RHEUMATOLOGY

Rheumatology 2021;60:2231–2237
doi:10.1093/rheumatology/keaa561
Advance Access publication 16 November 2020

Original article

Individual responses to topical ibuprofen gel or capsaicin cream for painful knee osteoarthritis: a series of n-of-1 trials

Monica S. M. Persson [1,2], Joanne Stocks[1,2,3], Aliya Sarmanova[1], Gwen Fernandes[1,3], David A. Walsh[1,2,3], Michael Doherty[1,2,3] and Weiya Zhang[1,2,3]

Abstract

Objectives. To determine individual responses to ibuprofen gel or capsaicin cream for painful, radiographic knee OA using a series of n-of-1 trials.

Methods. Twenty-two participants were allocated 5% ibuprofen gel (A) and 0.025% capsaicin cream (B) in random sequence (AB or BA). Patients underwent up to 3 treatment cycles, each comprising one treatment for 4 weeks, an individualized washout period (maximum 4 weeks), then the other treatment for 4 weeks. Differential (ibuprofen or capsaicin) response was defined when change-from-baseline pain intensity scores (0–10 NRS) differed by $\geq$1 between treatments in $\geq$2 cycles within a participant.

Results. A total of 104 treatment periods were aggregated. Mean pain reduction was 1.2 (95% CI: 0.5, 1.8) on ibuprofen and 1.6 (95% CI: 0.9, 2.4) on capsaicin ($P = 0.221$). Of 22 participants, 4 (18%) had a greater response to ibuprofen, 9 (41%) to capsaicin, 4 (18%) had similar responses, and 5 (23%) were undetermined.

Conclusion. Irrespective of equal efficacy overall, 59% of people displayed a greater response to one treatment over the other. Patients who do not benefit from one type of topical treatment should be offered to try another, which may be more effective. N-of-1 trials are useful to identify individual response to treatment.

Clinical trial registration. https://clinicaltrials.gov, NCT03146689

Key words: topical NSAID, capsaicin, osteoarthritis, n-of-1, response

Rheumatology key messages

- Response to topical NSAIDs and capsaicin varies between individuals.
- Over half of people with knee OA respond better to one treatment over the other.
- Application of an effective treatment should be tailed to individual responses.

Introduction

Topical NSAIDs and capsaicin are commonly recommended and effective treatments for pain relief in OA [1, 2]. Whilst topical NSAIDs reduce pain primarily through cyclo-oxygenase inhibition [3], capsaicin, the principal warming component of chilli peppers, is thought to act by defunctionalization of spontaneously active nociceptors [4]. Despite reliance on mechanistically disparate methods for pain relief, indirect study-level evidence through network meta-analysis indicates that the treatments are equally effective overall for pain relief in OA [2]. No head-to-head comparison of individual or average responses to the treatments is currently available. However, it is hypothesized that despite the study-level equivalence, treatment efficacy varies between individuals as anecdotally observed in clinical practice.

Evidence synthesis for treatment efficacy has largely focussed on average treatment effects, but in order to improve care it is important to also examine individual responses to treatment. This is the basis of precision

[1]Pain Centre Versus Arthritis, Academic Rheumatology, University of Nottingham, Nottingham, UK, [2]NIHR Nottingham Biomedical Research Centre, Nottingham, UK and [3]Centre for Sport, Exercise and Osteoarthritis Research Versus Arthritis, University of Nottingham, Nottingham, UK

Submitted 11 June 2020; accepted 31 July 2020

Correspondence to: Weiya Zhang, Academic Rheumatology, Clinical Sciences Building, Nottingham City Hospital, Hucknall Road, Nottingham NG5 1PB, UK. E-mail: Weiya.zhang@nottingham.ac.uk

CLINICAL SCIENCE

medicine, and the movement away from a one-size-fits-all approach to individual patient management. N-of-1 trials, or single participant randomized trials, have previously been used to establish the relative efficacy between treatments using within-person comparisons [5, 6]. However, perhaps the greatest benefit of this design is that it allows examination of individual responses to different treatments, and can therefore help optimize treatment at an individual level [7]. The present study aimed to determine individual responses to ibuprofen gel (topical NSAID) or capsaicin cream in knee OA and to explore the use of n-of-1 trials for this purpose.

