

Title: Duration of treatment effect should be considered in the design and interpretation of clinical trials: results of a discrete choice experiment

Authors: B Copsey ^{a*}, J Buchanan ^b, R Fitzpatrick ^b, SE Lamb ^a, SJ Dutton ^a, JA Cook ^a.

Affiliations:

^a Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences (NDORMS), University of Oxford, Botnar Research Centre, Windmill Road, Headington, Oxford, OX3 7LD.

^b Nuffield Department of Population Health, University of Oxford, Richard Doll Building, Old Road Campus, Oxford, OX3 7LF.

*Corresponding author (BC): Email: bethan.copsey@csm.ox.ac.uk, tel: +44 (0)1865 737923.

Email addresses for co-authors: james.buchanan@dph.ox.ac.uk, raymond.fitzpatrick@dph.ox.ac.uk, sarah.lamb@ndorms.ox.ac.uk, susan.dutton@csm.ox.ac.uk, jonathan.cook@ndorms.ox.ac.uk.

Funding:

Financial support for this study was provided entirely by a doctoral studentship grant with the EPSRC (Engineering and Physical Sciences Research Council) and Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences via the Medical Sciences Division of the University of Oxford. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report. Sarah E Lamb, Ray Fitzpatrick and Bethan Copsey also receive funding from the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care Oxford at Oxford Health NHS Foundation Trust.

1 Jonathan Cook is funded by the National Institute for Health Research (NIHR) Oxford Biomedical
2 Research Centre. The views expressed are those of the authors and not necessarily those of the NHS,
3 the NIHR or the Department of Health.

4

5

6

Abstract (274/275 words)

Objective

This study examined whether the duration of treatment effect should be considered in a benefit-risk assessment, using a case study of osteoarthritis medications.

Study design and setting

A discrete choice experiment was completed by 300 residents of the United Kingdom with hip and/or knee osteoarthritis. In 16 choice tasks, participants selected their preferred option from 2 medications. Medications were described in terms of the effect on pain, stiffness, and function, duration of treatment effect, and risk of heart attack and stomach ulcer bleeding. Analysis used mixed-effects logistic regression.

Results

Pain, severity and duration of treatment effect had the greatest influence on medication preferences, whereas stiffness did not significantly affect medication choice. Participants were willing to accept an increase in the risk of heart attack of 2.6% (95% CI: 2.0% to 3.2%) to increase duration of the treatment effect from one month to 12 months. Reducing pain from moderate to mild was valued the same as increasing duration of effect from one month to three months; both were seen as equivalent to an absolute reduction of 1.2% in the risk of heart attack in the next year. Subgroup analysis suggested disease severity influenced patient preferences.

1 Conclusions

2 Along with treatment benefits and risks, the results suggest that duration of treatment effect is an
3 important factor in the medication choices of people with osteoarthritis. This could have
4 implications for the design and interpretation of clinical trials, for example, incorporating longer-
5 term surveillance of trials participants and account for duration of treatment effect in risk-benefit
6 assessments. Future research is needed to assess whether these findings are generalisable to other
7 samples, disease areas and levels of duration of effect.

8 **Keywords:** Osteoarthritis, WOMAC, patient preference, duration, treatment effect,
9 discrete choice experiment.

10

11 Manuscript word count: 4699 words

12

13

14

Introduction

Researchers and clinical guidelines have emphasised the importance of long-term follow-up of patients in clinical studies, to assess safety and the sustainability of treatment effects over time [1-4]. However, the benefit-risk assessment of treatments often considers the treatment effect at a single time point. In reality, treatment effects are often not sustained indefinitely; a treatment may provide large initial benefits that diminish over time [4, 5]. Patients' preferences (i.e. which medication they would choose) may differ depending on how long the treatment effect is sustained. A patient could be willing to take a medication that provides reduced symptom relief if the medication is likely to provide benefits for a longer period of time. However, it is unclear how the duration of treatment benefit affects treatment preferences. If duration of treatment effect is important, this could have implications for the design of clinical trials and the interpretation of their results. As well as implications for medical research, preferences on duration could also inform decision-making and impact on the implementation of treatments in clinical practice.

The existing literature on the impact of duration of treatment effect on medication preferences is limited. Providers are aware that it is desirable for medication effects to be quick, strong and long-lasting [6]. Clinicians have suggested that adoption of pharmacological treatments is more likely for medications that were tested in trials of longer follow-up duration [7]. A second study found that patients with psoriasis expressed a strong negative preference for treatments with short duration of beneficial effects [8], and a meta-analysis found that patients would accept a much higher risk of adverse events to increase duration of effect compared to reducing symptoms [9]. However, this is only one example from a single condition. Overall, there is no consensus on the strength of preferences for duration relative to other aspects of treatment.

1

2 A discrete choice experiment (DCE) could be used to quantify patient preferences for the duration of
3 treatment effect and explore how this impacts on treatment decision-making. A DCE is a patient
4 preference study where participants select their preferred alternative from multiple options in
5 hypothetical scenarios. In the context of treatment preferences, a DCE could provide insight into
6 whether duration of treatment effect is important to patients, relative to the benefits and risks
7 associated with the treatment.

8

9 Patient preference studies have previously explored the importance of time by examining the trade-
10 off between treatment effect and survival time or time until the patient is cured. However, few
11 studies have examined the importance of the duration of treatment effect from a course of
12 medication. Among osteoarthritis patients, Posnett et al. found that duration of pain relief was an
13 important factor for treatment preferences [10]. However, the duration of pain relief was based on
14 one dose of oral medication (8 or 12 hours), rather than whether the treatment effect was sustained
15 in long-term use of the medication. Long-term assessment of treatment outcomes in clinical practice
16 and medical research is recommended for chronic conditions, such as arthritis. Randomised trials
17 often find that short-term treatment effects are not sustained over time and do not always translate
18 into long-term benefits [5, 11]. Therefore, the duration of treatment effect could vary between
19 medications and potentially influence the decision of which medication should be prescribed to a
20 patient.

