

# Statistical Analysis Plan (SAP)

## Footwear for Osteoarthritis of the Lateral Knee: the FOLK Trial

### Contents

Section 1. Administrative Information .....	3
1. Title .....	3
2. Trial registration .....	3
3. SAP version.....	3
4. Protocol Version.....	3
5. SAP Revisions .....	3
6. Names and affiliations .....	3
Section 2: Introduction.....	4
7. Background and rationale .....	4
8. Objectives Research hypothesis: .....	4
Section 3: Trial Methods .....	4
9. Trial design .....	4
10. Randomisation .....	5
11. Sample size .....	5
12. Framework .....	5
13. Statistical interim analyses and stopping guidance .....	5
14. Timing of final analysis .....	5
15. Timing of outcome assessments .....	5
Section 4: Statistical Principles .....	5
16. Level of statistical significance .....	5
17. Description of any planned adjustment for multiplicity, and if so, including how the type 1 error is to be controlled.....	5
18. Confidence intervals to be reported .....	6
19. Adherence and protocol deviations .....	6
20. Analysis populations.....	6
Section 5: Trial Population .....	6
21. Screening Data .....	6
22. Eligibility .....	6
23. Recruitment .....	6
24. Withdrawal/follow-up.....	6
25. Baseline characteristics .....	7

Section 6: Analysis.....	7
26. Outcome definitions .....	7
27. Analysis methods.....	8
28. Statistical Methods – adjustment for covariates .....	9
29. Statistical Methods – sensitivity analyses.....	9
30. Statistical Methods – subgroup analyses.....	9
31. Missing data reporting and assumptions/statistical methods to handle missing data .....	9
32. Additional Analyses .....	10
33. Harms.....	10
34. Statistical Software.....	10
References.....	11

## Section 1. Administrative Information

### 1. Title

Footwear for Osteoarthritis of the Lateral Knee: the FOLK Trial

### 2. Trial registration

This trial has been prospectively registered by the Australian New Zealand Clinical Trials Registry on 15/11/2018 (reference: ACTRN12618001864213).

### 3. SAP version

Version: 1.0 Date: 20 July 2021

### 4. Protocol Version

This document has been written based on information contained in the FOLK study protocol version 1.0 dated 23/10/18. The protocol was published as follows:

Paterson, K. L., Bennell, K. L., Metcalf, B. R., Campbell, P. K., Kasza, J., Wrigley, T. V., & Hinman, R. S. (2020). Footwear for osteoarthritis of the lateral knee: protocol for the FOLK randomised controlled trial. *BMC Musculoskeletal Disorders*, 21(1), 247. doi:10.1186/s12891-020-03275-5.

### 5. SAP Revisions

Not applicable

### 6. Names and affiliations

Document prepared by Dr Kade L. Paterson, Ms Fiona McManus, Professor Kim L. Bennell, Mr Ben R. Metcalf, Ms Penny K. Campbell, and Professor Rana S. Hinman, University of Melbourne.

Emails: KL Paterson [kade.paterson@unimelb.edu.au](mailto:kade.paterson@unimelb.edu.au); F McManus [fmcmamus@unimelb.edu.au](mailto:fmcmamus@unimelb.edu.au); KL Bennell [k.bennell@unimelb.edu.au](mailto:k.bennell@unimelb.edu.au); BR Metcalf [b.metcalf@unimelb.edu.au](mailto:b.metcalf@unimelb.edu.au); PK Campbell [penelope.campbell@unimelb.edu.au](mailto:penelope.campbell@unimelb.edu.au); RS Hinman [ranash@unimelb.edu.au](mailto:ranash@unimelb.edu.au)

Signatures:

