knee OA who do and do not develop functional limitations two years later (18).

Given that people with OA who have overweight or obesity tend to experience more severe symptoms (19, 20) and are more likely to undergo joint replacement surgery (21), weight loss is an important component of management. However, although weight loss is recommended by evidence-based guidelines (4, 22, 23) there is only limited evidence to support its efficacy in improving symptoms of pain and dysfunction. In the literature, interventions often combine both exercise and weight loss components, making it difficult to identify the independent effects of weight loss on outcomes. One pivotal trial involving 454 adults with knee OA who had overweight or obesity investigated the effects of an intensive diet program, with and without exercise, conducted in a rigorously monitored environment (24). The combined diet and exercise group lost approximately 10% of body weight, compared to 2% in the exercise only group, leading to reduced knee loads and inflammation compared to the exercise only group. While statistically significant benefits on pain (1-point difference (95% confidence interval: 0.3 to 1.7) on 20-point scale) and function (4.3-point difference (2.1 to 6.5) on 68-point scale) were also seen, the magnitude of these benefits are small and may be of questionable clinical importance. Furthermore, although the diet and exercise intervention was reported to be more cost-effective than usual care (pharmacological NSAIDs regimen followed by total knee replacement surgery) (25), the cost-effectiveness analyses did not include data from the exercise only group. Thus, it remains unknown whether a combined intervention of diet and exercise is more cost-effective than exercise alone.

Although the findings of Messier and colleagues (24) support the efficacy of combined weight loss and exercise programs, their interventions were time-intensive and potentially costly, with the weight loss component involving 12 individual and 42 group sessions for nutrition education and behavioural support over 18-months, and the exercise component involving three group exercise sessions per week for 6-months. Given that many barriers to self-management of OA relate to inaccessibility and/or costs of healthcare (26, 27), implementation of such an intensive program outside of a research setting is likely to be difficult in most community or hospital settings, either public or private, without

considerable cost and time burdens. Identifying effective programs that minimise active involvement by clinicians is therefore important. One way in which to improve the accessibility and scalability of care is to provide it remotely via technology (telehealth), allowing patients to consult from their own home or workplace. There is emerging evidence that telehealth for people with musculoskeletal conditions improves pain and function and is equivalent to outcomes following traditional in-person consultations (28). More recently, there is evidence that exercise advice and prescription delivered by physiotherapists via videoconferencing leads to improvements in pain and function in people with knee OA (29). Importantly, such a service is also acceptable to both the people with knee pain receiving care and the physiotherapists delivering it (30). Similarly, dietary interventions for weight loss can also be effective when delivered remotely. A recent systematic review and meta-analysis found that dietary interventions for people with chronic diseases (i.e. obesity, diabetes, heart disease, hypertension, stroke, or kidney disease) delivered by dietitians or nurses via video or telephone improved diet quality and led to significant weight loss and reduced waist circumference (31). No previous studies have investigated the effectiveness of remotely-delivered exercise and weight management programs by physiotherapists and dietitians for people with knee OA who have overweight or obesity.

For many people with OA, undertaking an exercise and/or weight management program requires considerable changes in lifestyle behaviours over prolonged periods of time. As such, it is important that intervention development is underpinned by behaviour change theory, and informed by best practice in chronic disease management and self-management support programs. Interventions must include provision of high quality information. However, education alone is insufficient to support behavioural change. Healthcare clinicians must also have the communication and psychosocial skills necessary to facilitate long-term changes in behaviour (32, 33). Incorporating behavioural counselling (34, 35) and specific behaviour change techniques that are effective for supporting change in exercise and eating behaviours in this population (36) into intervention design will facilitate participants to achieve and sustain effective self-management practices in the long-term.

The aim of this study, conducted in a private health insurance setting, is to compare the effectiveness of two remotely-delivered, evidence- and theory-informed, behaviour change interventions to provide information and behaviour change support for i) exercise and; ii) exercise plus weight management, to each other, and to an information-only control for people with knee OA who have overweight or obesity. We hypothesise that the exercise plus weight management intervention will lead to greater improvements in pain and function than the exercise intervention and that both interventions will be more effective than information only.

## 8. Objectives

**Study objectives:**

Our primary research questions are:

For people with clinically diagnosed knee OA who have overweight or obesity:

(1) Is the *Exercise plus weight management* intervention more effective (greater pain relief and improved function) than the *Exercise* intervention at 6 months?

(2) Are both interventions more effective than information alone at 6 months?

