**Shoe-stiffening inserts for first metatarsophalangeal joint osteoarthritis: a randomised trial**

Shannon E. Munteanu, *PhD*1,2 Karl B. Landorf, *PhD*1,2 Jodie A. McClelland, *PhD*2,3 Edward Roddy, *DM*4,5 Flavia M. Cicuttini, *PhD*6 Alan Shiell, *PhD*7 Maria Auhl, *BSc*1,2 Jamie J. Allan, *MPodPrac*1,2 Andrew K. Buldt, *PhD* 1,2 Hylton B. Menz, *DSc*1,2,4🖂

1Discipline of Podiatry, School of Allied Health, Human Services and Sport, La Trobe University, Melbourne, Victoria 3086, Australia

2La Trobe Sport and Exercise Medicine Research Centre, School of Allied Health, Human Services and Sport, La Trobe University, Melbourne, Victoria 3086, Australia

3Discipline of Physiotherapy, School of Allied Health, Human Services and Sport, La Trobe University, Melbourne, Victoria 3086, Australia

4Primary Care Centre Versus Arthritis, School of Primary, Community and Social Care, Keele University, Keele, Staffordshire, ST5 5BG, United Kingdom

5Haywood Academic Rheumatology Centre, Midlands Partnership NHS Foundation Trust, Haywood Hospital, Burslem, Staffordshire, ST6 7AG, United Kingdom

6Department of Epidemiology and Preventive Medicine, Monash University, Alfred Hospital, Melbourne, Victoria 3004, Australia

7School of Psychology and Public Health, La Trobe University, Melbourne, Victoria 3086, Australia

## 🖂Correspondence: Prof Hylton B Menz, School of Allied Health, Human Services and Sport, La Trobe University, Melbourne, Victoria 3086, Australia. Ph +61-3-9479-5801, Fax: +61-3-9479-5768, Email: h.menz@latrobe.edu.au

## Abstract

*Objective:* To evaluate the efficacy of carbon-fibre shoe-stiffening inserts in individuals with first metatarsophalangeal joint osteoarthritis.

*Design:* This was a randomised, sham-controlled, participant- and assessor-blinded trial. One hundred participants with first metatarsophalangeal joint osteoarthritis received rehabilitation therapy and were randomised to receive either carbon fibre shoe-stiffening inserts (n=49) or sham inserts (n=51). The primary outcome measure was the Foot Health Status Questionnaire (FHSQ) pain domain assessed at 12 weeks.

*Results:* All 100 randomised participants (mean age 57.5 (SD 10.3) years; 55 (55%) women) were included in the analysis of the primary outcome. At the 12 week primary endpoint, there were 13 drop-outs (7 in the sham insert group and 6 in the shoe-stiffening insert group), giving completion rates of 86 and 88%, respectively. Both groups demonstrated improvements in the FHSQ pain domain score at each follow-up period, and there was a significant between-group difference in favour of the shoe-stiffening insert group (adjusted mean difference of 6.66 points, 95% CI 0.65 to 12.67, p=0.030). There were no between-group differences for the secondary outcomes, although global improvement was more common in the shoe-stiffening insert group compared to the sham insert group (61 versus 34%, RR 1.73, 95% CI 1.05 to 2.88, p=0.033; number needed to treat 4, 95% CI 2 to 16).

*Conclusion:* Carbon-fibre shoe-stiffening inserts were more effective at reducing foot pain than sham inserts at 12 weeks. These results support the use of shoe-stiffening inserts for the management of this condition, although due to the uncertainty around the effect on the primary outcome, some individuals may not experience a clinically worthwhile improvement.

Key words: foot, osteoarthritis, orthosis, randomised trial

## Introduction

Osteoarthritis of the first metatarsophalangeal joint (first MTP joint OA) is the most common form of foot OA and affects approximately 8% of people aged 50 years and over.1 The condition is characterised by pain and stiffness in the first MTP joint, which is frequently disabling1 and impairs both foot-specific and general health-related quality of life.2 Radiographic severity of first MTP joint OA is associated with increased pain, deformity and decreased joint range of motion, suggesting that it may be a progressive disorder that has an accumulative impact on the weight-bearing function of the foot.3

Shoe-stiffening inserts are semi-rigid insoles that are commonly recommended as an intervention for first MTP joint OA.4 These inserts are placed inside the shoe with the objective of reducing the rate and magnitude of dorsiflexion at the MTP joints during walking, which is speculated to reduce symptoms by decreasing compression at the dorsal aspect of the joint.5,6 Indeed, there is evidence to suggest that shoe-stiffening inserts can reduce dorsiflexion of the first MTP joint during gait7 and preliminary evidence from a recent case series study that reported clinically worthwhile improvements in foot pain and foot-related disability.8

The primary aim of this trial was to determine if shoe-stiffening inserts are more effective at reducing pain associated with first MTP joint compared to sham shoe inserts. The secondary aims were to determine if shoe-stiffening inserts are: (i) more effective at reducing first MTP joint motion when walking compared to sham shoe inserts, and (ii) more cost-effective compared to sham shoe inserts.

## Methods

### Study design

The trial was prospectively registered with the Australian New Zealand Clinical Trials Registry (ACTRN12616000552482). The La Trobe University Human Ethics Committee provided ethical approval (number HEC15-128) and all participants provided written informed consent prior to enrolment. The full trial protocol has been published.9 The design was a parallel group, participant- and assessor-blinded, randomised controlled superiority trial with a 52 week follow-up. Participants were randomised to the control group (sham shoe inserts) or the intervention group (shoe-stiffening insert(s) [Carbon Fibre Spring Plate, Paris Orthotics, Vancouver, BC, Canada]). Both groups also received a rehabilitation therapy program. Participants were blinded to their group allocation by being informed that they would receive one of two different shoe insert treatments (i.e. the characteristics of the inserts were not disclosed) and a rehabilitation therapy program for their condition. Due to the nature of the intervention, research staff administering the shoe inserts could not be blinded to group allocation. However, research staff prescribing the rehabilitation therapy, assessing outcomes, and entering and analysing data were blinded.