21

22 This study uses a case study of osteoarthritis medications to explore this issue. NSAIDs (non-steroidal
23 anti-inflammatory drugs) can provide long-lasting relief for osteoarthritis symptoms [12]. However,
24 NSAIDs also have known risks of side effects and serious adverse events, including cardiovascular

and gastrointestinal problems [13-17]. As such, these medications are a good exemplar to explore preferences for treatment duration since they are used to treat a long-term condition where sustained symptom relief is desirable but treatment effects may diminish over time. A previous DCE by Hauber *et al.* explored the relative importance of the benefits and risks of medications to treat osteoarthritis symptoms whilst considering the treatment effect at a fixed time point [18]. This study found that reductions in the risk of adverse events, pain levels and function were associated with medication choice. However, this DCE only considered the effectiveness of treatment one hour after taking the medication and did not consider any long-term benefit or decline in treatment effect.

The discrete choice experiment presented in this paper aims to build on the study by Hauber *et al.* by evaluating the importance of the duration of the treatment effect, relative to the level of treatment benefit and potential risks of treatment, for medications targeting relief of osteoarthritis symptoms.

Methods

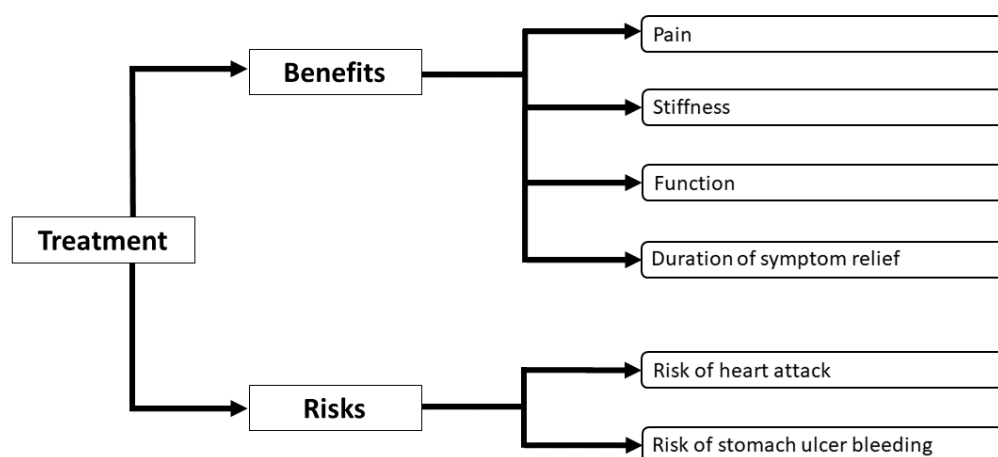
The methods section includes information on attribute and level selection, experimental design, data collection and statistical analyses, aligning with the practice guidelines developed by Bridges *et al.* [19].

Development of the DCE questionnaire

A long-list of potential attributes to be included in the discrete choice experiment was identified by searching the literature on existing patient preference studies in people living with osteoarthritis.

Research has suggested that discrete choice experiments should include no more than 6-8 attributes [20, 21]. The included attributes were selected based on the results of rating and ranking exercises performed by 10 patient representatives, with consideration of the existing literature on patient preferences and the clinical relevance of potential attributes (Appendix 1). The attributes identified broadly covered: treatment effectiveness, duration of treatment effect and risk of serious events (Figure 1).

Figure 1. Attributes included in the discrete choice experiment



Treatment effectiveness was characterised using the three domains of the WOMAC index (Western Ontario and McMaster Universities Osteoarthritis Index): pain, stiffness and function. The WOMAC Index is a commonly used outcome measure in lower-limb osteoarthritis made up of 24 items across three domains [22]. For the pain, stiffness and function attributes, the levels used in this study were: none, mild, moderate, or severe. These levels matched the Likert scale set of responses used in the WOMAC Index. For the duration of treatment effect, levels were selected based on commonly used follow-up time points in randomised trials of NSAIDs for osteoarthritis [23-26].

1

2 The results of the rating and ranking exercises indicated that patients placed similar importance on
3 the risk of serious cardiovascular and gastrointestinal events. However, clinical literature suggested
4 that cardiovascular risk varies between different NSAID medications [27, 28] and there is a high
5 incidence of gastrointestinal events in NSAID users compared to the general population [29, 30].
6 Therefore, risk of heart attack and risk of stomach ulcer bleeding were included as attributes. Levels
7 for the risk attributes were selected based on existing literature on risks associated with the general
8 population, osteoarthritis population and NSAID users (Table 1) [14, 17, 29-34].

9

10

1

2 Table 1. Summary of attributes and levels

Attribute	Levels
Pain	None (0 out of 100mm)
	Mild (25 out of 100mm)
	Moderate (50 out of 100mm)
	Severe (75 out of 100mm)
Functional difficulties	None (0 out of 100mm)
	Mild (25 out of 100mm)
	Moderate (50 out of 100mm)
	Severe (75 out of 100mm)
Stiffness	Stiffness None (0 out of 100mm)
	Mild (25 out of 100mm)
	Moderate (50 out of 100mm)
	Severe (75 out of 100mm)
Duration of symptom relief	1 month
	3 months
	6 months
	12 months
Risk of heart attack in the next year	0% (0 out of 100 people)
	0.5% (1/2 out of 100 people i.e. 1 out of 200 people)
	1% (1 out of 100 people)
	2% (2 out of 100 people)
Risk of stomach bleed in the next year	0% (0 out of 100 people)
	1% (1 out of 100 people)
	2% (2 out of 100 people)
	4% (4 out of 100 people)

3











4

Experimental design and data collection

Each choice task asked participants to choose between two hypothetical medications (A and B). An example choice task is shown in Figure 2. Attributes and levels were described and participants completed a practice choice task before beginning the main questionnaire. The practice choice task included a dominated alternative, where one medication was the same or better than the other medication for all attributes. Each participant completed the same set of 16 choice tasks in an identical order. This has been found to be a feasible number of choice tasks, including for older populations, accounting for cognitive burden and participant fatigue [19, 35, 36]. A duplicate choice task was also included (4th and 17th tasks) to test the consistency of participant responses. After the final choice task, participants rated the overall difficulty of choosing between two medications and considering a hypothetical patient scenario when completing the choice tasks. The questionnaire was piloted in a group of 30 participants to evaluate participant burden and elicit prior values for the coefficients.