**Research hypotheses:**

Primary hypotheses:
The *Exercise plus weight management* intervention will lead to greater improvement in knee pain and physical function when compared to the *Exercise* intervention at 6 months.

The *Exercise plus weight management* and *Exercise* interventions will lead to greater improvements in

knee pain and physical function when compared to the control (information-only) group at 6 months.

Secondary hypotheses:
The *Exercise plus weight management* intervention will lead to greater improvement in knee pain and physical function when compared to the *Exercise* intervention at 12 months.

The *Exercise plus weight management* and *Exercise* interventions will lead to greater improvements in knee pain and physical function when compared to the control (information-only) group at 12 months.

The *Exercise plus weight management* intervention will have greater benefits for other clinical outcomes (other pain measures, physical function, quality of life, global rating of change) compared to the *Exercise* and control (information-only) groups at 6 and 12 months.

# Section 3: Trial Methods

## 9. Trial design

The Better Knee, Better Me study is a three-arm, pragmatic, superiority parallel-design randomised controlled trial.

## 10. Randomisation

Randomisation will occur according to a 5:5:2 ratio (due to the different pre-specified effect sizes across comparisons). On completion of baseline questionnaire, participants will be randomly allocated into one of the three groups; i) *Exercise* intervention ii) *Exercise and weight management* intervention iii) Information only (Control)

The randomisation schedule will be computer generated, using random permuted blocks of varying sizes by a person not involved in recruitment of participants (JK, biostatistician). Allocation will be stratified by history of knee surgery (arthroscopy or contralateral arthroplasty). To conceal allocation, the randomisation schedule will be accessed via a password-protected, computer program by a researcher not involved in participant recruitment scheduling or assessment. The person who will determine if a potential participant is eligible for inclusion in the trial will be unaware, when this decision is made, to which group the participant will be allocated.

Participants allocated to the *Exercise* group will be randomly allocated to a physiotherapist. Participants allocated to the *Exercise plus weight management* group will be randomly allocated to a pairing of physiotherapist/dietitian.

## 11. Sample size

The primary outcomes of this trial include changes in pain (NRS) and physical function (WOMAC) over 6 months and the intervention will be considered beneficial if it changes either one or both of the primary outcomes. A Cochrane Review shows that exercise has a moderate effect size for pain and function in knee OA (10). Calculations thus assume an effect size between the least-performing treatment group (presumed to be the *Exercise* group) and the control group of 0.4. Given that the effect size between two treatments will be less than between treatment and control, we assumed an effect size of 0.3 between the *Exercise* group and the *Exercise plus weight management* group. We chose 0.3 because any smaller effect is unlikely to be clinically relevant or cost-effective to implement in real-world settings given the considerable expense of the weight management program relative to the exercise components of this combined intervention. Calculations also assume a correlation between baseline and follow-up measurements of 0.4, power 80%, 15% loss to follow-up (based on our previous

research (37, 38)), and two-sided significance level of 0.05. Based on these calculations we require 173 participants in each treatment group, and 69 participants in the control group arm, giving a total sample size of 415 participants.

There are different levels of clustering for each of the three groups; in the *Exercise* group, participants are clustered by physiotherapist, in the *Exercise plus weight management* group, people are clustered by physiotherapist and dietitian, and in the control group, there is no clustering. Given the geographical and demographic spread of participants, we expect that the intra-cluster correlation of outcomes treated by the same dietitian or physiotherapist to be of the order of 0.01. Modifying the formulae (39) to allow for correlations between baseline and follow-up measurements, and assuming that there are five dietitians delivering the *Exercise plus weight management* treatment, the effect size that can be detected between the *Exercise* and the *Exercise plus weight management* groups with 80% power given the originally calculated sample size increases slightly to 0.32.

### 12. Framework

This trial uses a superiority hypothesis testing framework between groups for all outcomes.

### 13. Statistical Interim analyses and stopping guidance

Nil

### 14. Timing of final analysis

Final analysis will be performed after all participants have reached the 12-month (secondary) timepoint and completed their outcome measures.

### 15. Timing of outcome assessments

Outcomes are collected at baseline, 6, and 12-months.

## Section 4: Statistical Principles

### 16. Level of statistical significance

All applicable statistical tests will be 2-sided and will be performed using a 5% significance level.

### 17. Description of any planned adjustment for multiplicity, and if so, including how the type 1 error is to be controlled

We have two co-primary outcomes: knee pain and physical function. We will not adjust for dual outcomes but instead report all effect sizes, confidence intervals, and p values in order to let readers use their own judgment about the relative weight of the conclusions on the effect of exercise and weight management for symptoms. In addition, we will not adjust for multiplicity due to multiple arms in this trial, since these three arms are distinct treatments (40, 41). This approach aligns with the usage of p-values favoured by the American Statistical Association (42).