### Participants

Between June 2016 and September 2017 we recruited participants via advertisements placed in local newspapers, by posters placed in senior citizens’ centres and retirement villages, mail-out advertisements to health-care practitioners in Melbourne, mail-out to people currently accessing podiatry services at the La Trobe University Health Sciences Clinic, and through social media.

To be included in the trial, participants had to: (i) be aged at least 18 years, (ii) have pain in the first MTP joint on most days for at least 12 weeks, (iii) have pain rated at least 30 mm on a 100 mm visual analogue scale (VAS), (iv) have pain upon palpation of the dorsal aspect of the first MTP joint, (v) be able to walk household distances (>50 metres) without the aid of a walker, crutches or cane, (vi) be willing to have their foot x-rayed, (vii) agree to attempt to not receive additional interventions (such as shoe modifications, physiotherapy, foot orthoses, intra-articular injections, or surgery) for the first MTP joint pain during the course of the trial, (viii) be able to reach their feet to perform rehabilitation therapy, and (ix) be willing to attempt to discontinue consuming any pain relieving medications for first MTP joint OA (except paracetamol [up to 4 grams per day] as rescue medication) for at least 14 days prior to the baseline assessment and during the trial period. Participants who consumed paracetamol for first MTP joint pain prior to recruitment were advised to discontinue its use at least 24 hours prior to the baseline assessment and follow-up assessments at 4, 12, 24 and 52 weeks.

Exclusion criteria included: (i) previous first MTP joint surgery, (ii) currently pregnant, (iii) significant first MTP joint deformity including hallux valgus (defined as a score of 2 or 3 using the Manchester scale),10,11 (iv) presence of one or more conditions within the foot or ankle that could confound pain and functional assessments of the first MTP joint, such as forefoot pain that is not first MTP joint OA, (v) presence of any systemic inflammatory condition such as gout or rheumatoid arthritis, (vi) any medical condition that, in the opinion of the investigators, made the participant unsuitable for inclusion (e.g. clinically important pain in the musculoskeletal system other than the first MTP joint), (vii) an inability to speak and read English, (viii) cognitive impairment (ix) intra-articular injections (such as corticosteroids) of the first MTP joint in the previous 3 months, (x) unwilling to discontinue use of any foot orthotic devices if currently wearing them, (xi) currently wearing shoe-stiffening inserts, and (xii) regularly wear shoes unable to accommodate the shoe-stiffening inserts.

### Randomisation

Participants were allocated to the intervention or control groups using minimisation12 incorporating stratifications by age (18 to 40, 41 to 60, > 61 years) and sex, using an interactive voice response telephone service provided by the NHMRC Clinical Trials Centre at the University of Sydney, Sydney, Australia.

### Clinical and radiographic assessment

All assessments were performed at the La Trobe University Health Sciences Clinic, Melbourne, Victoria, Australia. Participant characteristics, medical conditions and medications were obtained via structured questionnaire. Participants underwent a clinical assessment including measurements of height, weight and body mass index (BMI), Foot Posture Index,13 passive non-weightbearing dorsiflexion range of motion at the first MTP joint14 and observation to determine the presence of pain on palpation, a dorsal exostosis, pain during motion, a hard-end feel when the joint was fully dorsiflexed, and crepitus.15 Radiographic first MTP joint OA was documented using the La Trobe University atlas.16

### Interventions

Interventions were administered at the La Trobe University Health Sciences Clinic, Melbourne, Australia. Participants allocated to the intervention group were provided with full-length shoe-stiffening inserts (Carbon Fibre Spring Plate, Paris Orthotics Ltd, Vancouver, BC, Canada) at the baseline assessment (a single insert if symptoms were unilateral, or a pair of inserts if symptoms were bilateral). The shoe-stiffening inserts were lightweight (32 to 48 grams across the size range), 1.5 mm thick, and covered with 3.2 mm foam (PPT® 2 809 Blue with an Ultralux top layer, Langer Biomechanics, USA). A full length piece of nylon woven textile covering (Cambrelle®, Camtex Fabrics Ltd, United Kingdom) was applied to the underside of the insert to make its appearence similar to the sham insert.

## Participants allocated to the control group received sham inserts (a single insert if symptoms were unilateral, or a pair of inserts if symptoms were bilateral) designed to not affect first MTP joint dorsiflexion. To achieve this, the distal end of the insert was removed so that the anterior edge finished at the anterior margin of the heel. As with the inserts provided to the intervention group, the sham inserts were covered with a full length layer of 3.2 mm foam, and a full length piece of nylon woven textile covering was applied to the underside of the insert. See Figure 1. Mechanical testing of the sham insert confirmed that it had a negligible effect on shoe bending stiffness.9 The inserts were dispensed by one of the trial investigators (JJA), a registered podiatrist.

Both groups also received standardised rehabilitation therapy.17 This home-based program was performed daily for approximately 30 minutes for 12 weeks, then 3 times per week for the reminder of the trial. Each session involved heat pack application, self-mobilisation of the first MTP joint, toe flexor strengthening exercises and cold pack application. The rehabilitation therapy was reviewed at 1 and 12 weeks. After the 12 week review, participants were telephoned at monthly intervals to remind them to perform the therapy.

Treatment credibility and expectancy were quantified using the Credibility/Expectancy Questionnaire18 which consists of six items; three related to credibility and three related to expectancy. Participants were asked to rate the credibility of the intervention and their expectations on a 9-point Likert scale. High scores on the scale indicate that the participant considers the intervention to be credible and expects it to be effective (maximum score for each is 27).