1

2 Figure 2. Example choice task

	Medication A	Medication B
The level of pain	Moderate (50 / 100) 	Mild (25 / 100) 
The level of stiffness	Severe (75 / 100) 	None (0 / 100) 
The level of difficulty doing daily activities	Moderate (50 / 100) 	None (0 / 100) 
The length of the symptom relief	1 month	6 months
The risk of heart attack in the next year	0% (0 out of every 100 people) 	2% (2 out of every 100 people) 
The risk of stomach ulcer bleed in the next year	1% (1 out of every 100 people) 	2% (2 out of every 100 people) 

3

4

5

1
2 Ngene software was used to select a fractional factorial design to optimise d-efficiency whilst
3 excluding any tasks with dominated alternatives, unbalanced levels or level overlap [20, 37, 38].
4 Fixed prior values for the coefficients were used, based on the results from pilot data. A model-
5 averaging approach was used where two models were equally weighted; in model 1, coefficients
6 from the pilot results were used for all priors and, in model 2, priors were assigned to be coefficients
7 from the pilot results if they had logically consistent direction and zero otherwise.

8
9 Participants were recruited through an online panel using market research company, ResearchNow.
10 Eligible participants were adults (aged 18 years or older) who resided in the United Kingdom, had a
11 self-reported diagnosis of hip and/or knee osteoarthritis and provided informed consent.
12 Participants received point-based rewards for completing the survey, which could be redeemed for
13 online shopping vouchers. Ethical approval was obtained from the University of Oxford Central
14 University Research Ethics Committee [reference R55785/RE002].

15
16 A sample size of 300 participants was used. The majority of discrete choice experiments include 100-
17 300 participants and a sample size of 300 has been recommended as a rule of thumb for main
18 effects analysis of discrete choice experiments [39-41].

19
20 Data were collected on the age, gender, work status, disease severity, prior treatments and co-
21 morbidities of participants. Disease severity was measured using the WOMAC Index Likert scale
22 version 3.1 [42]. Single item questions examined participant's self-assessed level of optimism and
23 attitude to risk-taking.

Statistical analysis

For descriptive information on the study sample, continuous variables were summarised using the mean and standard deviation, and categorical variables were described using the frequency and proportions of responses. No missing data was anticipated because participants were required to respond to all questions of the online survey.

The responses to the choice questions were used to estimate models of choice behaviour. Attributes of pain, stiffness, function and duration of symptom relief were included in these models as effects-coded variables. The base levels were severe for the WOMAC domain attributes and one month for duration. Risk attributes for heart attack and stomach ulcer bleeding were included as continuous variables, assuming a linear form. Two models were estimated. First, a fixed effects model was estimated using conditional logistic regression analysis, as this is the workhorse model of DCE analysis [38]. Second, a mixed effects logistic model was estimated to accommodate respondent heterogeneity, with all terms initially included as random effects. Those with significant standard deviation ($p < 0.05$) indicating between-participant heterogeneity were included as random-effects terms in the final model. The random effect terms were assumed to be correlated. All other levels were included as fixed effects variables. The mixed effects analysis was conducted in Stata IC 15, using the `mixlogit` command [43].

Coefficients were presented with the corresponding 95% confidence interval. A positive coefficient represents an improvement in utility, with a higher coefficient indicating a greater probability of choosing the medication. The rate at which respondents are willing to trade-off between each attribute and risk of heart attack is presented. Confidence intervals for base levels (severe or 1

month) were calculated using the delta method. Statistical significance was defined as a two-sided p-value less than 0.05.

Subgroup and sensitivity analyses

Subgroup analyses were considered exploratory due to the anticipated low level of precision. Pre-specified subgroup analyses were performed using interaction terms between key demographics (age, gender and WOMAC score) and all attribute levels. Age and baseline WOMAC total score were included as continuous linear terms, centred on the mean. Additional exploratory subgroup analyses examined interactions between (a) prior stomach bleed and risk of stomach bleed, (b) prior heart attack with risk of heart attack, (c) risk-taking attitude with risk of stomach bleed and risk of heart attack, and (d) time since diagnosis with duration of symptom relief. Time since diagnosis was included as two dummy variables, one for three to five years and one for six or more years since diagnosis (base case response of less than three years since diagnosis or 'don't know'). Interaction terms for each baseline covariate were initially included in separate models. The significant terms were combined into a single model with multiple interactions and non-significant interaction terms were sequentially dropped.

Sensitivity analysis was performed to test the robustness of the results based on the included sample of participants [44-46]. Sensitivity analyses were conducted excluding:

- (i) Participants who selected a dominated alternative in the practice choice task,
- (ii) Participants who selected different medications when the same choice task was presented twice at different points in the questionnaire,

- (iii) Participants who found it 'quite difficult' or 'very difficult' to choose between medications in the choice tasks, and
- (iv) Participants who found it 'quite difficult' or 'very difficult' to imagine a hypothetical patient scenario when completing the choice tasks.

Results

Sample characteristics

Of the 342 responders who were eligible and consented to participate, 300 participants completed the survey and were included in the analysis (88%, $n=300/342$). Participants had a mean age of 60 years and there were more females ($n=164/300$, 55%) than males. Around half of the participants were retired. The majority of participants had moderate levels of symptoms on the WOMAC pain, function and stiffness sub-scales (Table 2), and had experienced osteoarthritis symptoms for over three years ($n=242/300$, 81%). Many participants had experienced previous cardiovascular issues with 54% having high blood pressure ($n=162/300$). Most participants had received medication and physiotherapy to treat their osteoarthritis symptoms. Around a quarter of participants had undergone joint replacement surgery ($n=83/300$, 28%). The majority of participants had good mental health status, with high levels of optimism (Table 3).