### 18. Confidence intervals to be reported

All confidence intervals will be 95% confidence intervals.

### 19. Adherence and Protocol Deviations

The primary analysis will be based on the principle of intention-to-treat, whereby participants are included in the groups to which they were originally assigned, regardless of their adherence to their assigned treatments. Any protocol deviations (if they occur), including errors applying inclusion/exclusion criteria and/or administration of the wrong intervention will be summarised in trial results (patient flow diagram/text) by treatment group. Randomisation errors resulting from these errors will be handled according to recommendations (43).

Multiple measures of adherence to the two active interventions are used in this trial and data from all measures will be reported using means, standard deviations and proportions (number and percentage) as appropriate for each intervention group. Measures of adherence include:

- Number of consultations with the physiotherapist (where attendance at 4 or more consultations out of 6 consultations will be considered acceptable adherence)
- Number of consultations with the dietitian (where attendance at 4 or more consultations out of 6 will be considered acceptable adherence)
- Self-rated adherence to strengthening exercise (on a 11-point NRS)
- Self-rated adherence to physical activity (on a 11-point NRS)
- Self-rated adherence to weight management (on a 11-point NRS)

### 20. Analysis Populations

The primary analysis will be based on the principle of intention-to-treat, whereby participants are included in the groups to which they were originally randomised, regardless of their adherence to their assigned treatments.

## Section 5: Trial Population

### 21. Screening Data

Screening data will be collected and summarized. A CONSORT flow diagram will be used (44). The following summaries will be presented in text and/or flow diagram: time frame for recruitment, the number of participants screened, the number of participants recruited, the number of screened participants not recruited, and the reasons for non-recruitment.

### 22. Eligibility

Trial inclusion and exclusion criteria are described in the trial protocol. Reasons for exclusion will be summarized in the CONSORT (44) flow diagram.

### 23. Recruitment

A CONSORT flow diagram (44) will be used to describe the number of people enrolled, randomized, allocated to each treatment group, lost to follow up (including reasons) and analysed.

### 24. Withdrawal/follow-up

If a participant withdraws from the study, the nature, timing of and reasons for withdrawal will be described (provided the participant responds to contact made by the research team). Any data provided up to the point of withdrawal will be analysed in accordance with intention to treat analyses, unless the participant specifically requests to withdraw their data from the study. Losses to follow-up (including reasons) will be summarised in the CONSORT flow diagram by treatment group.

### 25. Baseline characteristics

Baseline characteristics will be summarised by treatment group and presented in a table:

- Age
- Sex
- Height, body mass, body mass index
- Country of birth
- Geographical location
- Aboriginal or Torres Strait Islander heritage
- Level of education
- Current employment status
- Duration of knee OA symptoms
- Length of time since first visit to doctor for knee pain
- Most painful knee
- Problems in other joints
- Current medication use
- Previous history of arthroscopy and arthroplasty
- Previous treatments and surgery
- Medical history

Baseline characteristics will be summarised as appropriate (means and standard deviations for continuous variables that appear to be distributed approximately symmetrically, medians and interquartile ranges for other continuous variables, counts and percentages for categorical variables). Tests of statistical significance will not be undertaken for comparing baseline characteristics of treatment groups; rather the clinical importance of any imbalance will be noted.

An appendix table will provide summaries of baseline characteristics and baseline levels of primary and secondary outcomes and compare these characteristics between two groups: those participants who provide both primary outcomes at 6 months, and those participants who are missing one or both primary outcomes at 6 months. T-tests will be used to compare continuous characteristics between these groups, and chi-squared tests will be used to compare categorical characteristics.

## Section 6: Analysis

### 26. Outcome definitions

**Co-primary outcomes:**

- Change in severity of knee pain: Average overall knee pain in the past week is self-assessed using a 11-point numeric rating scale (NRS) with terminal descriptors of ‘no pain’ (score 0) and ‘worst pain possible’ (score 10). Change score at 6 months (primary time point) and 12 months (secondary time point) will be calculated as follow-up minus baseline.

- Change in physical function: Physical function is self-assessed using the WOMAC physical function subscale, where scores range from 0 to 68 (lower scores indicate worse function). Change score at 6 months (primary time point) and 12 months (secondary time point) will be calculated as follow-up minus baseline.