*Signature of senior statistician responsible (Ms Fiona McManus):*

*Date: 28/6/2021*



*Signature of chief investigator (Dr Kade Paterson):*

*Date: 29/2/2021*



## Section 2: Introduction

### 7. Background and rationale

Osteoarthritis (OA) is the leading cause of musculoskeletal pain and disability in Australia and the knee joint is most often affected. Knee OA is extremely debilitating. Pain is dominant, becoming persistent and more limiting as OA progresses. The medial tibiofemoral (TF) compartment is more frequently affected by OA than the lateral compartment.<sup>1</sup> Nonetheless, structural features of lateral TF joint OA occur in 10-55% of cases of knee OA,<sup>1-5</sup> and research has shown that co-existing lateral TF OA is associated with worse knee pain in people with mixed compartmental OA.<sup>6</sup>

Knee OA is a chronic disease with no cure thus people with OA have little choice but to self-manage their condition. Accordingly, advice about self-management is the cornerstone of conservative treatment, along with exercise and weight control.<sup>7,8</sup> As abnormal biomechanics are central to disease pathogenesis,<sup>9,10</sup> clinical guidelines advocate clinicians provide advice on “appropriate” footwear as part of core treatment for knee OA.<sup>7,11</sup> However, there is scant evidence from clinical trials to guide footwear choice. Unfortunately, all research into footwear for OA has focussed on people with predominantly medial TF OA, and there are no randomised controlled trials (RCTs) evaluating the efficacy of any footwear for people with predominantly TF knee OA. This is a problem given that the biomechanics of people with lateral knee OA differ from those with medial knee OA,<sup>12</sup> and thus any evidence about (in)effectiveness of biomechanical treatments for medial TF OA cannot be directly translated to the lateral compartment.

There is indirect RCT evidence to suggest that “motion control” footwear, which possess medially stiff midsoles and arch support, may be beneficial for people with lateral TF OA. A single small RCT showed improvements in pain and function with medially-wedged foot orthoses compared to neutral insoles in 30 women with predominantly lateral TF OA and bilateral valgus deformity.<sup>13</sup> There have been no RCTs testing the efficacy of motion control shoes on symptoms in people with predominantly lateral TF OA, and there is no evidence to inform clinical guidelines about which type of footwear is best for this important subgroup of patients with knee OA.

### 8. Objectives

#### Research hypothesis:

Primary alternative hypothesis: That motion-control walking shoes will lead to significantly greater reductions in knee pain with walking, compared to neutral walking shoes at 6 months.

Secondary alternative hypothesis: That motion-control walking shoes will have significantly greater benefits on other clinical outcomes (physical function, other measures of knee pain, global ratings of change, health-related quality of life, physical activity levels) compared to neutral walking shoes at 6 months.

#### Study objective:

Primary objective: To determine whether motion-control walking shoes lead to significantly greater reductions in knee pain with walking, compared to neutral walking shoes at 6 months.

Secondary objective: To determine whether motion-control walking shoes will have significantly greater benefits on other clinical outcomes (physical function, other measures of knee pain, global ratings of change, health-related quality of life, physical activity levels) compared to neutral walking shoes at 6 months.

## Section 3: Trial Methods

### 9. Trial design

The FOLK trial is two-arm, superiority, participant- and assessor-blinded RCT. Participants are randomized to either motion-control walking shoes or neutral walking shoes.

## **10. Randomisation**

The randomisation schedule was prepared by the biostatistician (permuted block sizes 6 to 12) stratified by KL grade (2, 3 or 4). Treatment allocation was using a 1:1 ratio. The schedule was stored on a password-protected website (REDCap) maintained by a researcher not involved in either participant recruitment or administration of primary/secondary outcome measures. Group allocation was revealed by this same researcher after baseline primary/secondary outcomes were completed.

## **11. Sample size**

We originally aimed to detect the minimal clinically important difference (MCID) in the primary outcome (change in severity of knee pain on walking (baseline minus follow up)) between groups (1.8 numerical rating scale (NRS) units). We conservatively assumed a between-subject standard deviation of 2.7 and a baseline to 6-month correlation of 0.21 based on previous similar trials.<sup>14,15</sup> Using analysis of covariance (ANCOVA) adjusted for baseline score, we needed 46 per arm to achieve 90% power to detect the MCID in the primary outcome. Allowing for 15% attrition, we aimed to recruit 55 people per arm in total (n=110).