1

2 Table 2. Participant demographics: Key characteristics (n=300)

	Mean (SD) or n	Range or %
Age (years)	60.5 (SD 13.3)	23-92
Sex:		
Male	136	45%
Female	164	55%
Work status:		
Working full-time	76	25%
Working part-time	38	13%
Retired	141	47%
Other	45	15%
Osteoarthritis sites: (participants could select multiple options)		
Shoulder	111	37%
Elbow	65	22%
Hip	186	62%
Hand or wrist	158	53%
Knee	241	80%
Foot or ankle	97	32%
Prior injury of knee or hip before OA*	95	32%
Time since OA diagnosis*:		
Less than a year	14	5%
1 - 2 years	39	13%
3 - 5 years	79	26%
6 - 10 years	70	23%
Over 10 years	93	31%
Don't know	5	2%
General health:		
Excellent	24	8%
Very good	59	20%
Good	97	32%
Fair	74	25%
Poor	46	15%
WOMAC scores*:		
WOMAC pain	8.1 (SD 4.7)	0-20
WOMAC stiffness	3.5 (SD 1.9)	0-8
WOMAC function	27.4 (SD 16.6)	0-68
WOMAC total	39.0 (SD 22.3)	0-96

3 * OA: osteoarthritis; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index

4

1

2 Table 3. Participant demographics: Previous treatments, co-morbidities and mental health

3 (n=300)

	n	%
Medications: (participants could select multiple options)		
Warfarin	53	18%
Glucocorticoid	87	29%
Co-morbidities: (participants could select multiple options)		
Stomach bleed	42	14%
Peptic ulcer disease	45	15%
Stroke	43	14%
Heart attack	40	13%
High blood pressure	162	54%
Other heart problems	66	22%
Prior treatment: (participants could select multiple options)		
Medication (tablet)	229	76%
Medication (cream or gel)	156	52%
Injection	96	32%
Physiotherapy	155	52%
Exercise	152	51%
Joint replacement surgery	83	28%
Time during the past four weeks felt calm and peaceful		
None of the time	21	7%
A little of the time	73	24%
Some of the time	72	24%
Good bit of the time	50	17%
Most of the time	66	22%
All of the time	18	6%
Time during the past four weeks felt downhearted and blue		
None of the time	60	20%
A little of the time	79	26%
Some of the time	59	20%
Good bit of the time	43	14%
Most of the time	33	11%
All of the time	26	9%
Level of optimism:		

Very optimistic	64	21%
Quite optimistic	121	40%
Neither optimistic, nor pessimistic	66	22%
Quite pessimistic	25	8%
Very pessimistic	17	6%
Don't know	7	2%

Risk-taking attitude:

Risk loving	42	14%
Risk neutral	109	36%
Risk averse	146	49%
Don't know	3	1%

1

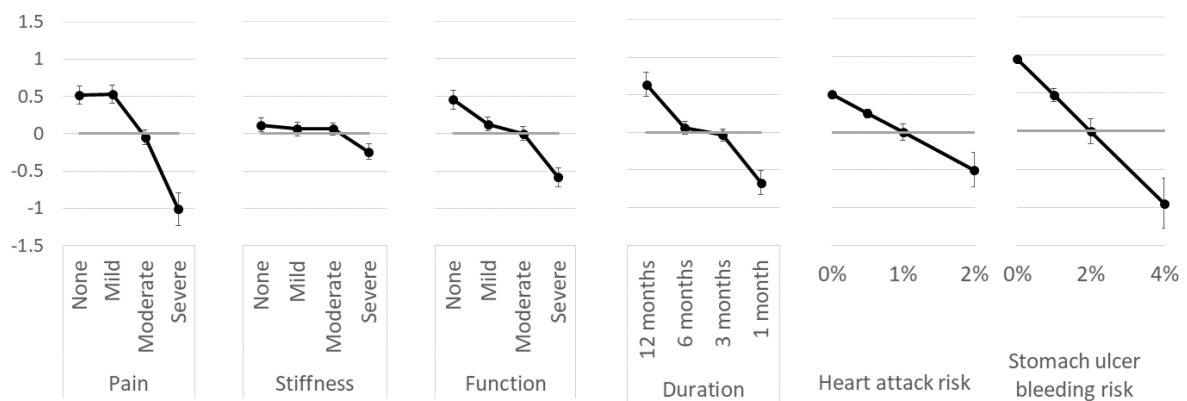
2

Results of mixed effects model

Fixed-effects and mixed-effects models produced very similar results and predicted the same medication choice for all choice tasks and participants. Therefore, due to the presence of preference heterogeneity, only the mixed-effects model is presented in this article.

The relative importance of the different attributes is shown in Figure 3. All attributes significantly impacted on treatment choice except stiffness. Duration of overall treatment effect, effect on pain symptoms, and risk of stomach ulcer bleeding were seen as the most important, with effect on physical function and risk of heart attack seen as important to a lesser extent.

Figure 3. Results of mixed effects model: coefficients



1 Comparing benefits and risks associated with the treatment, participants were willing to accept an
2 increase in the risk of heart attack of 3% to increase the duration of the treatment effect from one
3 month to 12 months. To reduce symptoms from 'severe' to 'none', respondents would be willing to
4 accept an increased risk of heart attack of 2% for improvements in function, and 3% for
5 improvements in pain symptoms (Table 4).