**Secondary outcomes:**

- Change in weight: Self-reported weight in kilograms. Change score at 6 months and 12 months will be calculated as follow-up minus baseline.

- Change in physical activity: The Incidental and Planned Exercise Questionnaire (IPEQ-W) will be used. Scores are calculated as the product of the frequency score and duration score to create a total duration for the week score. Higher scores indicate higher levels of activity. Change scores at 6 and 12 months will be calculated as follow-up minus baseline.

- Change in health-related quality of life: The AQoL questionnaire (version AQoL-8D) measures health-related quality of life. Scores range from -0.04 to 1.00 with 1.00 indicating full health-related quality of life. Change scores at 6 and 12 months will be calculated as follow-up minus baseline.

- Global rating of change at 6 and 12 months: Global improvement overall will be scored using a 7-point Likert scale with response options ranging from "much worse" to "much better" when compared to baseline. Participants indicating they are "moderately better" or "much better" will be classified as improved. All other respondents will be classified as not improved. N (%) improved and not improved will be reported for each group.

- Satisfaction with care at 6 and 12 months: Self-reported using a 7-point Likert scale for "Overall satisfaction with the care you received in this study" from "Extremely unsatisfied" to "Extremely satisfied". Participants indicating they are "moderately satisfied" or "extremely satisfied" will be classified as satisfied. All other respondents will be classified as not satisfied. N (%) satisfied and not satisfied will be reported for each group.

- Appointment with an orthopaedic surgeon: Self-reported as yes or no. Participants will be asked if they have visited with an orthopaedic surgeon or have an appointment scheduled for their knee pain in the past 6 months. N (%) answering yes and no will be reported for each group at 6- and 12-months.

- Change in depression, anxiety, and stress: assessed using the Depression, Anxiety, and Stress Scale (DASS-21). Change score at 6 months and 12 months will be calculated as follow-up minus baseline.

- Knee surgery at 6 and 12 months: Self-reported as yes or no. Participants will be asked if they have had a knee arthroscopy and/or arthroplasty since baseline. N (%) answering yes and no for i) surgery (either arthroscopy and/or arthroplasty); ii) arthroscopy alone, and iii) arthroplasty alone will be reported for each group.

- Willingness to undergo surgery: Self-reported. Participants will be asked about their willingness to have surgery on their knee in the near future, ranging from "definitely not willing" to "definitely willing" on a 5-point ordinal scale at baseline, 6, and 12-months. At each time, the n (%) willing (those who answered "probably willing" or "definitely willing") and unsure/not willing (those who answered "unsure", "probably unwilling", or "definitely unwilling") to undergo surgery will be calculated. The proportions in each category at each time point will be reported, as will the numbers and proportions of participants who have made each change from baseline to 6 months and from baseline to 12 months.

- Proportion who meet or exceed the minimal clinically important difference (MCID) for primary outcomes: the proportion of participants in each group whose improvements meet or exceed the MCID for primary outcomes of pain (a change of ≥1.8 units on 11-point NRS) and function (a change of ≥6 units out of 68 on the WOMAC) at i) 6 and ii) 12 months.

## 27. Analysis methods

We will use intention-to-treat analyses, whereby all participants will be included in their randomised treatment groups regardless of their post-randomisation behaviour. Analyses will be conducted by a statistician blinded to treatment group, with two-sided hypothesis tests. If the proportion of participants missing both primary outcomes at 6 months exceeds 5%, missing outcome data will be imputed using multiple imputation methodology, and sensitivity to the missing at random assumption will be investigated (45). Changes from baseline will be presented for each group at each time point using the mean change and 95% confidence intervals.

Outcomes (primary, secondary, and determinants) will be compared between groups as follows: i) *Exercise* vs *Exercise plus weight management*; ii) *Exercise* vs control, and; iii) *Exercise plus weight management* vs control. For continuous data, mean differences in change over time between these groups will be estimated using linear regression models fit to data from both follow-up time points, with random effects for participants, and accounting for clustering by physiotherapist and dietitian in the *Exercise* and *Exercise plus weight management* groups. Models will be adjusted for the stratification variable (history of knee surgery) and values of the outcome at baseline. Terms for time and treatment will be included, with an interaction between the two. Regression assumptions will be assessed using standard diagnostic plots.

For binary outcomes, logistic regression models will be fit using generalised estimating equations to account for clustering, with risk differences and 95% confidence intervals calculated. The stratifying variable of history of knee surgery will be adjusted for.