### Outcome measures

Outcome measures were collected at baseline and at 4, 12, 24 and 52 weeks. To minimise participant burden, postal questionnaires were used for the week 4, 24 and week 52 assessments. To avoid over-testing and to minimise the risk of Type I error associated with serial measurements, statistical analysis of the efficacy of the interventions specifically focused on the change in primary outcome measures between baseline and 12 weeks.19,20 Participants with bilateral symptoms were asked to describe the symptoms of their more painful foot (or right foot if they could not determine the more painful foot).21

The primary outcome measure was the Foot Health Status Questionnaire (FHSQ) pain domain.22 The FHSQ consists of 13 questions that assess foot health in four domains. There are a total of four questions under the pain domain. Questions are scored using a Likert response format and the responses are transformed into a score ranging from 0 to 100 for each domain (with 100 representing optimal foot health).22 Participants were instructed to specifically focus on their big toe joint when answering the questions. The FHSQ has undergone extensive psychometric development22 23 24,25 and has been used previously in clinical trials of interventions for first MTP joint OA.21,26

Secondary outcome measures included: (i) foot-related disability (the FHSQ function domain),22 (ii) severity of pain at the first MTP joint while walking over a flat surface and during rest over the past week, (iii) self-reported magnitude of symptom change (using a 15-point Likert scale where the responses range from “a very great deal better” to “a very great deal worse”). This variable was dichotomised, with ‘effective’ defined as “somewhat better” or above,27 (iv) physical activity (the Incidental and Planned Activity Questionnaire),28 (v) health status (the Short-Form-12 Version 2 questionnaire29 and EuroQol questionnaire [EQ-5D-5L]),30 (vi) use of paracetamol rescue medication and cointerventions, documented with a monthly diary.21,26

### Biomechanical evaluation

Biomechanical evaluation was performed in a sub-sample of participants to evaluate changes in function of the first MTP joint using the inserts.9 Measurement of first MTP joint motion was performed during walking7 using a 10-camera infrared motion analysis system (Vicon Motion Systems Ltd, Oxford, UK). Two force plates were used to identify gait cycle events. Retro-reflective markers were attached to the foot as required for calculation of first MTP joint kinematics.31 Marker trajectory (100 Hz) and force plate (4,000 Hz) data were collected synchronously using Nexus software (Vicon Motion Systems Ltd, Oxford, UK). Comparisons between intervention groups were made for differences with and without the intervention (shoe-stiffening insert or sham insert).

### Adherence

Adherence to the interventions (shoe inserts and rehabilitation therapy) in both groups was assessed at monthly intervals up to 52 weeks via postal survey. For the shoe insert interventions, participants provided information regarding the number of hours per day and number of days that they had worn their inserts during the previous 4 weeks. For the rehabilitation therapy intervention, participants provided information regarding the average number of days per week that they performed their exercises during the previous 4 weeks.

### Adverse events

Adverse events were assessed at monthly intervals up to 52 weeks via postal survey. Participants were asked to document the type of adverse event, the body location, the frequency and/or severity of the effect. Serious adverse events were defined as events that were life-threatening, required hospitalisation, or resulted in persistent or significant disability or incapacity.32

### Sample size

Sample size was determined *a priori* using SPSS Sample Power 3.0 (IBM Corporation, USA). Using a superiority design (one-sided test), power of 90%, minimal important difference of 12.5 points in the foot pain domain of the FHSQ,27 standard deviation of 16.8,21 assuming a 10% drop-out rate, and a significance level set at α < 0.05, we estimated that 90 participants would be required.

### Statistical analysis

Statistical analysis was undertaken using IBM SPSS Statistics version 25.0 (IBM Corp, Armonk, NY, USA) using the intention-to-treat principle.33 Multiple imputation was used to replace missing outcome data using five iterations, with sex, age, baseline scores, and group allocation as predictors. Continuously-scored outcome measures were analysed using analysis of covariance with the intervention group and baseline scores entered as independent variables. Dichotomous-scaled outcome measures were compared using relative risk, risk difference, and number needed to treat (NNT).

### Economic evaluation

An economic analysis was undertaken from a health service perspective. Participants documented their health service use, use of cointerventions, days unable to work and impairments in physical activity in the 4 weeks prior to each survey round. Health service encounters were converted to costs.34 The cost-effectiveness analysis estimated the additional cost per participant who achieved a meaningful improvement in the FHSQ pain score.35 The cost-utility analysis used the EQ-5D-5L to generate health-related quality of life utility scores at baseline and each follow-up, which was used to estimate the difference in quality of life across the two groups. An incremental cost per quality-adjusted life year was then derived. Probabilistic sensitivity analysis was conducted using @Risk software (Palisade, NY, USA).

## Results

### Participant characteristics

Figure 2 shows the flow of participants through the trial. One hundred participants were recruited and randomised (45 men and 55 women, age 24 to 82 years, mean 57.5 [SD 10.3]). Fifty-one participants were allocated to the sham insert group and 49 to the shoe-stiffening insert group. Participants in the two groups had similar baseline characteristics, although the median duration of symptoms differed between the groups (sham insert group: 48 months, shoe-stiffening insert group: 30 months) (Table 1).

### Participant retention

Overall retention rates at each follow-up time point were as follows: 4 weeks (94%), 12 weeks (87%), 24 weeks (80%) and 52 weeks (78%). At the 12 week primary endpoint, there were 13 drop-outs (7 in the sham insert group and 6 in the shoe-stiffening insert group), giving completion rates of 86 and 88%, respectively.