6

Table 4. Results of mixed effects model: Willingness to increase risk of heart attack[#]

		Maximum risk increase willing to accept (%)		
		Willingness to risk	LCI*	UCI*
Pain	Severe to moderate	1.9	1.4	2.5
	Moderate to mild	1.2	0.8	1.6
	Mild to none	0.0	-0.3	0.2
	Total: Severe to none	3.1	2.4	3.7
Stiffness	Severe to moderate	0.6	0.3	0.9
	Moderate to mild	0.0	-0.3	0.3
	Mild to none	0.1	-0.2	0.4
	Total: Severe to none	0.7	0.4	1.0
Functional difficulties	Severe to moderate	1.2	0.8	1.6
	Moderate to mild	0.3	0.0	0.6
	Mild to none	0.7	0.3	1.0
	Total: Severe to none	2.1	1.6	2.6
Duration	1 month to 3 months	1.3	0.9	1.7
	3 months to 6 months	0.2	-0.1	0.4
	6 months to 12 months	1.2	0.8	1.5
	Total: 1 month to 12 months	2.6	2.0	3.2

*LCI: Lower confidence interval, UCI: Upper confidence interval.

[#] Coefficients were translated into willingness to risk of heart attack due to subgroup differences in the risk of stomach ulcer bleeding due to disease severity

For the duration of the treatment effect, the initial increase from one to three months was seen as important, after which duration must be extended to 12 months to be important. For treatment benefit, reducing pain levels to 'mild' was highly important, however respondents did not differentiate between 'mild pain' and 'no pain'. Reducing functional difficulties from 'severe' to 'moderate' was also important, but after this point, functional problems must be reduced to 'no difficulty' to have a large effect on treatment choice.

Random effects terms were included for pain (none and mild), duration (12 months), stiffness (mild), risk of heart attack and risk of stomach ulcer bleeding. Significant random effects indicated preference heterogeneity between respondents on the importance of large improvements in pain and function, increasing the length of the treatment effect to 12 months and risks of cardiovascular and gastrointestinal events (Appendix 2).

The risk of experiencing a serious cardiovascular or gastrointestinal event was considered to be important. The effect of increasing the incidence risk of a serious event by 1% was very similar for both risk of heart attack and risk of stomach ulcer bleeding (the marginal rate of substitution was close to one). The risk of stomach ulcer bleeding was seen as more important overall due to the high risk incidence (the maximum risk was 4% for stomach ulcer bleeding, compared to 2% for heart attack risk).

Predictive ability and reliability of the model

The main model without interaction terms had an area-under-the-curve value of 0.68 (95% CI 0.67 to 0.70). Coefficients exhibited a logical direction in almost all cases and otherwise had very small differences between adjacent levels. For each individual choice task, the predicted medication choice was the same for all participants. For 15 of the 16 choice tasks, the model predicted that participants would choose the most highly-selected medication. Sensitivity analyses did not produce any substantive differences in the model coefficients or any changes to the predicted medication choice.

When performing the choice tasks, a third of participants found it 'quite difficult' to choose between medications in the choice tasks (n=100, 33%) but few found it 'very difficult' (n=10, 3%). A quarter of participants reported that it was difficult to consider an imaginary patient scenario, rather than their own situation (59 (20%) found it 'quite difficult' and 12 (4%) found it 'very difficult'). When presented with a duplicate task later in the survey (tasks 4 and 17 being identical), around a fifth of respondents selected a different medication (n=55, 18%). During a practice choice task, a relatively small proportion of respondents demonstrated irrational preferences, by selecting a dominated alternative (n=20, 7%).

Exploratory subgroup analyses

Subgroup analyses using the final combined model indicated that a significant interaction term was only present between the baseline WOMAC score covariate and risk of stomach ulcer bleeding (Appendix 3a). This suggested that participants with a higher baseline WOMAC score (i.e. more severe disease) placed less importance on the risk of stomach ulcer bleeding (Appendix 3b).

Discussion

Summary of findings

This study aimed to examine whether duration of treatment effect is an important factor for patients when choosing between different osteoarthritis medications, and to quantify the level of importance relative to treatment benefits and risks. The results suggest that the duration of the treatment effect is an important factor in treatment decision-making for osteoarthritis patients. The overall importance of duration was similar to the importance of improvements in pain levels and function and risks of serious cardiovascular and gastrointestinal events. However, improvements in stiffness were seen as relatively unimportant. The results suggest that participants would be willing to accept a medication which is less effective in relieving their pain symptoms if the effect of the medication was longer-lasting. For example, participants viewed a medication reducing pain from 'severe' to 'moderate' for 3 months as equivalent to a medication reducing pain from 'severe' to 'mild' for 1 month.

The importance of moving between different levels of the WOMAC domains was not consistent across the levels. For example, improving pain levels to 'mild' compared to 'moderate' was important but reduction to 'no pain' was seen as equivalent to 'mild pain'. There was preference heterogeneity on the importance of large improvements in pain and duration and risks of serious events. Heterogeneity in the importance of risk of stomach ulcer bleeding was partially explained by differences in disease severity.

Comparison with existing literature

1 The age and gender profile of participants in this study was very similar to the UK CPRD sample of
2 early OA and late-stage OA requiring joint replacement [47, 48]. The pain levels of the sample were
3 also similar to the baseline scores in clinical trials of NSAIDs for osteoarthritis, with moderate disease
4 symptoms on pain, function and stiffness on the WOMAC scale [23, 26, 49]. The DCE sample was
5 also representative in terms of the sites of OA occurrence, with many participants having OA at
6 multiple sites with knee OA being most common, followed by hip OA then hand OA [50-53].

7
8 Overall, the results on the importance of symptom relief and risks of serious events are in line with
9 the results of a previous discrete choice experiment on osteoarthritis medication by Hauber et al.
10 [18]. The discrete choice experiment by Hauber et al. was based on symptom relief one hour after
11 taking osteoarthritis medication and assumed that the treatment effect did not degrade over time.
12 Both studies found that stiffness was relatively unimportant. A key difference was that the results of
13 this study suggested that the importance of improvements from 'moderate' to 'mild' were more
14 important than 'mild' to 'none'. This conflicts with the findings of Hauber and colleagues. This
15 discrepancy could be due to differences in patient characteristics of the sample, variations in how
16 participants interpreted the levels of the WOMAC domains or differences in how participants
17 interpret changes in symptoms over a longer time period.