The outcome of willingness to undergo surgery at 6 months will be analysed using multinomial logistic regression models, adjusting for history of knee surgery, willingness at baseline, treatment, and an interaction between willingness at baseline and treatment. Fixed effects will be included for physiotherapist and dietitian. The outcome of willingness to undergo surgery at 12 months will be analysed similarly, again adjusting for willingness at baseline. For each time point, relative risk ratios for the transition between states of willingness at baseline and follow-up will be presented, for both treatment groups, together with the probabilities of each transition.

To evaluate whether there is a relationship between amount of weight lost at 6 months and changes in pain and/or function at 6 months, linear regression models for changes in the pain and function outcomes at 6 months will be fit and will include a term for weight lost at 6 months (as percentage of baseline weight), adjusting for baseline level of the outcome, assigned treatment group, stratifying variables, and accounting for clustering in the *Exercise* and *Exercise plus weight management* groups. Demographic variables thought to affect both weight loss and outcome levels, including age and sex, will also be included in the model. Fractional polynomial terms for weight lost will be assessed to investigate if the relationship between outcomes and weight loss is non-linear. Standard diagnostic plots will be used to assess regression assumptions.

We will also evaluate whether there is a relationship between adherence to exercise (self-rated on an 11-point NRS) and changes in pain and/or function at 6-months by estimating a complier-average causal effect using a two-stage least squares approach (46).

## 28. Statistical Methods – adjustment for covariates

As described above, analyses will be conducted adjusting for covariates which include baseline values and the stratifying variable. For the regression models that include data from both 6 and 12 months as the outcome, an interaction between month and treatment group will also be included as a covariate.

## 29. Statistical Methods – sensitivity analyses

A sensitivity analysis will estimate treatment effects on the primary outcomes at 6 months assuming full adherence to the assigned treatment (for each arm relative to the control arm separately, where acceptable adherence is defined as attendance at 4 or more of 6 physiotherapy consultations for participants assigned to the *Exercise* arm, and as attendance at 4 or more of 6 dietitian consultations for patients assigned to the *Exercise and weight management* arm). Complier average causal effects will be estimated using an instrumental variables approach (where randomization is the instrument for adherence). Two-stage least squares models will be fit [20] with complier average causal effects reported with 95% confidence intervals and p-values.

## 30. Statistical Methods – subgroup analyses

We will conduct exploratory analyses to investigate whether the relative effects of each intervention on change in each of the primary outcomes at 6-months is moderated by: i) pain self-efficacy; ii) BMI; iii) self-efficacy for eating control. Models will include terms for treatment, the moderator variables, and an interaction between the two. Fractional polynomials will be included for the continuous moderators [88]. The estimated effects of treatment and 95% confidence intervals will be presented for each moderator, with visual representations included for continuous moderators.

The hypothesis, and rationale, for each chosen moderator is below:

- Pain self-efficacy: Participants in the intervention groups who have higher self-efficacy at baseline will report greater improvement in primary outcomes than those who have lower self-efficacy (relative to control group).

- Body mass index: Participants in the *Exercise plus weight management* group who have a higher BMI will report greater improvement in primary outcomes than those with a lower BMI (relative to *Exercise* group and to control group).

- Self-efficacy for eating control: Participants in the *Exercise plus weight management* group who have higher self-efficacy for eating control will report greater improvement in primary outcomes than those with lower self-efficacy for eating control (relative to *Exercise* group and to control group).

## 31. Missing data reporting and assumptions/statistical methods to handle missing data

Baseline characteristics of participants with one or both primary outcomes missing at 6 months will be compared to those of participants with both primary outcomes, as outlined in Section 25. If more than 5% of participants have at least one primary outcome missing at 6 months, multiple imputation will be applied. The number of imputed datasets will be approximately equal to the proportion of participants with missing primary outcomes for the primary outcome with the most missing data. Missing baseline characteristics will be imputed using single mean imputation. Missing outcome values will be imputed separately by treatment group, using chained equations and predictive mean matching, using the five nearest neighbours. Imputation models will include baseline levels of outcomes and baseline characteristics that appear to be different between participants who provide complete follow up data and participants who do not. Initially imputation models for all outcomes will be chained together, with outcomes broken into subsets if imputation models do not converge. Imputed datasets will be compared to complete data using density plots for continuous outcomes and plots of proportions for binary outcomes.

To assess the potential impact of the violation of the missing-at-random assumption on conclusions for the primary outcomes, a pattern-mixture approach (as in White et al (47)) will be applied. We will explore the impact of the violation of the missing-at-random assumption if the assumption was violated in all groups, two groups, or in one group only.