### Treatment credibility and expectancy

Participants considered their treatments to be credible and expected to benefit from them. There was no difference in treatment credibility (sham insert group 22.0 [SD 3.8], shoe-stiffening insert group 22.7 [SD 3.3], *t*=-1.06, *p*=0.294) or expectancy (sham insert group 18.7 [SD 4.1], shoe-stiffening insert group 19.4 [SD 4.8], *t*=-1.06, *p*=0.294) between the two groups (maximum score for each is 27 points).

### Intervention adherence

There was no difference in the total number of hours participants wore their allocated inserts (sham insert group 2,027 [SD 1,102] hours, shoe-stiffening insert group 1,924 [SD 1,224] hours; *t*=0.36, *p*=0.721) or the total number of days participants performed their rehabilitation therapy (sham insert group 181 [SD 65] days, shoe-stiffening insert group 176 [SD 60] days, *t*=0.36, *p*=0.722). However, there was a decrease in insert and rehabilitation therapy adherence in both groups over the 52 week follow-up period (Supplementary file 1).

### Primary outcome

Figure 3 shows the mean (standard error) FHSQ pain domain scores and Table 2 shows the mean (SD) scores and adjusted mean differences (95% CIs) between groups at baseline and at 4, 12, 24 and 52 weeks follow-up. Both groups demonstrated increases in the FHSQ pain domain score (i.e. less pain) at each follow-up time-point. At the 12 week follow-up (the prespecified primary end-point), there was a significant between-group difference in favour of the shoe-stiffening insert group (ANCOVA adjusted mean difference of 6.66 points, 95% CI 0.65 to 12.67, *p*=0.030). Improvements in pain at 12 weeks were maintained at 24 and 52 weeks.

### Secondary outcomes

Table 2 shows the mean (SD) scores and adjusted mean differences (95% CIs) between groups for the secondary outcome measures. There were no between-group differences at the 12 week follow-up for any of these measures. However, the perception of global improvement at 12 weeks, defined as at least moderate improvement (score ≥ 4) on the 15-point Likert scale, was higher in the shoe-stiffening insert group (61 *versus* 34%, RR 1.73, 95% CI 1.05 to 2.88, *p*=0.033). The NNT was 4 (95% CI 2 to 16).

Complete case analysis produced similar results to the imputed dataset for all outcome measures (Supplementary file 2).

### Biomechanical evaluation

Biomechanical evaluation data were available for a sub-sample of 48 participants (23 in the sham insert group and 25 in the shoe-stiffening insert group, mean age 57.8 years, SD 10.8). Compared to the shoe only condition, maximum first MTP joint dorsiflexion during gait was reduced by 0.5 (SD 2.0) degrees with the sham insert and 3.8 (SD 4.1) degrees with the shoe-stiffening insert (adjusted mean difference -3.1, 95% CI -4.8 to -1.4, *p*<0.001). See Figure 4.

### Economic analysis

Mean costs incurred by participants in the sham insert group and shoe-stiffening insert group over the 52 week follow-up period were AU$219 (SD 217) and 270 (SD 111), respectively (cost difference AU$51, 95% CI -18 to 120; *p*=0.144). The additional cost per participant who achieved a meaningful improvement in the FHSQ pain score was AU$86 (95% CI 37 to 135; *p*=0.001). The QALY difference was 0.0039 in favour of the shoe-stiffening insert group. The resulting incremental cost effectiveness ratio was AU$12,980 per QALY gained, with a 55% probability of the insert being cost-effective at willingness-to-pay thresholds greater than $6,500 per QALY gained. The cost effectiveness plane and acceptability curves are provided in Supplementary file 3.

### Use of co-interventions

There were no statistically significant between-group differences in the proportion of participants who used cointerventions (Supplementary file 4).

### Adverse events

There were no serious adverse events. The most common adverse events were pain or soreness within the feet, back or hips. There were no statistically significant between-group differences in the proportion of participants reporting at least one adverse event at any time-point (see Supplementary file 5). However, there was an increased likelihood of foot pain (other than first MTP joint pain) during the first 12 weeks (50 *versus* 28%, RR 1.79, 95% CI 1.05 to 3.03, *p*=0.032) for the shoe-stiffening insert group. Overall, six participants (three from each group) withdrew from the trial due to adverse events (five for foot pain and one for back pain).

## Discussion

The objective of this study was to evaluate the efficacy of carbon-fibre shoe-stiffening inserts in reducing pain in individuals with first MTP joint OA. We found that shoe-stiffening inserts were more effective than sham inserts at reducing joint pain, with the number needed to treat indicating that for every four participants treated with the inserts, one would achieve a successful treatment outcome. Adherence to the shoe-stiffening inserts was very good and no serious adverse events were reported. When viewed from a health service perspective, the shoe-stiffening inserts were considered cost-effective according to the UK National Institute for Health and Care Excellence threshold.36

The primary outcome measure improved in both groups throughout the trial, with the adjusted mean difference favouring the shoe-stiffening insert group at at all follow-up periods. At 12 weeks – our pre-specified primary endpoint – the adjusted mean difference was 7 points, which is less than the recently reported minimal clinically important difference of 11 points35 and had a wide confidence interval. Due to this uncertainty, some individuals may not experience a clinically worthwhile improvement. However, the perception of global improvement at 12 weeks was higher in the shoe-stiffening insert group (61 *versus* 34%, RR 1.73, 95% CI 1.05 to 2.88), and the associated NNT was 4 (95% CI 2 to 16), meaning that for every four participants treated with shoe-stiffening inserts, one would achieve a successful treatment outcome. Of note, the adjusted mean difference was 11 points at the 4 week follow-up, suggesting that the inserts may achieve their peak effectiveness earlier than we expected but then lose some of their effectiveness over time. This could be partly explained by different patterns of adherence between the groups. At the 4 week follow-up, participants reported wearing their shoe-stiffening inserts for an average of 177 hours every 4 weeks, while at the 12 week follow-up, this had reduced to 155 hours every 4 weeks. In contrast, adherence in the sham insert group remained stable (approximately 179 hours every 4 weeks) up to 12 weeks.