18
19 Few patient preference studies have examined the importance of the duration of treatment effect.
20 The studies that have examined the duration of treatment effect in osteoarthritis patients have
21 found it to be important, supporting the results of our study. For example, Cordero-Ampuero *et al.*
22 found lasting symptom relief was important to patients, and Posnett *et al.* found that duration of
23 relief was more important than the amount of pain reduction from injections [10, 54].

Strengths and limitations

This is one of the first discrete choice experiments to examine the importance of the duration of treatment effect in decision-making. This study has demonstrated that duration can be included in patient preference studies, along with benefits and risks of treatment, without causing prohibitive increases in cognitive difficulty and participant burden. The results have implications for the design and interpretation of clinical trials, which could improve the relevance of such research to osteoarthritis patients.

However, this study has only examined the importance of duration in terms of choosing between osteoarthritis medications. The findings are also limited in that one in five participants gave inconsistent responses to two identical choice tasks, raising concerns about the robustness of the results; however sensitivity analysis found that the findings were still similar when those providing inconsistent responses were excluded. In addition, the proportion selecting medication B was 58% in the first display and 56% in the repetition so the lack of consistency could be due to uncertainty for this particular choice task.

A further limitation is that the relatively small sample size did not allow us to fully explore subgroups and interactions. Clearly there is some heterogeneity in these respondent preferences, and we have explored this by incorporating random effects parameters and conducting subgroup analyses to explore the association between patient characteristics and the importance of different attributes. For example, subgroup analyses examined whether having a previous serious cardiovascular or gastrointestinal event, or the participants' level of risk aversion, or was associated with the importance of risk attributes. However, there may be other factors (measured or unmeasured)

1 which could explain heterogeneity in patient responses. We have provided the dataset as an
2 appendix in case interested readers would like to further explore the interaction between the
3 characteristics and preferences of respondents. In addition, the large number of attributes and
4 levels provided limited scope to explore interactions between different attributes. Reliance on the
5 self-reported diagnosis of osteoarthritis is a key limitation that could have introduced additional
6 heterogeneity in the participant sample and that prevents exploration of differences in radiographic
7 measures of disease severity.

8
9 The levels for the duration attribute were chosen to correspond to commonly used assessment time
10 points in randomised trials. However, the majority of the study sample had experienced
11 osteoarthritis symptoms for more than 3 years. Therefore, a further limitation is that we are unable
12 to quantify preferences for interventions with very long term effects in clinical practice in this study.
13 For similar reasons, preferences for effects measured in hours, such as those arising from a single
14 dose of medication, also cannot be estimated. Only a limited number of attributes and attribute
15 levels can be included in a discrete choice experiment without impacting upon completion rates.
16 Therefore, we were unable to fully encompass all potential benefits and risks potentially associated
17 with all medications, such as effects on co-morbid conditions.

18
19 Discrete choice experiments are limited in that participants only make stated preferences. We
20 cannot know whether they would make the same choice if forced to act on their decision. The
21 design is also limited because the included attributes are somewhat reductive. For example, in this
22 experiment, the effect on pain is assumed to be constant over time, when in reality it is likely to
23 fluctuate with flare-ups and subsiding periods. The interpretation of the different levels may also
24 differ across participants, for instance, different participants may vary in what they would define as
25 'mild pain' due to the subjectivity of pain measurement and varying pain thresholds.

1

2 *Implications*

3

4 The importance of the duration of treatment effect suggests that phase III clinical trials of
5 medications for osteoarthritis should include follow-up longer than 6 months to measure long-term
6 outcomes and to establish whether a significant treatment benefit is sustained over time. When
7 interpreting the results of clinical studies, the duration of the treatment effect should be taken into
8 account during risk-benefit assessments and clinical decision-making. In addition, trialists should
9 consider incorporating longer-term surveillance of trial participants. Extended-trial follow-up would
10 potentially allow the exploration of treatment effects over a longer period and better consideration
11 of rare adverse events [55]. Future research is needed to assess whether these findings are
12 generalisable to other samples, disease areas and levels of duration of effect.

13

14 Since the duration of treatment effect was nearly as important as pain and as important as function,
15 duration of treatment benefit should be carefully considered when comparing trials with different
16 assessment time points. Trialists should also consider the optimal time point for patient outcome
17 measurement when designing future trials. The time point of assessment may impact on the
18 patients' view of what is an 'important difference' in the treatment outcome. This could have
19 implications for the sample size required for a randomised trial. Future research is needed to identify
20 the optimal assessment time points for different patient groups and patterns of recovery.

21

22 Primary outcomes in clinical studies should focus on pain and/or function. Trials using composite
23 outcome measures should unpick which domains are generating differences in treatment effect. The
24 WOMAC index incorporates pain, stiffness and function, however the results of this study indicate

1 that improvements in the stiffness domain were not valued by participants. Therefore, when
2 reporting the results of clinical trials, authors should report the treatment effect on the individual
3 sub-scale for each of the 3 WOMAC domains, as well as the overall WOMAC score. This would help
4 readers to understand which domain is driving a significant treatment effect and ensure that the
5 treatment will target the mechanisms that are of value to each individual patient.

8 *Future research*

10 Future research could evaluate the preferences of clinicians and commissioners to establish how
11 they value the duration of treatment effect, relative to the benefits and risks of osteoarthritis
12 medications. A larger patient sample would increase the power of subgroup analyses, increasing the
13 potential to explain reasons for preference heterogeneity. Qualitative methods could also
14 supplement these findings, providing a more in-depth assessment of patients' decision-making
15 processes. Future studies could also examine the importance of the duration of treatment effect for
16 non-pharmacological interventions, such as exercise or surgical treatments. It would also be
17 interesting to explore whether the importance of the duration of treatment effect is similar in other
18 chronic conditions and disease populations.

20 Since duration was highly important, future research should explore the different methods that
21 could be used to account for duration in a benefit-risk summary measure. Further work could also
22 compare different representations of treatment effects over time to assess how trialists can make
23 these results easy for patients to interpret.