## 32. Additional Analyses

A separate health economic evaluation will be undertaken and reported in a separate manuscript to the primary and secondary trial outcomes. Economic evaluations will assess and compare the cost-effectiveness of the interventions. This will involve analysis of: i) cost per extra person with a clinically significant improvement in pain (measured as 1.8 point reduction on the pain score) and function (6 unadjusted WOMAC units), and; ii) per quality-adjusted life years (QALYs) gained for the intervention, compared to control over 12 months. QALYs will be calculated based on utility scores using the AQoL-8D at baseline, 6 months and 12 months. The difference in health care usage and productivity lost between baseline and 12 months will be calculated and compared for intervention and control groups. AQoL-8D scores at 6 months and 12 months will be estimated using a linear regression model with treatment and interaction dummies adjusting for baseline values and clustering within participants and clinicians. Costs over the 12 months will be calculated using a generalised linear model with an appropriate family and link to account for any skewness, adjusted for baseline values and within participant clustering. Total costs and QALYs over 12 months will be calculated as the area under the curve based on the regression estimates. If more than 5% of participants have costs or AQoL-8D missing at 6 months, multiple imputation will be applied as in section 31. The incremental costs and incremental QALYs (differences between groups) will be calculated as the predicted difference in group means over 12 months and presented as: a) a cost effectiveness ratio and b) the net benefit range when QALYs are multiplied by a willingness to pay per QALY over a range from 0 to $100,000. Ninety-five percent CIs (and p-values for each net benefit) will be calculated using the bootstrap method. The productivity of participants in terms of labour force participation, days absent from work, and self reported productivity at work will also be compared between the intervention and control groups using the WHO Health and Work Performance Questionnaire.

## 33. Harms

The number (and percentage) of patients experiencing adverse events will be presented for each group and the nature of the event(s) described. Adverse events are any undesirable clinical occurrence, whether considered to be treatment-related or not, that includes any problems in either their knee or elsewhere in their body. Adverse events are self-reported in questionnaires at 6 and 12 months. Participants are asked to provide details on the nature of the event, how long it lasted for, and what action they took.

## 34. Statistical Software

Stata v15 or later will be used (StataCorp. Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC)

# References

1. Pereira D, Peleteiro B, Araujo J, Branco J, Santos R, Ramos E. The effect of osteoarthritis definition on prevalence and incidence estimates: a systematic review. Osteoarthritis and Cartilage. 2011;19(11):1270-85.
2. Ackerman IN, Bohenseky MA, Pratt C, Gorelik A, Liew D. Counting the cost: The current and future burden of arthritis. Arthritis Australia; 2016.