Method

Study design

This was a randomized, open label series of n-of-1 trials. The study was approved by the Faculty of Medicine and Health Sciences Research Ethics Committee of the University of Nottingham (reference no. B10022017) and registered with ClinicalTrials.gov (NCT03146689). The final approved protocol, participant information sheet, and consent form are available online (https://www.nottingham.ac.uk/research/groups/osteoarthritisandcrystalarthritis/studies/index.aspx).

Participants underwent up to three treatment cycles (six treatment periods; Fig. 1). Each treatment cycle consisted of two treatment periods of ibuprofen gel and capsaicin cream in a randomized order. Treatment periods were 4 weeks and were separated by a washout period until the participant felt their knee pain had returned to its usual pre-treatment level or to a maximum of 4 weeks. An interim analysis was conducted for each participant at the end of the second cycle to determine treatment response. Those who met the pre-specified response criteria (i.e. showed the same response in two cycles) were given the option to complete the study at that point. The n-of-1 trials were aggregated into a series.

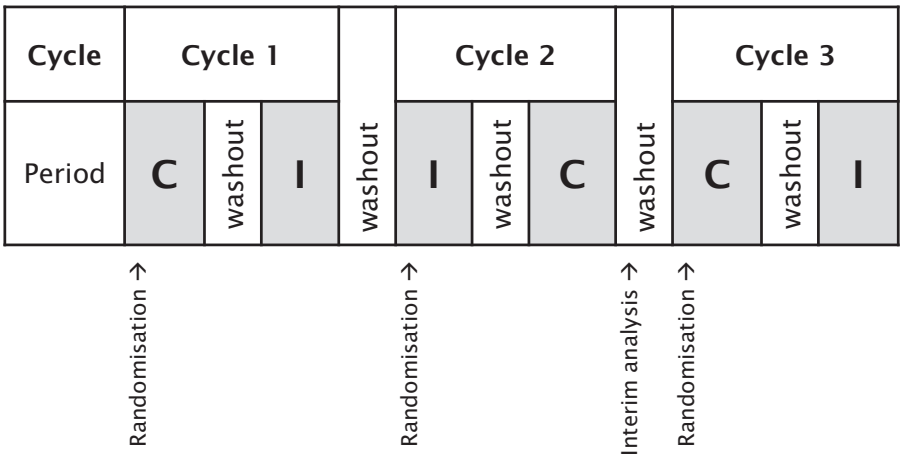
Randomization was conducted using a web-based program (www.randomizer.org) by a researcher not involved in participant recruitment, enrolment, assessment or outcome collection. Randomization data were kept strictly confidential and treatment allocation was recorded in sequentially numbered, sealed, opaque envelopes. Participants were sequentially allocated to the next assigned envelope at the beginning of each treatment cycle. Participants and research staff were not blinded to the treatments due to the initial warming sensation and erythema often experienced with capsaicin and the different appearance and amounts of applied treatment.

Participants

Participants were recruited from the Nottingham Knee Pain and Health in the Community (KPIC) cohort study from the East Midlands region of the UK [8]. Inclusion criteria were men and women aged 40 years and over with chronic knee pain and radiographic knee OA (i.e. definite narrowing and definite osteophyte in the tibiofemoral and/or patellofemoral compartments as per Nottingham line drawing atlas scoring) [9, 10]. Participants scoring between 4 and 8, inclusive, on the 0–10 numeric rating scale (NRS) for knee pain intensity in the index knee were eligible. The most painful knee was determined to be the index knee for local assessments and questionnaire responses.