The biomechanical sub-study showed that peak first MTP joint dorsiflexion, which occurs just prior to toe-off during the propulsive phase of the gait cycle,37 was significantly reduced in the shoe-stiffening insert group, but not in the sham insert group. This is the intended effect of shoe-stiffening inserts and may decrease the degree of compression at the dorsal aspect of the joint, which is considered to be responsible for joint space narrowing, cartilage erosion and osteophyte formation associated with this condition.3 The relationship between changes in first MTP joint kinematics resulting from inserts and the degree of symptomatic improvement has yet to be explored, but as with mechanical treatments for knee OA (such as lateral wedging38 and bracing39), such an analysis would assist in identifying participants who are most likely to benefit.

Adverse events, although common in both groups, were generally minor. The shoe-stiffening insert group were more likely to report new onset foot pain (other than first MTP joint pain) during the first 12 weeks, and 3 participants withdrew from the trial for this reason. In our pilot study,8 we found that 13% of participants developed new onset foot pain while wearing the shoe-stiffening insert, which we attributed to the stiffness of the inserts leading to an increase in weight-bearing pressure under the foot.40 However, it is also possible that the shoe-stiffening inserts created problems with shoe fit, leading to foot pain in some participants. Future studies should evaluate whether plantar pressures can be further reduced with softer covering materials.

These findings make an important contribution to the limited evidence base for interventions for first MTP joint OA. Our Cochrane review in 20104 found only one randomised trial of physical therapy,17 and only three additional trials of intra-articular hyaluronan,26 foot orthoses and rocker-soled footwear,21 and a synthetic cartilage implant41 have since been undertaken. The improvement we observed in the insert group at 12 weeks (23 points on the FHSQ) is similar to that previously reported with rocker-sole footwear (22 points)21 and greater than prefabricated foot orthoses (17 points)21 or intra-articular hyaluronan (12 points).26 Given that the appearance of rocker-sole footwear limits acceptability and adherence, it could be argued that on the basis of the available evidence, shoe-stiffening inserts should be considered the non-surgical treatment of choice for first MTP joint OA.

Key strengths of this trial include the rigorous design, incorporating participant blinding, a credible sham comparator, adequate statistical power, and excellent retention over a relatively long follow-up period. However, our findings need to be intepreteted in the context of several limitations. First, as is the case with all mechanical intervention/device trials, our control intervention cannot be considered truly inert, so is best considered a sham rather than a placebo.42 We designed the sham insert so it appeared to be a credible treatment, while having as little biomechanical effect as possible. Although our biomechanical analysis confirmed that the sham insert had negligible effects on first MTP joint dorsiflexion, we cannot rule out other biomechanical effects that may have contributed to the improvement observed in the sham group. Second, due to the nature of the interventions, research staff administering the shoe inserts could not be blinded to group allocation. Finally, our participants were volunteers recruited from the general community, so may not necessarily reflect the profile of patients who present with this condition in primary care. However, participant characteristics were very similar to those reported in a population-based study conducted in the UK.3

In conclusion, in individuals with first MTP joint OA, shoe-stiffening inserts are more effective than sham inserts at reducing joint pain, and are cost-effective from a health service perspective. These findings support the use of shoe-stiffening inserts as an effective non-surgical management approach for first MTP joint OA, with the caveat that the minimum important difference in the primary outcome measure was not achieved at the 12 week follow-up and the confidence interval was wide, suggesting that some individuals may not experience a clinically worthwhile improvement. Future studies should investigate methods for improving comfort of the inserts to optimise adherence and potentially extend the duration of their effectiveness.

## Acknowledgements

This study was funded by the National Health and Medical Research Council of Australia (grant ID: 1105244). HBM is currently a National Health and Medical Research Council of Australia Senior Research Fellow (ID: 1135995).

## Author contributions

SEM, HBM, KBL and JAM obtained funding. SEM, HBM, KBL, JAM, ER and FMC designed the study. JAM designed the biomechanical analysis and AS designed the economic evaluation. SEM and HBM coordinated the study. MA, JJA and AKB collected the data. SEM and HBM analysed the data. HBM wrote the first substantial draft of the article. All authors critically revised the manuscript and approved the final manuscript. SEM is guarantor. The corresponding author (SEM) attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

## Competing interests

All authors have completed the ICMJE uniform disclosure form at [www.icmje.org/coi\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) (available in request from the corresponding author) and declare that SEM, KBL, JAM, ER, FMC, AS, MA, JJA, AKB and HBM have no relationships with companies that might have an interest in the submitted work in the previous three years, and SEM, KBL, JAM, ER, FMC, AS, MA, JJA, AKB and HBM have no non-financial interests that may be relevant to the submitted work.

## Role of funding source

The funders did not participate in the study design, data collection, analysis, and interpretation, or the preparation or submission of this report.

## References

1. Roddy E, Thomas MJ, Marshall M, et al. The population prevalence of symptomatic radiographic foot osteoarthritis in community-dwelling older adults: the Clinical Assessment Study of the Foot. Ann Rheum Dis 2015; 74: 156-63.

2. Bergin SM, Munteanu SE, Zammit GV, Nikolopoulos N, Menz HB. Impact of first metatarsophalangeal joint osteoarthritis on health-related quality of life. Arthritis Care Res (Hoboken) 2012; 64(11): 1691-8.

3. Menz HB, Roddy E, Marshall M, et al. Demographic and clinical factors associated with radiographic severity of first metatarsophalangeal joint osteoarthritis: cross-sectional findings from the Clinical Assessment Study of the Foot. Osteoarthritis Cartilage 2015; 23(1): 77-82.