3. National Institute for Health and Care Excellence. Osteoarthritis - Care and management, clinical guideline CG177. In: National Institute for Health and Care Excellence, editor. London2014.
4. The Royal Australian College of General Practitioners. Guideline for the management of knee and hip osteoarthritis. 2nd edn ed. East Melbourne, VIC: RACGP; 2018.
5. Runciman WB, Hunt TD, Hannaford NA, Hibbert PD, Westbrook JI, Coiera EW, et al. CareTrack: assessing the appropriateness of health care delivery in Australia. Med J Aust. 2012;197(2):100-5.
6. Haskins R, Henderson JM, Bogduk N. Health professional consultation and use of conservative management strategies in patients with knee or hip osteoarthritis awaiting orthopaedic consultation. Australian Journal of Primary Health. 2014;20(3):305-10.
7. Australian Institute of Health and Welfare. Osteoarthritis snapshot: AIHW; 2018 [cited 2018 24 August]. Cat. no: PHE 232:[Available from: https://www.aihw.gov.au/reports/chronic-musculoskeletal-conditions/osteoarthritis/contents/what-is-osteoarthritis
8. Baker P, Van der Meulen J, Lewsey J, Gregg P. The role of pain and function in determining patient satisfaction after total knee replacement: data from the National Joint Registry for England and Wales. The Journal of Bone and Joint Surgery. 2007;89(7):893-900.
9. Arthritis and Osteoporosis Victoria. A problem worth solving. The rising cost of musculoskeletal conditions in Australia. . Melbourne: Arthritis and Osteoporosis Victoria2015.
10. Fransen M, McConnell S, Harmer AR, Van der Esch M, Simic M, Bennell KL. Exercise for osteoarthritis of the knee. Cochrane Database Syst Rev. 2015;1:CD004376.
11. Bannuru RR, Schmid CH, Kent DM, Vaysbrot EE, Wong JB, McAlindon TE. Comparative effectiveness of pharmacologic interventions for knee osteoarthritis: a systematic review and network meta-analysis. Ann Intern Med. 2015;162(1):46-54.
12. Quicke J, Foster N, Thomas M, Holden M. Is long-term physical activity safe for older adults with knee pain?: a systematic review. Osteoarthritis and cartilage. 2015;23(9):1445-56.
13. Hinman RS, Hunt MA, Creaby MW, Wrigley TV, McManus FJ, Bennell KL. Hip muscle weakness in individuals with medial knee osteoarthritis. Arthritis Care Res (Hoboken). 2010;62(8):1190-3.
14. Runhaar J, Luijsterburg P, Dekker J, Bierma-Zeinstra S. Identifying potential working mechanisms behind the positive effects of exercise therapy on pain and function in osteoarthritis; a systematic review. Osteoarthritis and Cartilage. 2015;23(7):1071-82.
15. Wallis J, Webster K, Levinger P, Taylor N. What proportion of people with hip and knee osteoarthritis meet physical activity guidelines? A systematic review and meta-analysis. Osteoarthritis and Cartilage. 2013;21(11):1648-59.
16. Dunlop DD, Song J, Arntson EK, Semanik PA, Lee J, Chang RW, et al. Sedentary time in US older adults associated with disability in activities of daily living independent of physical activity. Journal of Physical Activity & Health. 2015;12(1):93-101.
17. Lee J, Chang RW, Ehrlich-Jones L, Kwoh CK, Nevitt M, Semanik PA, et al. Sedentary behavior and physical function: objective evidence from the Osteoarthritis Initiative. Arthritis Care Res (Hoboken). 2015;67(3):366-73.
18. White DK, Tudor-Locke C, Zhang Y, Fielding R, LaValley M, Felson DT, et al. Daily walking and the risk of incident functional limitation in knee osteoarthritis: an observational study. Arthritis Care Res (Hoboken). 2014;66(9):1328-36.
19. Ackerman IN, Osborne RH. Obesity and increased burden of hip and knee joint disease in Australia: results from a national survey. BMC Musculoskelet Disord. 2012;13(1):1.
20. Vasilic-Brasnjevic S, Marinkovic J, Vlajinac H, Vasiljevic N, Jakovljevic B, Nikic M, et al. Association of body mass index and waist circumference with severity of knee osteoarthritis. Acta reumatologica portuguesa. 2016;41(3):226-31.
21. Smith WA, Zucker-Levin A, Mihalko WM, Williams M, Loftin M, Gurney JG. Physical Function and Physical Activity in Obese Adults After Total Knee Arthroplasty. Orthopedic Clinics of North America. 2017;48(2):117-25.
22. Brosseau L, Wells GA, Tugwell P, Egan M, Dubouloz C-J, Casimiro L, et al. Ottawa Panel evidence-based clinical practice guidelines for the management of osteoarthritis in adults who are obese or overweight. 2011;91(6):843-61.