However, due to COVID-19 restrictions in Melbourne (Australia) halting trial recruitment for a prolonged period of time and grant funding running out, the final total sample size was 40. Using ANCOVA adjusted for baseline score, we have 57.8% power to detect the MCID in change in severity of knee pain on walking (baseline minus follow up) with the final sample size of 40 participants (assuming 20 participants per arm).

## **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 on the final total sample size of 40.

## **15. Timing of outcome assessments**

Table 4.6 in the study protocol document details the timing of outcome assessments, the majority of which occur at baseline and at 6 months post-randomisation.

## **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 one primary outcome (change in knee pain with walking over 6 months). We have several secondary outcomes (physical function, other measures of knee pain, global ratings of change, health-related quality of life, physical activity levels). All secondary outcomes are exploratory and not powered for. We will therefore not adjust for multiple secondary 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 footwear (motion-control walking shoes versus neutral walking shoes) for change in knee pain on walking. This approach aligns with the usage of p-values favoured by the American Statistical Association.<sup>16</sup>

## **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 occurred), 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.<sup>17</sup>

Multiple measures of adherence are used in this trial (described in Table 4.6 of the study protocol document) and data from all measures will be reported using means, standard deviations and proportions (number and percentage) as appropriate for each treatment group.

In this study, participants were advised to wear their allocated footwear for at least 6 hours per day. Participants recorded the hours/day they wore their footwear daily for the fourth week of each month in log books. Adherence will be calculated as the average hrs/day spent wearing the study shoes recorded in each of the 6 log books. Participants will be classified as 'adherent' if they wear their footwear for an average of at least 6 hrs/day over 6 months, and all other participants classified as 'non-adherent'.

If a participant does not provide all log books, adherence will be calculated using the available completed log books. If a participant does not provide any log books, non-adherence will be assumed by performing no imputation for the missing adherence variable of average daily wear of allocated footwear.

## **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 created.<sup>18</sup> 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 section 5.2 of the trial protocol document. Reasons for exclusion will be summarized in the CONSORT<sup>18</sup> flow diagram.

## **23. Recruitment**

A CONSORT flow diagram<sup>18</sup> will be used to describe the number of people enrolled, randomised, 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 requests for information 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
- Gender
- Height, body mass, body mass index
- Radiographic disease severity
- Duration of knee symptoms
- Anatomical knee alignment
- Current employment status
- Expectation of treatment outcome- pre and post randomisation
- Arthritis Self Efficacy Scale
- Co-intervention use
- Current footwear characteristics
- Foot posture index
- Foot mobility magnitude
- Navicular drop

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.

If more than one participant is missing the primary outcome at 6 months, 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 primary outcomes at 6 months, and those participants who are missing primary outcomes.

## Section 6: Analysis

### 26. Outcome definitions

#### Primary outcome:

- Change in severity of knee pain on walking: average knee pain intensity on walking in the last week was assessed at baseline and 6 months using an 11-point NRS with terminal descriptors of 'no pain' (score = 0) and 'worst pain possible' (score = 10). Change score at 6 months will be calculated as baseline minus follow-up.

#### Secondary outcomes:

- Change in overall knee pain (KOOS pain subscale): The pain subscale of the KOOS is scored using nine questions regarding knee pain over the previous week, with Likert response options for each question ranging from none/never (score = 0) to extreme/always (score = 4).<sup>19</sup> Scores are then transformed to provide an overall value that ranges from 0 to 100, with 0 representing extreme knee pain and 100 representing no knee pain. Change scores at 6 months will be calculated as baseline minus follow-up.
- Change in physical function (KOOS physical function subscale): The physical function subscale of the KOOS is used to assess limitations with physical functioning.<sup>19</sup> The subscale contains 17 questions on knee function over the past week, with Likert response options from none (score = 0) to extreme (score = 4). Total score ranges from 0 to 100, with lower scores indicating worse function. Change scores at 6 months will be calculated as baseline minus follow-up.