4. Zammit GV, Menz HB, Munteanu SE, Landorf KB, Gilheany MF. Interventions for treating osteoarthritis of the big toe joint. Cochrane Database Syst Rev 2010; (9): CD007809.

5. Nawoczenski D. Nonoperative and operative intervention for hallux rigidus. J Orthop Sports Phys Ther 1999; 29(12): 727-35.

6. Shereff MJ, Baumhauer JF. Hallux rigidus and osteoarthrosis of the first metatarsophalangeal joint. J Bone Joint Surg Am 1998; 80A: 898-908.

7. Rao S, Baumhauer JF, Tome J, Nawoczenski DA. Orthoses Alter In Vivo Segmental Foot Kinematics During Walking in Patients With Midfoot Arthritis. Arch Phys Med Rehabil 2010; 91(4): 608-14.

8. Jerilyn TX, Menz HB, Murley GS, Munteanu SE. Effectiveness of Shoe Stiffening Inserts for First Metatarsophalangeal Joint Osteoarthritis: A Proof-of-Concept Study. Am J Phys Med Rehabil 2016; 95(2): 103-11.

9. Munteanu SE, Landorf KB, McClelland JA, et al. Shoe-stiffening inserts for first metatarsophalangeal joint osteoarthritis (the SIMPLE trial): study protocol for a randomised controlled trial. Trials 2017; 18(1): 198.

10. Garrow AP, Papageorgiou A, Silman AJ, Thomas E, Jayson MI, Macfarlane GJ. The grading of hallux valgus. The Manchester Scale. J Am Podiatr Med Assoc 2001; 91: 74-8.

11. Menz HB, Munteanu SE. Radiographic validation of the Manchester scale for the classification of hallux valgus deformity. Rheumatology 2005; 44: 1061-6.

12. Pocock SJ, Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. Biometrics 1975; 31(1): 103-15.

13. Redmond AC, Crosbie J, Ouvrier RA. Development and validation of a novel rating system for scoring standing foot posture: The Foot Posture Index. Clin Biomech 2006; 21: 89-98.

14. Buell T, Green DR, Risser J. Measurement of the first metatarsophalangeal joint range of motion. J Am Podiatr Med Assoc 1988; 78: 439-48.

15. Zammit GV, Munteanu SE, Menz HB. Development of a diagnostic rule for identifying radiographic osteoarthritis in people with first metatarsophalangeal joint pain. Osteoarthritis Cartilage 2011; 19(8): 939-45.

16. Menz HB, Munteanu SE, Landorf KB, Zammit GV, Cicuttini FM. Radiographic classification of osteoarthritis in commonly affected joints of the foot. Osteoarthritis Cartilage 2007; 15: 1333-8.

17. Shamus J, Shamus E, Gugel RN, Brucker BS, Skaruppa C. The effect of sesamoid mobilization, flexor hallucis strengthening, and gait training on reducing pain and restoring function in individuals with hallux limitus: a clinical trial. J Orthop Sports Phys Ther 2004; 34: 368-76.

18. Devilly GJ, Borkovec TD. Psychometric properties of the credibility/expectancy questionnaire. J Behav Ther Exp Psychiatry 2000; 31: 73-86.

19. Matthews JN, Altman DG, Campbell MJ, Royston P. Analysis of serial measurements in medical research. BMJ 1990; 300: 230-5.

20. Bender R, Lange S. Adjusting for multiple testing--when and how? J Clin Epidemiol 2001; 54: 343-9.

21. Menz HB, Auhl M, Tan JM, Levinger P, Roddy E, Munteanu SE. Effectiveness of Foot Orthoses Versus Rocker-Sole Footwear for First Metatarsophalangeal Joint Osteoarthritis: Randomized Trial. Arthritis Care Res (Hoboken) 2016; 68(5): 581-9.

22. Bennett PJ, Patterson C, Wearing S, Baglioni T. Development and validation of a questionnaire designed to measure foot-health status. J Am Podiatr Med Assoc 1998; 88: 419-28.

23. Menz HB, Auhl M, Ristevski S, Frescos N, Munteanu SE. Comparison of the responsiveness of the Foot Health Status Questionnaire and the Manchester Foot Pain and Disability Index in older people. Health Qual Life Outcomes 2014; 12(1): 158.

24. Martin RL, Irrgang JJ. A survey of self-reported outcome instruments for the foot and ankle. J Orthop Sports Phys Ther 2007; 37: 72-84.

25. Riskowski JL, Hagedorn TJ, Hannan MT. Measures of foot function, foot health, and foot pain. Arthritis Care Res (Hoboken) 2011; 63: S229-39.

26. Munteanu SE, Zammit GV, Menz HB, et al. Effectiveness of intra-articular hyaluronan (Synvisc, hylan G-F 20) for the treatment of first metatarsophalangeal joint osteoarthritis: a randomised placebo-controlled trial. Ann Rheum Dis 2011; 70: 1838-41.

27. Landorf KB, Radford JA, Hudson S. Minimal Important Difference (MID) of two commonly used outcome measures for foot problems. J Foot Ankle Res 2010; 3: 7.

28. Delbaere K, Hauer K, Lord SR. Evaluation of the incidental and planned activity questionnaire for older people. Br J Sports Med 2010; 44: 1029-34.

29. Ware J, Jr., Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996; 34: 220-33.

30. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med 2001; 33(5): 337-43.

31. Nester CJ, Jarvis HL, Jones RK, Bowden PD, Liu A. Movement of the human foot in 100 pain free individuals aged 18-45: implications for understanding normal foot function. J Foot Ankle Res 2014; 7(1): 51.