Exclusion criteria were: inability to give informed consent; terminal or untreated major mental illness; pregnancy or breastfeeding; daily use of oral NSAID in the last 2 weeks; prior regular use of ibuprofen gel or capsaicin cream on the affected knee(s); hypersensitivity or allergy to the interventions or other ingredients in the preparations; total joint replacement of the affected joint;

Fig. 1 N-of-1 trial design showing a hypothetical random sequence for capsaicin (C) and ibuprofen (I)



Treatment periods were 4 weeks' duration and washout periods were variable in length (until pain returned to pre-treatment levels, up to a maximum of 4 weeks).

current treatment for stomach or duodenal ulcers; renal failure; or current treatment with anticoagulants.

Interventions

Participants received 5% w/w ibuprofen gel (Care, Thornton & Ross Ltd) and 0·025% w/w capsaicin cream [Zacin, Cephalon (UK) Ltd]. The medications were applied four times per day to the painful knee(s). The recommended doses were an extruded inch of ibuprofen gel and a pea-sized amount of capsaicin cream. Participants continued to use their regular medications, including oral analgesics, throughout the trial provided the frequency/dose had remained stable for 3 months. Non-permitted concomitant therapies were additional topical analgesics for the affected knee, regular oral NSAIDs, joint injection or surgery.

Outcome measures

Participants recorded pre- and post-treatment pain intensity scores in their index knee for each treatment period using 0–10 NRS (0—'no pain' to 10—'worst imaginable pain'). Change-from-baseline pain scores were calculated per period.

Baseline characteristics assessed prior to randomization were: age, sex and obesity; knee pain intensity (0–10 NRS) and neuropathic-like knee pain (modified painDETECT questionnaire [11]); function (activities of daily living domain of the Knee Injury and OA Outcome Score questionnaire [12]); illness perception (modified Brief Illness Perception Questionnaire [13]); expectation of treatment [13]; anxiety and depression (Hospital Anxiety and Depression Scale [14]); fibromyalgia [15]; central pain mechanism traits [16]; inflammation (knee ultrasound: synovial thickness, effusion and power Doppler signal); abnormal pain processing [quantitative sensory testing: pressure pain thresholds (PPTs) and temporal summation (TS)]; quadriceps muscle strength; and radiographic severity from radiographs obtained at varying time points within the previous 2 years [8].

Sample size

The n-of-1 trials consisted of three treatment cycles (six paired periods). Three cycles is the commonest number of cycles used in n-of-1 trials in OA and other conditions [5, 6, 17]. After the first few cycles, each additional cycle contributes little to the precision of the trial [18]. Three cycles were therefore felt to provide sufficient data without lengthening the trial beyond 44 weeks, thereby maintaining trial eligibility and participant retention.

The sample size of the n-of-1 trial series was subsequently based on the number of *treatment periods* rather than participants. In order to detect a minimum clinically important difference (MCID) of 0.5 (S.D.) between treatments, if there was any, 66 participants were required in a traditional parallel comparison trial, 33 in a cross-over trial, or 11 in a series of n-of-1 trials with three cycles under the assumption of no carry-over and period effects. This would give 80% power at a significance level of 0.05 for the trial. Assuming only 50% of participants showed a differential response between treatments, 22 participants were required.

Statistical methods

Categorical variables were reported as frequencies and continuous variables as mean (S.D.) (if normally distributed) or median and interquartile range (IQR) (if not normally distributed). Statistical significance was set at $P < 0{\cdot}05$.

A difference of 1 point on 0–10 NRS was determined to be the threshold for clinically important pain relief. This is a conservative estimate of the MCID (0.5 S.D.) used by NICE OA guidelines [1] and reflects a 'slightly better' change in status [19].

Each participant's treatment response was determined by comparing the change-from-baseline scores for ibuprofen and capsaicin per cycle. A difference of $\geq$1 point between the treatments indicated a better response for one treatment over the other within the cycle. If a participant displayed the same differential response in two or more cycles, the overall differential treatment response was established for that participant, otherwise they were established as having an equal response. If the participant withdrew prior to meeting the above criteria, their response was 'undetermined'.