23. Alrushud AS, Rushton AB, Kanavaki AM, Greig CA. Effect of physical activity and dietary restriction interventions on weight loss and the musculoskeletal function of overweight and obese older adults with knee osteoarthritis: a systematic review and mixed method data synthesis. BMJ Open. 2017;7(6):e014537.
24. Messier SP, Mihalko SL, Legault C, Miller GD, Nicklas BJ, DeVita P, et al. Effects of intensive diet and exercise on knee joint loads, inflammation, and clinical outcomes among overweight and obese adults with knee osteoarthritis: the IDEA randomized clinical trial. JAMA. 2013;310(12):1263-73.
25. Losina E, Smith KC, Paltiel AD, Collins JE, Suter LG, Hunter DJ, et al. Cost-Effectiveness of Diet and Exercise for Overweight and Obese Knee Osteoarthritis Patients. Arthritis Care Res (Hoboken). 2019;71(7):855-64.
26. Papandony MC, Chou L, Seneviwickrama M, Cicuttini FM, Lasserre K, Teichtahl AJ, et al. Patients' perceived health service needs for osteoarthritis care: A scoping systematic review. Osteoarthritis and Cartilage. 2017;25(7):1010-25.
27. Dobson F, Bennell KL, French SD, Nicolson PJ, Klaasman RN, Holden MA, et al. Barriers and facilitators to exercise participation in people with hip and/or knee osteoarthritis: Synthesis of the literature using behavior change theory. American Journal of Physical Medicine & Rehabilitation. 2016;95(5):372-89.
28. Cottrell MA, Galea OA, O'Leary SP, Hill AJ, Russell TG. Real-time telerehabilitation for the treatment of musculoskeletal conditions is effective and comparable to standard practice: A systematic review and meta-analysis. Clin Rehabil. 2016;31(5):625-38.
29. Bennell KL, Nelligan R, Dobson F, Rini C, Keefe F, Kasza J, et al. Effectiveness of an internet-delivered exercise and pain-coping skills training intervention for persons with chronic knee pain: A randomised trial. Ann Intern Med. 2017;166(7):453-62.
30. Hinman R, Nelligan R, Bennell K, Delany C. "Sounds a bit crazy, but it was almost more personal": A qualitative study of patient and clinician experiences of physical therapist-prescribed exercise for knee osteoarthritis via Skype™. Arthritis Care Res (Hoboken). 2017;69(12):1834-44.
31. Kelly JT, Reidlinger DP, Hoffmann TC, Campbell KL. Telehealth methods to deliver dietary interventions in adults with chronic disease: a systematic review and meta-analysis, 2. The American Journal of Clinical Nutrition. 2016;104(6):1693-702.
32. Fernandes L, Hagen KB, Bijlsma JW, Andreassen O, Christensen P, Conaghan PG, et al. EULAR recommendations for the non-pharmacological core management of hip and knee osteoarthritis. Ann Rheum Dis. 2013;72(7):1125-35.
33. Pearson ML, Mattke S, Shaw R, Ridgely MS, Wiseman SH. Patient Self-Management Support Programs. In: Services USDoHaH, editor. 2007.
34. Martin E, Dobson F, Hall M, Marshall C, Egerton T. The effects of behavioural counselling on the determinants of health behaviour change in adults with chronic musculoskeletal conditions making lifestyle changes: a systematic review and meta-analysis. Musculoskelet. 2019;17(3):170-97.
35. O'Halloran PD, Blackstock F, Shields N, Holland A, Iles R, Kingsley M, et al. Motivational interviewing to increase physical activity in people with chronic health conditions: a systematic review and meta-analysis. 2014;28(12):1159-71.
36. Michie S, Ashford S, Sniehotta FF, Dombrowski SU, Bishop A, French DP. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. Psychol Health. 2011;26(11):1479-98.
37. Bennell KL, Ahamed Y, Jull G, Bryant C, Hunt MA, Forbes AB, et al. Physical therapist–delivered pain coping skills training and exercise for knee osteoarthritis: randomized controlled trial. Arthritis Care Res (Hoboken). 2016;68(5):590-602.
38. Bennell KL, Nelligan RK, Rini C, Keefe FJ, Kasza J, French S, et al. Effects of internet-based pain coping skills training before home exercise for individuals with hip osteoarthritis (HOPE trial): a randomised controlled trial. Pain. 2018;159(9):1833-42.
39. Moerbeek M, Wong WK. Sample size formulae for trials comparing group and individual treatments in a multilevel model. Statistics in Medicine. 2008;27(15):2850-64.
40. Wason JM, Stecher L, Mander AP. Correcting for multiple-testing in multi-arm trials: is it necessary and is it done? Trials. 2014;15(1):1-7.

41. Parker RA, Weir CJ. Non-adjustment for multiple testing in multi-arm trials of distinct treatments: Rationale and justification. Clinical Trials. 2020;17(5):562-6.
42. Wasserstein RL, Schirm AL, Lazar NA. Moving to a world beyond "p< 0.05". Taylor & Francis; 2019.
43. Yelland LN, Sullivan TR, Voysey M, Lee KJ, Cook JA, Forbes AB. Applying the intention-to-treat principle in practice: Guidance on handling randomisation errors. Clinical Trials. 2015;12(4):418-23.
44. Moher D, Hopewell S, Schulz KF, Montori V, Gotzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. Bmj. 2010;340:c869.
45. Carpenter J, Kenward M. Multiple imputation and its application. West Sussex, UK: John Wiley & Sons, Ltd.; 2013.
46. Stuart EA, Perry DF, Le H-N, Ialongo NS. Estimating intervention effects of prevention programs: Accounting for noncompliance. Prevention Science. 2008;9(4):288-98.
47. White IR, Kalaitzaki E, Thompson SG. Allowing for missing outcome data and incomplete uptake of randomised interventions, with application to an Internet-based alcohol trial. Statistics in medicine. 2011;30(27):3192-207.