- Change in sport and recreation activities (KOOS sport and recreation subscale): The sport and recreation subscale is assessed using five questions on function during sport and recreational activities over the previous week.<sup>19</sup> Likert responses for each question range from none (score = 0) to extreme (score =4). Scores are then transformed to provide an overall value that ranges from 0 to 100, with 0 representing extreme problems with sport and recreation and 100 representing no problems with sports and recreation. Change scores at 6 months will be calculated as baseline minus follow-up.
- Change in knee-related quality of life (KOOS quality of life subscale): This subscale is assessed using four questions on knee-related quality of life experienced in the previous week.<sup>19</sup> There are five Likert response options for each question, ranging from none/never/not at all (score = 0) to extreme/constantly/extremely (score = 4). Scores are then transformed to provide an overall value that ranges from 0 to 100, with 0 representing extreme problems with quality of life and 100 representing no problems with quality of life. Change scores at 6 months will be calculated as baseline minus follow-up.
- Change in patellofemoral pain and OA (KOOS patellofemoral pain and OA subscale): The patellofemoral pain and OA subscale includes 11 questions on knee pain and function experienced in the last week, each with five Likert response options, ranging from none/never/not at all (score = 0) to extreme/always/totally (score = 4).<sup>19</sup> Scores are then transformed to provide an overall value that ranges from 0 to 100, with 0 representing extreme patellofemoral problems and 100 representing no patellofemoral problems. Change scores at 6 months will be calculated as baseline minus follow-up.
- Global improvement at 6 months: Global improvement in i) pain and ii) physical function will each be scored using a 7-point global rating of change 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.
- Achievement of the MCID in improvement in pain (1.8 NRS units): Improvement in severity of knee pain on walking by 1.8 NRS units (the MCID in pain) from baseline to 6-month follow-up will be classified as improved. All other respondents will be classified as not improved.
- Change in health-related quality of life: The AQoL questionnaire (version AQoL-6D) measures health-related quality of life.<sup>20</sup> This is a 20-item questionnaire and scores range from -0.04 to 1.00 with 1.00 indicating maximum health-related quality of life. Change scores at 6 months will be calculated as baseline minus follow-up.
- Change in physical activity: The Physical Activity Scale for the Elderly was used to assess physical activity over the previous week.<sup>21</sup> This is a 10-item questionnaire which collects responses for the frequency, duration and intensity level of a range of activities typically chosen by older adults. Scores range 0 to >400 with higher scores indicating greater levels of physical activity. Change scores at 6 months will be calculated as baseline minus follow-up.

## 27. Analysis methods

### Primary outcome:

Main comparative analyses between groups will be performed using intention-to-treat. If more than 5% of primary outcomes are missing, multiple imputation will be applied. For the primary hypothesis, differences in mean change in pain (baseline minus follow-up) will be compared between groups using linear regression modelling adjusted for the primary outcome at baseline and the stratifying variable of KL grade (as 3 categories: 2, 3 or 4). Results will be presented as mean differences between groups with 95% confidence intervals, and p-values will also be reported. Complete-case analyses will also be conducted. Standard diagnostic plots will be used to check model assumptions.



### **Secondary outcomes:**

Analyses between groups will be performed using intention-to-treat. For continuous outcomes, analyses will be similar to those for the primary outcomes. Improvement based on global change scores and the achievement of the MCID in improvement in pain (1.8 NRS units) will each be compared between groups separately using logistic regression, adjusting for the stratifying variable of KL grade (as 3 categories: 2, 3 or 4), with results reported as risk ratios and risk differences. Counts and percentages of participants experiencing improvements based on global change scores and of participants achieving the MCID in improvement in pain (1.8 NRS units) will be reported in each treatment group. For all between-group comparisons, 95% confidence intervals for comparisons and p-values will be reported. Standard diagnostic plots will be used to check model assumptions.

### **28. Statistical Methods – adjustment for covariates**

For all outcomes, adjustment is as described in the relevant section (Section 27, 29 and 30).