32. Keech AC, Wonders SM, Cook DI, Gebski VJ. Balancing the outcomes: reporting adverse events. Med J Aust 2004; 181(4): 215-8.

33. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010; 340: c332.

34. Pharmaceutical Benefits Advisory Committee. Manual of Resource Items and their Associated Costs Canberra: Australian Government, 2009.

35. Menz HB, Auhl M, Tan JM, Levinger P, Roddy E, Munteanu SE. Comparative Responsiveness of Outcome Measures for the Assessment of Pain and Function in First Metatarsophalangeal Joint Osteoarthritis. Arthritis Care Res (Hoboken) 2019; 72(5): 679-684.

36. McCabe C, Claxton K, Culyer AJ. The NICE cost-effectiveness threshold: what it is and what that means. PharmacoEconomics 2008; 26(9): 733-44.

37. Carson MC, Harrington ME, Thompson N, O'Connor JJ, Theologis TN. Kinematic analysis of a multi-segment foot model for research and clinical applications: a repeatability analysis. J Biomech 2001; 34(10): 1299-307.

38. Jones RK, Chapman GJ, Forsythe L, Parkes MJ, Felson DT. The Relationship Between Reductions in Knee Loading and Immediate Pain Response Whilst Wearing Lateral Wedged Insoles in Knee Osteoarthritis. J Orthop Res 2014; 32: 1147-54.

39. Moyer RF, Birmingham TB, Bryant DM, Giffin JR, Marriott KA, Leitch KM. Biomechanical effects of valgus knee bracing: a systematic review and meta-analysis. Osteoarthritis Cartilage 2015; 23(2): 178-88.

40. Queen RM, Verma R, Abbey AN, Nunley JA, Butler RJ. Plantar loading during jumping while wearing a rigid carbon graphite footplate. Gait Posture 2014; 39(2): 707-11.

41. Baumhauer JF, Singh D, Glazebrook M, et al. Prospective, Randomized, Multi-centered Clinical Trial Assessing Safety and Efficacy of a Synthetic Cartilage Implant Versus First Metatarsophalangeal Arthrodesis in Advanced Hallux Rigidus. Foot Ankle Int 2016; 37(5): 457-69.

42. Felson DT, Redmond AC, Chapman GJ, et al. Recommendations for the conduct of efficacy trials of treatment devices for osteoarthritis: a report from a working group of the Arthritis Research UK Osteoarthritis and Crystal Diseases Clinical Studies Group. Rheumatology 2016; 55(2): 320-6.

## Figure legends

**Figure 1.** Sham inserts (left) and carbon fibre shoe-stiffening inserts (right) used in the trial. Upper images show superior view, lower images show inferior view.

**Figure 2.** Flow of participants through the trial.

**Figure 3.** Mean (standard error) Foot Health Status Questionnaire (FHSQ) pain scores in the sham insert and shoe-stiffening insert groups. Higher scores represent less pain.

**Figure 4.** First metatarsophalangeal joint dorsiflexion during walking in the sham insert group (top) and shoe-stiffening insert group (bottom).

**Table 1.** Participant characteristics at baseline. Values are mean (SD) unless otherwise noted.

|  |  |  |
| --- | --- | --- |
|  | Sham insert group (n=51) | Shoe-stiffening insert group (n=49) |
| Demographics and anthropometrics |  |  |
| Age – years | 56.1 (9.9) | 59.0 (10.7) |
| Female – n (%) | 29 (56.9) | 26 (53.1) |
| Height – cm | 169.0 (8.9) | 167.5 (7.3) |
| Weight – kg | 79.6 (11.3) | 79.2 (14.6) |
| Body mass index – kg/m2 | 28.0 (4.3) | 28.3 (4.9) |
| General health |  |  |
| SF-12 – physical | 45.9 (9.1) | 46.8 (9.9) |
| SF-12 – mental | 54.0 (9.1) | 53.0 (9.3) |
| Total physical activity – hours / week (IPAQ) | 13.8 (9.9) | 13.1 (11.1) |
| Clinical features |  |  |
| Pain duration – months, median (range) | 48.0 (4 – 432) | 30.0 (3 – 264) |
| Foot Posture Index – mean (SD) [range] | 3.5 (2.6) [-1 – 10] | 4.3 (2.6) [-2 – 9] |
| Maximum 1st MTP joint dorsiflexion – degrees | 44.6 (11.3) | 45.7 (11.2) |
| Pain on palpation – n (%) | 51 (100.0) | 49 (100.0) |
| Palpable dorsal exostosis – n (%) | 51 (100.0) | 48 (98.0) |
| Pain on motion of 1st MTP joint – n (%) | 43 (84.3) | 33 (67.3) |
| Hard-end feel when dorsiflexed – n (%) | 49 (96.1) | 45 (91.8) |
| Crepitus – n (%) | 11 (21.6) | 12 (24.5) |
| Radiographic features – n (%)\*† |  |  |
| Dorsal osteophytes | 47 (95.9) | 45 (97.8) |
| Dorsal joint space narrowing | 44 (89.8) | 42 (91.3) |
| Lateral osteophytes | 46 (93.9) | 42 (93.3) |
| Lateral joint space narrowing | 46 (93.9) | 41 (91.1) |
| Radiographic 1st MTP joint OA‡ | 44 (89.8) | 42 (91.3) |
| SF-12=Short Form 12 Health Survey; IPAQ=International Physical Activity Questionnaire; MTP=metatarsophalangeal; OA=osteoarthritis\* score >0 using La Trobe Radiographic Atlas16† x-ray data not available for 5 participants‡ at least one score of 2 for osteophytes or joint space narrowing from either view, using case definition from La Trobe Radiographic Atlas16 |