The mean treatment effect was determined using a three-level model, clustered at the cycle, period- and participant level. End-of-period pain scores were the dependent variable and adjustment was undertaken for period baseline pain scores. Fixed treatment (ibuprofen = 0 and capsaicin = 1) were assumed. The significance of period and treatment sequence effects was examined through the addition of a treatment-by-period or treatment-by-sequence interaction term (fixed effect). One interaction was examined per model.

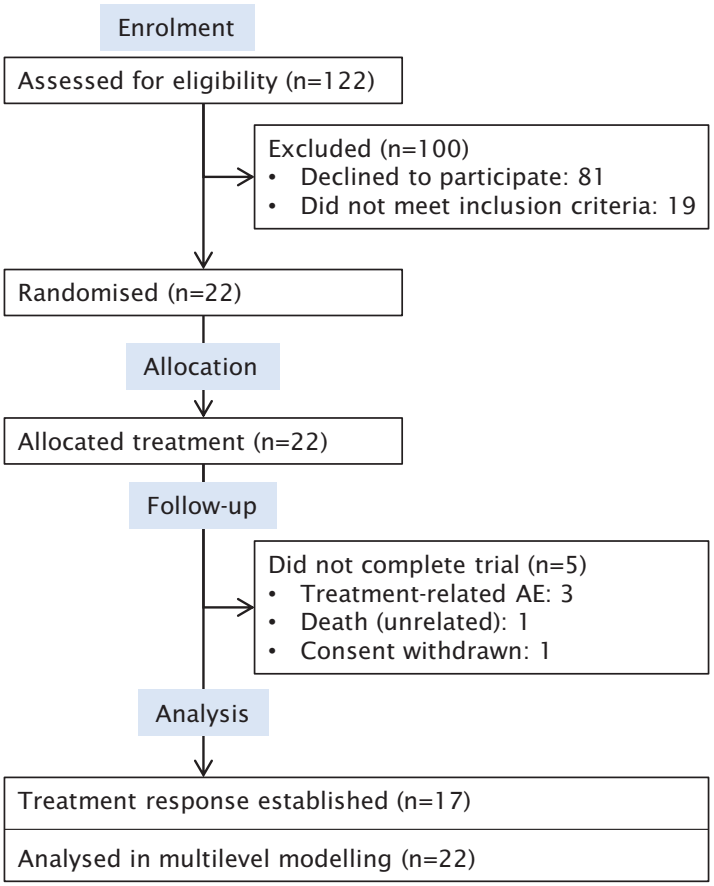
A sensitivity analysis was conducted to exclude participant pain levels remaining outside the pre-specified inclusion criteria (4–8 NRS) after 4 weeks washout. For this, all periods where baseline pain was not within 4–8 points (inclusive) were excluded from the multilevel regression modelling.

Analyses were conducted in Stata Version 15.

Results

Between August 2017 and December 2018, 22 participants were enrolled and completed at least one treatment cycle (Fig. 2). Five participants withdrew before trial completion: 3 due to erythema and skin irritation following ibuprofen gel use, 1 withdrew consent, and 1 died during the third washout period. Cause of death (ischaemic and hypertensive heart disease) was deemed unrelated to the study medications. Baseline characteristics of all participants are presented in Table 1.

Fig. 2 Flow of subjects through the trial



Individual treatment responses were established for the 17 participants who completed the trial. Multilevel modelling for average treatment effects was conducted in all 104 completed treatment periods (22 participants).

Average level of pain reduction for topical ibuprofen *vs* capsaicin

In the 104 completed treatment periods, mean pain reductions were 1.2 (95% CI: 0.5, 1.8) for ibuprofen and 1.6 (95% CI: 0.9, 2.4) for capsaicin (*P*-value for difference = 0.271). No significant period or treatment sequence effects were identified (−0.07; 95% CI: −0.42, 0.28; $P = 0.691$ and 0.15; 95% CI: −0.88, 1.19; $P = 0.773$, respectively).