1

2

3

4

5

6

7

8

9

10

11

12

13

14

15

16

17

18

19

20

21

22

Future studies should also explore the properties of the WOMAC Index. The stiffness domain lacks patient relevance and participants viewed moving between adjacent levels differently depending on the starting level. Researchers should explore whether different weightings of the WOMAC items and levels could produce an outcome measure where changes in the total score more accurately reflect whether an important change has occurred from the patient’s perspective.

Conclusion

The results of this discrete choice experiment suggest that the duration of the treatment effect is an important factor on medication choice for people living with osteoarthritis and that it is viewed with similar importance to the amount of symptom relief provided and risks associated with treatment. Therefore, the duration of treatment effect should be considered when interpreting the results of clinical studies of osteoarthritis medications. However, future research is needed to assess whether these findings are generalisable to other samples, disease areas and levels of duration of effect. Further studies should also explore whether the duration of treatment effect is an important factor in treatment decision-making in other disease areas or non-pharmacological interventions. Clinical trials in osteoarthritis populations should be designed to measure long-term clinical outcomes to monitor whether treatment effects are sustained.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

Acknowledgements:

The authors wish to acknowledge Dr Simon Abram for his helpful advice and comments during the early conception of this study. The authors also wish to acknowledge the project management team and panel members at ResearchNow for providing the sample of participants.

Authors’ contributions:

All authors read and approved the final manuscript. BC designed the study, performed the data analysis and drafted the article. SD, RF and SL contributed to the design of the study, critical revision of manuscript and project supervision. JB and JC contributed to the design and conception of the study, data analysis and interpretation, critical revision of manuscript and project supervision.

Declaration of conflicts of interest:

The Authors declares that there is no conflict of interest.

References:

1. Askling J, Holmqvist M, Ljung L: **Editorial: Is Rheumatoid Arthritis a Mortal Disease?** *Arthritis & rheumatology (Hoboken, NJ)* 2017, **69**(8):1509-1511.
2. Chan S, Jonsson A, Bhandari M: **Planning a clinical research study.** *Indian journal of orthopaedics* 2007, **41**(1):16-22.
3. Lequesne M, Dougados M, Abiteboul M, Bontoux D, Bouvenot G, Chicheportiche V, Dreiser RL, Dropsy R, Maheu E, Mazieres B *et al*: **[How to evaluate the long-term course of osteoarthritis. Tests for trials of fundamental treatments (spine excluded)].** *Revue du rhumatisme et des maladies osteo-articulaires* 1990, **57**(9 (Pt 2)):24s-31s.
4. Pincus T, Sokka T: **Should contemporary rheumatoid arthritis clinical trials be more like standard patient care and vice versa?** *Annals of the rheumatic diseases* 2004, **63** Suppl 2:ii32-ii39.
5. Heneghan C, Goldacre B, Mahtani KR: **Why clinical trial outcomes fail to translate into benefits for patients.** *Trials* 2017, **18**(1):122.
6. Study IoMUCftSAC: **Treating Drug Problems: Volume 1: A Study of the Evolution, Effectiveness, and Financing of Public and Private Drug Treatment Systems.** Washington (DC): National Academies Press (US) 1990.
7. Backhouse ME, Fenn P: **The use of discrete choice analysis in the design of randomised controlled trials.** In: *CRIS Discussion Paper Series – 2006VII.* University of Nottingham; 2006.
8. Schaarschmidt ML, Schmieder A, Umar N, Terris D, Goebeler M, Goerdts S, Peitsch WK: **Patient preferences for psoriasis treatments: process characteristics can outweigh outcome attributes.** *Archives of dermatology* 2011, **147**(11):1285-1294.
9. Gonzalez JM: **Evaluating Risk Tolerance from a Systematic Review of Preferences: The Case of Patients with Psoriasis.** *The patient* 2018, **11**(3):285-300.
10. Posnett J, Dixit S, Oppenheimer B, Kili S, Mehin N: **Patient preference and willingness to pay for knee osteoarthritis treatments.** *Patient preference and adherence* 2015, **9**:733-744.
11. Pincus T: **Rheumatoid arthritis: disappointing long-term outcomes despite successful short-term clinical trials.** *Journal of clinical epidemiology* 1988, **41**(11):1037-1041.
12. da Costa BR, Reichenbach S, Keller N, Nartey L, Wandel S, Juni P, Trelle S: **Effectiveness of non-steroidal anti-inflammatory drugs for the treatment of pain in knee and hip osteoarthritis: a network meta-analysis.** *Lancet (London, England)* 2017, **390**(10090):e21-e33.
13. Cannon CP, Curtis SP, FitzGerald GA, Krum H, Kaur A, Bolognese JA, Reicin AS, Bombardier C, Weinblatt ME, van der Heijde D *et al*: **Cardiovascular outcomes with etoricoxib and diclofenac in patients with osteoarthritis and rheumatoid arthritis in the Multinational Etoricoxib and Diclofenac Arthritis Long-term (MEDAL) programme: a randomised comparison.** *Lancet (London, England)* 2006, **368**(9549):1771-1781.
14. Hernandez-Diaz S, Varas-Lorenzo C, Garcia Rodriguez LA: **Non-steroidal antiinflammatory drugs and the risk of acute myocardial infarction.** *Basic & clinical pharmacology & toxicology* 2006, **98**(3):266-274.
15. McGettigan P, Henry D: **Cardiovascular risk with non-steroidal anti-inflammatory drugs: systematic review of population-based controlled observational studies.** *PLoS medicine* 2011, **8**(9):e1001098.