### **29. Statistical Methods – sensitivity analyses**

A sensitivity analysis will estimate treatment effects on the primary outcome assuming full adherence, where full adherence is as defined in Section 19, adjusted for the outcome at baseline and the stratifying variable of KL grade (as three categories: 2, 3 or 4). That is, complier average causal effects will be estimated using an instrumental variables approach (where randomisation is the instrument for adherence). Two-stage least squares models will be fit<sup>22</sup> with complier average causal effects reported with 95% confidence intervals and p-values.

### **30. Statistical Methods – subgroup analyses**

To assess whether the effect of shoe group on the primary outcome of pain is moderated by KL grade, a linear regression model will be fit for the primary outcome with the outcome at baseline, treatment group, and KL grade (as a binary moderator, Grade 4 versus Grade 2 and 3) as covariates, including an interaction between treatment group and KL grade.

To assess whether the effect of shoe group on the primary outcome of pain is moderated by any of FPI score, knee alignment and KOOS PFJ pain and OA, separate linear regression models will be fit for the primary outcome for each potential moderator with the outcome at baseline, treatment group, the relevant potential moderator and KL grade (as 3 categories, Grade 2, 3 and 4), as covariates, including an interaction between treatment group and the potential moderator

The rationale for the *a priori* choice of treatment effect modifiers is as follows:

- KL grade- we hypothesise that pain reduction with motion-control shoes (relative to neutral) will be greater in those with more severe radiographic disease severity, based on our prior research that showed that people with more severe radiographic disease had greater pain relief with unloading shoes in a sample of people with medial tibiofemoral OA.<sup>23</sup>
- FPI score- we hypothesise that pain reduction with motion-control shoes (relative to neutral) will be moderated by FPI, given the association between rearfoot eversion and medial-to-lateral knee load distribution.<sup>23</sup>
- Knee alignment- we hypothesise that pain reduction with motion-control shoes (relative to neutral) will be greater in those with more valgus knee alignment, as these people are likely to have greater lateral TF knee loading and thus greater scope for improvement with motion-control shoes.<sup>23</sup>
- KOOS patellofemoral pain and OA subscale score- we hypothesise that pain reduction with motion-control shoes (relative to neutral) will be greater in people with more severe concurrent patellofemoral symptoms, given that medially-posted motion-controlling foot orthoses can alleviate patellofemoral pain.<sup>23</sup>

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

If more than 5% of participants have primary outcomes 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. 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<sup>24</sup>) will be applied. We will explore the impact of the violation of the missing-at-random assumption if the assumption was violated in both groups, or in one group only.

### **32. Additional Analyses**

Nil

### **33. Harms**

The number (and percentage) of patients experiencing any adverse events will be presented for each treatment group and the nature of the event(s) described. An adverse event is defined as any problem experienced in the study knee or elsewhere in the body because of wearing the study shoes.

### **34. Statistical Software**

Stata v16.1 will be used (StataCorp. 2020. *Stata Statistical Software: Release 16.1*. College Station, TX: StataCorp LLC)