Pain levels returned to 4–8 points (inclusive) after the washout for 71% of treatment periods. For the periods where pain levels did not return to 4–8 points after washout, 21% were >8 points and 79% were <4 points. No difference was found between topical NSAIDs and capsaicin in the sensitivity analysis limited only to periods where pre-treatment pain was 4–8 NRS ($n = 22$, 80 periods, $P = 0.068$).

Individual responses to treatment

Treatment responses favoured topical ibuprofen in 4 participants (18%), capsaicin in 9 (41%), and found no difference in 4 (18%). Five participants (23%) withdrew prior to a response being established ('undetermined'). Pre-treatment characteristics of patients favouring ibuprofen *vs* capsaicin are presented in Table 1, and some variation in baseline characteristics is seen between the groups.

Discussion

This is the first n-of-1 trial series aiming to identify individual responses to topical NSAIDs and capsaicin. We found that irrespective of the average equivalence between the treatments, more than half of individuals (59%) responded better to one treatment over the other. This suggests that it is feasible to use n-of-1 trials to examine individual responses to treatment.

The best evidence available thus far for the relative efficacy of topical NSAIDs and capsaicin in OA concluded that they provide equal levels of pain relief [2]. However, in order to provide evidence to guide a clinician's question about an individual, there is a need to move away from group averages and towards individual responses [20, 21]. Unlike between-group comparisons, such as randomized controlled trials, n-of-1 trials determine the difference between treatments *within* an individual. They

Table 1 Pre-treatment characteristics of all trial participants and according to their response to treatment

Characteristic	All participants ($n = 22$)	Treatment response	
		Ibuprofen ($n = 4$)	Capsaicin ($n = 9$)
Basic demographics			
Age; mean (S.D.), years	67.0 (9.3)	67.6 (2.2)	63.8 (12.5)
Sex; *n* (%) women	12 (55%)	1 (25%)	7 (78%)
BMI; mean (S.D.), kg/m^2	30.7 (5.6)	27.9 (1.6)	32.3 (7.1)
Questionnaire – comorbidities			
Anxiety; *n* (%) HADS anxiety subscale $\geq$8	9 (40.9%)	1 (25.0%)	4 (44.4%)
Depression; *n* (%) HADS depression subscale $\geq$8	4 (18.2%)	1 (25.0%)	2 (22.2%)
Fibromyalgia; *n* (%) criteria met	2 (9.1%)	0 (0%)	2 (100.0%)
Questionnaire – OA features			
Index knee; *n* (%) left knee pain	11 (50.0%)	2 (50.0%)	5 (55.5%)
Baseline pain; median (IQR), NRS severity	5.5 (4.0 to 7.0)	6.5 (6.0 to 7.5)	5.0 (4.0 to 7.0)
Neuropathic-like pain			
n (%) with definite NP (PDQ $\geq$ 19)	3 (13.6%)	0 (0%)	2 (22.2%)
n (%) with definite or possible NP (PDQ $\geq$ 13)	7 (31.8%)	0 (0%)	5 (55.5%)
Physical function; median (IQR), KOOS	58.1 (47.1 to 73.5)	55.9 (51.5 to 66.2)	58.8 (48.5 to 73.5)
Central mechanisms trait; median (IQR), range (0–24)	9.4 (7.8 to 14.3)	9.0 (7.8 to 11.5)	10.8 (9.0 to 14.3)
Examination findings (index knee)			
Static quadriceps strength; median (IQR), kg	16.1 (14.3 to 21.0)	24.4 (17.6 to 27.9)	16.0 (14.5 to 19.4)
Radiographic severity			
Total NLDA score; median (IQR)	13 (9 to 18)	13 (10 to 17)	14 (6 to 20)
NLDA osteophyte score; median (IQR)[a]	9 (5 to 15)	8 (6 to 11)	9 (3 to 16)
NLDA JSN score; median (IQR)	4 (3 to 5)	5 (4 to 7)	4 (3 to 5)
n (%) per tibiofemoral KL grade			
0	2 (9%)	1 (25%)	0 (0%)
1	3 (14%)	1 (25%)	2 (22%)
2	7 (32%)	2 (50%)	3 (33%)
3	9 (41%)	0 (0%)	3 (33%)
4	1 (5%)	0 (0%)	1 (11%)
n (%) per patellofemoral KL grade			
0	4 (19%)	0 (0%)	4 (44%)
1	9 (43%)	2 (50%)	2 (22%)
2	3 (14%)	0 (0%)	1 (11%)
3	5 (24%)	2 (50%)	2 (22%)
US features			
Synovial thickness; mean (S.D.), mm	5.52 (2.93)	6.48 (1.43)	3.96 (3.16)
n (%) with SH (SH $\geq$ 4mm)	15 (68.2%)	4 (100%)	4 (44.4%)
Effusion; mean (S.D.), mm	8.9 (3.7)	8.4 (3.7)	8.5 (4.0)
n (%) with effusion (effusion $\geq$4 mm)	21 (94.5%)	4 (100%)	8 (88.9%)
n (%) Power Doppler positive	3 (13.6%)	1 (25%)	0 (0%)
QST features			
Localized PPT (MJL); median (IQR), kPa	393.0 (154.4 to 610.6)	424.6 (223.4 to 664.0)	342.7 (151.6 to 610.6)
Distal PPT (proximal tibia); median (IQR), kPa	390.6 (200.0 to 529.6)	390.9 (163.4 to 733.4)	411.8 (200.0 to 518.8)
Remote PPT (sternum); median (IQR), kPa	280.9 (137.7 to 401.2)	282.6 (131.3 to 422.0)	287.2 (118.8 to 430.5)
TS; mean (S.D.), 0–100 NRS	37.9 (24.2)	44.4 (33.8)	34.5 (22.2)