**Table 2.** Primary and secondary outcome measures at baseline and follow-up. Values are mean (SD).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Sham insert group (n=51) | Shoe-stiffening insert group (n=49) | Adjusted mean difference (95% CI)\* | *p* |
| FHSQ – pain domain† |  |  |  |  |
| Baseline | 47.8 (17.1) | 47.1 (16.3) |  |  |
| 4 weeks | 56.2 (18.6) | 66.6 (19.8) | 10.79 (4.33, 17.26) |  |
| 12 weeks | 63.8 (17.4) | 70.1 (17.3) | 6.66 (0.65, 12.67) | 0.030 |
| 24 weeks | 64.7 (20.1) | 74.1 (21.1) | 9.59 (2.00, 17.18) |  |
| 52 weeks | 66.4 (23.1) | 73.1 (18.9) | 6.97 (-0.60, 14.53) |  |
| FHSQ – function domain† |  |  |  |  |
| Baseline | 67.8 (23.4) | 64.9 (23.3) |  |  |
| 4 weeks | 76.9 (19.8) | 79.8 (17.2) | 3.71 (-1.62, 9.04) |  |
| 12 weeks | 81.0 (19.2) | 84.9 (17.3) | 4.36 (-2.36, 11.08) | 0.202 |
| 24 weeks | 81.1 (17.5) | 87.9 (16.8) | 7.11 (0.98, 13.24) |  |
| 52 weeks | 82.7 (23.5) | 85.6 (20.1) | 3.31 (-4.58, 11.20) |  |
| Pain severity while walking - VAS‡ |  |  |  |  |
| Baseline | 51.8 (15.6) | 50.8 (14.9) |  |  |
| 4 weeks | 39.0 (21.2) | 31.7 (21.0) | -6.70 (-14.13, 0.72) |  |
| 12 weeks | 33.0 (24.9) | 26.1 (22.9) | -6.49 (1.75, -15.94) | 0.175 |
| 24 weeks | 28.8 (23.3) | 19.2 (19.0) | -9.46 (-18.44, -0.48) |  |
| 52 weeks | 26.1 (23.2) | 19.7 (22.8) | -6.17 (-14.51, 2.16) |  |
| Pain severity at rest - VAS‡§ |  |  |  |  |
| Baseline | 27.3 (19.9) | 29.9 (17.0) |  |  |
| 4 weeks | 22.6 (21.6) | 20.7 (19.1) | -3.23 (-10.7, 4.27) |  |
| 12 weeks | 13.9 (15.0) | 14.5 (16.4) | 0.08 (-5.01, 6.06) | 0.516 |
| 24 weeks | 18.2 (20.7) | 12.8 (19.1) | -5.73 (-12.76, 1.31) |  |
| 52 weeks | 16.8 (20.7) | 14.9 (21.9) | -2.31 (-10.76, 6.14) |  |
| Physical activity - IPAQ§ |  |  |  |  |
| Baseline | 13.8 (9.9) | 13.2 (11.1) |  |  |
| 4 weeks | 15.8 (12.0) | 14.1 (10.6) | -1.31 (-5.11, 2.49) |  |
| 12 weeks | 14.4 (13.5) | 16.3 (15.0) | 2.28 (-2.94, 7.50) | 0.878 |
| 24 weeks | 15.0 (13.9) | 14.5 (11.8) | 0.01 (-3.95, 3.96) |  |
| 52 weeks | 13.8 (11.3) | 15.3 (11.3) | 1.82 (-2.00, 5.63) |  |
| SF12 - physical† |  |  |  |  |
| Baseline | 45.9 (9.1) | 46.8 (9.9) |  |  |
| 4 weeks | 48.2 (8.5) | 48.5 (10.4) | -0.35 (-2.97, 2.27) |  |
| 12 weeks | 46.6 (9.7) | 50.0 (9.2) | 2.74 (-0.34, 5.83) | 0.081 |
| 24 weeks | 49.5 (8.7) | 50.0 (11.6) | -0.12 (-3.45, 3.21) |  |
| 52 weeks | 49.6 (8.6) | 49.5 (9.5) | -0.68 (-3.65, 2.30) |  |
| SF12 - mental† |  |  |  |  |
| Baseline | 54.0 (9.1) | 53.0 (9.3) |  |  |
| 4 weeks | 52.7 (9.1) | 53.3 (9.4) | 1.13 (-1.68, 3.95) |  |
| 12 weeks | 53.2 (9.2) | 53.3 (10.1) | 1.57 (-2.41, 3.80) | 0.658 |
| 24 weeks | 53.9 (8.4) | 52.1 (9.7) | -1.20 (-3.86, 1.46) |  |
| 52 weeks | 53.1 (10.3) | 52.8 (9.9) | 0.37 (-2.86, 3.60) |  |
| EQ-5D-5L VAS† |  |  |  |  |
| Baseline | 76.0 (15.7) | 80.1 (13.7) |  |  |
| 4 weeks | 77.5 (15.5) | 80.8 (13.3) | -0.05 (-0.35, 0.25) |  |
| 12 weeks | 80.1 (11.9) | 82.8 (15.2) | 0.95 (-3.81, 5.71) | 0.692 |
| 24 weeks | 80.1 (15.3) | 81.8 (13.8) | -0.48 (-5.19, 4.24) |  |
| 52 weeks | 80.8 (16.1) | 80.6 (18.5) | -1.76 (-8.42, 4.89) |  |
| FHSQ=Foot Health Status Questionnaire; VAS=visual analog scale; SF-12=Short Form 12 Health Survey; IPAQ=International Physical Activity Questionnaire; EQ-5D-5L VAS=EuroQol 5 dimension-5-level visual analogue scale\* adjusted for baseline score and intervention group using analysis of covariance† higher scores indicate better symptoms/function‡ higher scores indicate worse symptoms§ data skewed so analysis performed on log transformed variable |