## References

1. Ledingham J, Regan M, Jones A, Doherty M. Radiographic patterns and associations of osteoarthritis of the knee in patients referred to hospital. *Annals of Rheumatic Disease* 1993; 52: 520-526.
2. Fascingbauer M, Renner L, Waldstein W, Boettner F. Are lateral compartment osteophytes a predictor for lateral cartilage damage in varus osteoarthritic knees?: Data from the Osteoarthritis Initiative. . *Bone Jt J* 2015; 97B: 1634-1639.
3. Cotofana S, Buck R, Wirth W, Roemer F, Duryea J, Nevitt M, et al. Cartilage thickening in early radiographic knee osteoarthritis: a within-person, between-knee comparison. . *Arthritis Care & Research* 2012; 64: 1681-1690.
4. Yamabe E, Ueno T, Miyagi R, Watanabe A, Guenzi C, Yoshioka H. Study of surgical indication for knee arthroplasty by cartilage analysis in three compartments using data from Osteoarthritis Initiative (OAI). . *BMC Musculoskelet Disord* 2013; 14: 194.
5. Nagaosa Y, Lanyon P, Doherty M. Characterisation of size and direction of osteophyte in knee osteoarthritis: a radiographic study. *Ann Rheum Dis* 2002; 61: 319-324.
6. Iijima H, Aoyama T, Nishitani K, Ito H, Fukutani N, Isho T, et al. Coexisting lateral tibiofemoral osteoarthritis is associated with worse knee pain in patients with mild medial osteoarthritis. *Osteoarthritis Cartilage* 2017; 25: 1274-1281.
7. National Clinical Guideline Centre. Osteoarthritis. Care and management in adults. Clinical guideline CG177. Methods, evidence and recommendations. London; National Institute for Health and Care Excellence 2014.
8. McAlindon TE, Bannuru RR, Sullivan MC, Arden NK, Berenbaum F, Bierma-Zeinstra SM, et al. OARSI guidelines for the non-surgical management of knee osteoarthritis. *Osteoarthritis Cartilage* 2014; 22: 363-388.
9. Andriacchi TP, Mundermann A. The role of ambulatory mechanics in the initiation and progression of knee osteoarthritis. *Current Opinion in Rheumatology* 2006; 18: 514-518.
10. Chehab EF, Favre J, Erhart-Hledik JC, Andriacchi TP. Baseline knee adduction and flexion moments during walking are both associated with 5 year cartilage changes in patients with medial knee osteoarthritis. *Osteoarthritis Cartilage* 2014; 22: 1833-1839.
11. 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. *Annals of the Rheumatic Diseases* 2013; 72: 1125-1135.
12. Butler R, Barrios J, Royer T, Davis I. Frontal-plane gait mechanics in people with medial knee osteoarthritis are different from those in people with lateral knee osteoarthritis. *Phys Ther* 2011; 91: 1235-1243.
13. Rodrigues P, Ferreira A, Pereira R, Bonfa E, Borba E, Fuller R. Effectiveness of medial-wedge insole treatment for valgus knee osteoarthritis. *Arthritis & Rheumatism (Arthritis Care & Research)* 2008; 59: 603-608.
14. Hinman RS, Wrigley TV, Metcalf BR, Campbell PK, Paterson KL, Hunter DJ, et al. Unloading Shoes for Self-management of Knee Osteoarthritis: A Randomized Trial. *Ann Intern Med* 2016; 165: 381-389.
15. Bennell KL, Bowles KA, Payne C, Cicuttini F, Williamson E, Forbes A, et al. Lateral wedge insoles for medial knee osteoarthritis: 12 month randomised controlled trial. *BMJ* 2011; 342: d2912.
16. Wasserstein RL, Schirm AL, Lazar NA. Moving to a world beyond “ $p < 0.05$ ”. *The American Statistician* 2019; 73: 1-19.
17. 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. *Clin Trials* 2015; 12: 418-423.
18. 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.
19. Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS)--development of a self-administered outcome measure. *The Journal of orthopaedic and*

- sports physical therapy 1998; 28: 88-96.
20. Osborne RH, Hawthorne G, Lew EA, Gray LC. Quality of life assessment in the community-dwelling elderly: validation of the Assessment of Quality of Life (AQoL) Instrument and comparison with the SF-36. *Journal of Clinical Epidemiology* 2003; 56: 138-147.
  21. Martin K, Rejeski W, Miller M, James M, Ettinger Jr W, Messier S. Validation of the PASE in older adults with knee pain and physical disability. *Medicine & Science in Sports & Exercies* 1999; 31: 627-633.
  22. Maracy M, Dunn G. Estimating dose-response effects in psychological treatment trials: the role of instrumental variables. *Statistical Methods in Medical Research* 2008; 20: 191-215.
  23. Paterson KL, Bennell KL, Metcalf BR, Campbell PK, Kasza J, Wrigley TV, et al. Footwear for osteoarthritis of the lateral knee: protocol for the FOLK randomised controlled trial. *BMC Musculoskeletal Disorders* 2020; 21: 247.
  24. 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. *Stat Med* 2011; 30: 3192-3207.