[a]Skyline views were not available for one participant for whom an aggregated osteophyte score was not calculated. Osteophyte scores in the available graded compartment were $\geq$2. HADS, Hospital Anxiety and Depression Scale; IQR, interquartile range; KL, Kellgren-Lawrence; KOOS, Knee Injury and Osteoarthritis Outcome Score; MJL, medial joint line; NLDA, Nottingham Line Drawing Atlas; NP, neuropathic pain; NRS, numeric rating scale; PDQ, painDETECT questionnaire; PPT, pressure pain detection threshold; QST, quantitative sensory testing; SH, synovial hypertrophy; TS, temporal summation; US, ultrasound.

are therefore optimal for precision medicine. Treatments are repeated randomly in order to establish individual treatment responses. Cases can be accumulated gradually to the number sufficient to examine overall treatment responses or for the analysis of predictors of response. Over half of participants in the present study had a clinically important greater response to one treatment over the other, suggesting that this study design is

able to differentiate treatment responses, even when two treatments are equally effective in between-group comparisons.

For precision medicine to be implemented, three conditions need to be met: (i) the disease must be variable due to a multifactorial aetiology; (ii) there must exist several treatment options for which there are heterogeneous responses; and (iii) a clinical biomarker, indicating a differential response for a certain treatment in a patient subpopulation, must be identified [22]. It is widely accepted that OA is a heterogeneous condition and this is reflected in the baseline characteristics of the trial population. Over 50 therapies are available for OA [23] but the present study focussed on two widely available topical therapies. Heterogeneous responses to the topical treatments were identified in the present work. Finally, examination of baseline characteristics according to treatment response indicates that there may be sufficient variation in pre-treatment responses to search for clinical biomarkers. A large and adequately powered n-of-1 trial may be used to establish clinical biomarkers.

The present work is subject to limitations. First, baseline pain levels remained low for some participants despite 4 weeks' washout. This may be due to (i) insufficient washout durations, (ii) the Hawthorne effect, i.e. alterations in behaviour as a result of being observed, (iii) regression to the mean, (iv) natural fluctuations in OA pain, or (v) the effect of potential lifestyle changes during the long observation period. Adjustment for baseline pain allowed this to be accounted for in the comparison of the average pain reduction between treatments. In addition, we conducted a sensitivity analysis restricted to periods where baseline pain scores were 4–8 and found that this did not significantly alter the findings. Second, imbalances in the differential treatment response of the dropouts may have biased the examination of pre-treatment characteristics. Eighty per cent of dropouts displayed greater treatment response to ibuprofen in their first cycle, but were not classified as ibuprofen responders as they withdrew prior to meeting response criteria. However, the purpose of repeatedly comparing treatments in multiple cycles is to establish differential treatment response. Taking only one cycle makes the findings more prone to bias. This is why we chose to classify this group as 'undetermined'. Third, participants and trial personnel were not blinded. Due to inherent difficulties in blinding capsaicin and to more closely reflect clinical care, the trial was open label. However, this reflects clinical practice, where the patient and clinician are not blinded when selecting optimal treatments. Finally, although n-of-1 trials require fewer participants for aggregated analysis and for separation between response and non-response, a larger sample size is needed to compare the characteristics between responders and non-responders. Further study in this regard is useful.

In conclusion, despite topical NSAIDs and capsaicin being equally effective across the whole population, treatment responses varied between individuals with painful knee OA. Over half of participants showed a greater response to one treatment over the other. Where one topical treatment provides insufficient pain relief, clinicians should advise their patient to try a different topical alternative as patients may have a better response.

Acknowledgements

Daniel F. McWilliams is acknowledged for conducting the trial randomization. M.S.M.P. was involved in the conceptualization and design of the work, participant recruitment and screening, study coordination, data collection and analysis, and writing the manuscript. J.S. participated in the conceptualization and design of the work, data collection, and revised the paper. A.S. and G.S.F. participated in study recruitment, data collection, and revised the paper. D.A.W. and M.D. participated in the conceptualization and design of the work, interpretation of the data, and writing the manuscript. W.Z. participated in the conceptualization and design of the work, development of the statistical analysis plan, interpretation of the data, and writing the manuscript. W.Z. is the guarantor. All authors discussed the results, commented on the manuscript, and have approved the final version of the paper. The de-identified data for trial participants can be requested from the corresponding author. Baseline characteristics and follow-up pain data will be available upon request.

Funding: The work was supported by the Nottingham University Hospitals Charity [grant number 2009–01/08] and *Vs* Arthritis [grant number 20777]. The funders had no role in the study design, data collection, data synthesis, data interpretation, or writing the report.

Disclosure statement: M.S.M.P. declares a grant from the Nottingham University Charity for the conduct of the reported work; J.S. reports grants from *Vs* Arthritis and Nottingham University Charity during the conduct of the study; D.A.W. reports grants from *Vs* Arthritis and Nottingham University Charity during the conduct of the study; and other from AbbVie Ltd, Pfizer Ltd, Eli Lilly & Co Ltd, GlaxoSmithKline Medscape Education (New York), Love Productions (UK) outside the submitted work; M.D. reports grants from *Vs* Arthritis and from Nottingham University Hospital Charitable Trust during the conduct of the study; W.Z. reports grants from *Vs* Arthritis and Nottingham University Hospital Charity during the conduct of the study; and personal fees from Grunenthal and Regeneron Inc. outside the submitted work. The other authors have declared no conflicts of interest.

References

1 National Institute for Health and Care Excellence (NICE). Osteoarthritis: Care and management in adults. Clinical guideline [CG177]. Published 12 February 2014. https://www.nice.org.uk/guidance/indevelopment/gid-ng10127.

2 Persson MSM, Stocks J, Walsh DA, Doherty M, Zhang W. The relative efficacy of topical non-steroidal anti-inflammatory drugs and capsaicin in osteoarthritis: a network meta-analysis of randomised controlled trials. Osteoarthritis Cartilage 2018;26:1575–82.

3 Hagen M, Baker M. Skin penetration and tissue permeation after topical administration of diclofenac. Curr Med Res Opin 2017;33:1623–34.

4 Anand P, Bley K. Topical capsaicin for pain management: therapeutic potential and mechanisms of action of the new high-concentration capsaicin 8% patch. Br J Anaesth 2011;107:490–502.

5 March L, Irwig L, Schwarz J *et al.* n of 1 trials comparing a non-steroidal anti-inflammatory drug with paracetamol in osteoarthritis. BMJ 1994;309:1041–6.

6 Yelland MJ, Nikles CJ, McNairn N *et al.* Celecoxib compared with sustained-release paracetamol for osteoarthritis: a series of n-of-1 trials. Rheumatology 2007;46:135–40.

7 Lillie EO, Patay B, Diamant J *et al.* The n-of-1 clinical trial: the ultimate strategy for individualizing medicine? Per Med 2011;8:161–73.

8 Fernandes GS, Sarmanova A, Warner S *et al.* Knee pain and related health in the community study (KPIC): a cohort study protocol. BMC Musculoskelet Disord 2017;18:404.

9 Nagaosa Y, Mateus M, Hassan B, Lanyon P, Doherty M. Development of a logically devised line drawing atlas for grading of knee osteoarthritis. Ann Rheum Dis 2000;59:587–95.

10 Wilkinson CE, Carr AJ, Doherty M. Does increasing the grades of the knee osteoarthritis line drawing atlas alter its clinimetric properties? Ann Rheum Dis 2005;64:1467–73.

11 Freynhagen R, Baron R, Gockel U, Tölle TR. PainDETECT: a new screening questionnaire to identify neuropathic components in patients with back pain. Curr Med Res Opin2006;22:1911–20.

12 Roos EM, Lohmander LS. The Knee injury and Osteoarthritis Outcome Score (KOOS): from joint injury to osteoarthritis. Health Qual Life Outcomes 2003;1:64.

13 Broadbent E, Petrie KJ, Main J, Weinman J. The brief illness perception questionnaire. J Psychosom Res 2006;60:631–7.

14 Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983;67:361–70.

15 Wolfe F, Clauw DJ, Fitzcharles M-A *et al.* Revisions to the 2010/2011 fibromyalgia diagnostic criteria. Semin Arthritis Rheum 2016;46:319–29.

16 Akin-Akinyosoye K, Frowd N, Marshall L *et al.* Traits associated with central pain augmentation in the Knee Pain In the Community (KPIC) cohort. Pain 2018;159:1035–44.

17 Gabler NB, Duan N, Vohra S, Kravitz RL. N-of-1 trials in the medical literature: a systematic review. Med Care 2011;49:761–8.

18 Zucker DR, Ruthazer R, Schmid CH. Individual (N-of-1) trials can be combined to give population comparative treatment effect estimates: methodologic considerations. J Clin Epidemiol 2010;63:1312–23.

19 Salaffi F, Stancati A, Silvestri CA, Ciapetti A, Grassi W. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. Eur J Pain 2004;8:283–91.

20 Hogben L, Sim M. The self-controlled and self-recorded clinical trial for low-grade morbidity. Br J Prev Soc Med 1953;7:163–79.

21 Schork NJ. Personalized medicine: time for one-person trials. Nature 2015;520:609–11.

22 Trusheim MR, Berndt ER, Douglas FL. Stratified medicine: strategic and economic implications of combining drugs and clinical biomarkers. Nat Rev Drug Discov 2007;6:287–93.

23 Zou K, Wong J, Abdullah N *et al.* Examination of overall treatment effect and the proportion attributable to contextual effect in osteoarthritis: meta-analysis of randomised controlled trials. Ann Rheum Dis 2016;75:1964–70.