16. Scheiman JM, Fendrick AM: **Practical approaches to minimizing gastrointestinal and cardiovascular safety concerns with COX-2 inhibitors and NSAIDs.** *Arthritis research & therapy* 2005, **7 Suppl 4**:S23-29.
17. Paulus HE: **FDA arthritis advisory committee meeting: serious gastrointestinal toxicity of nonsteroidal antiinflammatory drugs; drug-containing renal and biliary stones; diclofenac and carprofen approved.** *Arthritis and Rheumatology* 1988, **31**(11):1450-1451.
18. Hauber AB, Arden NK, Mohamed AF, Johnson FR, Peloso PM, Watson DJ, Mavros P, Gammaitoni A, Sen SS, Taylor SD: **A discrete-choice experiment of United Kingdom patients' willingness to risk adverse events for improved function and pain control in osteoarthritis.** *Osteoarthritis and cartilage* 2013, **21**(2):289-297.
19. Bridges JF, Hauber AB, Marshall D, Lloyd A, Prosser LA, Regier DA, Johnson FR, Mauskopf J: **Conjoint analysis applications in health--a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force.** *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research* 2011, **14**(4):403-413.
20. Mangham LJ, Hanson K, McPake B: **How to do (or not to do) ... Designing a discrete choice experiment for application in a low-income country.** *Health policy and planning* 2009, **24**(2):151-158.
21. Vanniyasingam T, Cunningham CE, Foster G, Thabane L: **Simulation study to determine the impact of different design features on design efficiency in discrete choice experiments.** *BMJ open* 2016, **6**(7):e011985.
22. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW: **Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee.** *The Journal of rheumatology* 1988, **15**(12):1833-1840.
23. Kongtharvonskul J, Anothaisintawee T, McEvoy M, Attia J, Woratanarat P, Thakkinstian A: **Efficacy and safety of glucosamine, diacerein, and NSAIDs in osteoarthritis knee: a systematic review and network meta-analysis.** *European journal of medical research* 2015, **20**:24.
24. Puljak L, Marin A, Vrdoljak D, Markotic F, Utrobicic A, Tugwell P: **Celecoxib for osteoarthritis.** *The Cochrane database of systematic reviews* 2017, **5**:Cd009865.
25. Singh JA, Wilt T, MacDonald R: **Chondroitin for osteoarthritis.** *The Cochrane Library* 2006.
26. Smith SR, Deshpande BR, Collins JE, Katz JN, Losina E: **Comparative pain reduction of oral non-steroidal anti-inflammatory drugs and opioids for knee osteoarthritis: systematic analytic review.** *Osteoarthritis and cartilage* 2016, **24**(6):962-972.
27. Shau WY, Chen HC, Chen ST, Chou HW, Chang CH, Kuo CW, Lai MS: **Risk of new acute myocardial infarction hospitalization associated with use of oral and parenteral non-steroidal anti-inflammation drugs (NSAIDs): a case-crossover study of Taiwan's National Health Insurance claims database and review of current evidence.** *BMC cardiovascular disorders* 2012, **12**:4.
28. Varas-Lorenzo C, Riera-Guardia N, Calingaert B, Castellsague J, Salvo F, Nicotra F, Sturkenboom M, Perez-Gutthann S: **Myocardial infarction and individual nonsteroidal anti-inflammatory drugs meta-analysis of observational studies.** *Pharmacoepidemiology and drug safety* 2013, **22**(6):559-570.
29. Goldstein JL, Silverstein FE, Agrawal NM, Hubbard RC, Kaiser J, Maurath CJ, Verburg KM, Geis GS: **Reduced risk of upper gastrointestinal ulcer complications with celecoxib, a novel COX-2 inhibitor.** *The American journal of gastroenterology* 2000, **95**(7):1681-1690.
30. Silverstein FE, Faich G, Goldstein JL, Simon LS, Pincus T, Whelton A, Makuch R, Eisen G, Agrawal NM, Stenson WF *et al*: **Gastrointestinal toxicity with celecoxib vs nonsteroidal anti-inflammatory drugs for osteoarthritis and rheumatoid arthritis: the CLASS study: A randomized controlled trial. Celecoxib Long-term Arthritis Safety Study.** *Jama* 2000, **284**(10):1247-1255.

31. Dahlberg LE, Holme I, Hoyer K, Ringertz B: **A randomized, multicentre, double-blind, parallel-group study to assess the adverse event-related discontinuation rate with celecoxib and diclofenac in elderly patients with osteoarthritis.** *Scandinavian journal of rheumatology* 2009, **38**(2):133-143.
32. Latimer N, Lord J, Grant RL, O'Mahony R, Dickson J, Conaghan PG: **Value of information in the osteoarthritis setting: cost effectiveness of COX-2 selective inhibitors, traditional NSAIDs and proton pump inhibitors.** *PharmacoEconomics* 2011, **29**(3):225-237.
33. Ramey DR, Watson DJ, Yu C, Bolognese JA, Curtis SP, Reicin AS: **The incidence of upper gastrointestinal adverse events in clinical trials of etoricoxib vs. non-selective NSAIDs: an updated combined analysis.** *Current medical research and opinion* 2005, **21**(5):715-722.
34. Smalley WE, Ray WA, Daugherty JR, Griffin MR: **Nonsteroidal anti-inflammatory drugs and the incidence of hospitalizations for peptic ulcer disease in elderly persons.** *American journal of epidemiology* 1995, **141**(6):539-545.
35. Bech M, Kjaer T, Lauridsen J: **Does the number of choice sets matter? Results from a web survey applying a discrete choice experiment.** *Health economics* 2011, **20**(3):273-286.
36. Stamuli E, Torgerson D, Northgraves M, Ronaldson S, Cherry L: **Identifying the primary outcome for a randomised controlled trial in rheumatoid arthritis: the role of a discrete choice experiment.** *Journal of foot and ankle research* 2017, **10**(1):57.
37. Metrics C: **Ngene 1.1. 1 user manual and reference guide.** Sydney, Australia: ChoiceMetrics 2012.
38. Ryan M, Gerard K, Amaya-Amaya M: **Using discrete choice experiments to value health and health care**, vol. 11: Springer Science and Business Media; 2007.
39. de Bekker-Grob EW, Donkers B, Jonker MF, Stolk EA: **Sample Size Requirements for Discrete-Choice Experiments in Healthcare: a Practical Guide.** *The patient* 2015, **8**(5):373-384.
40. Marshall D, Bridges JF, Hauber B, Cameron R, Donnalley L, Fyie K, Johnson FR: **Conjoint Analysis Applications in Health - How are Studies being Designed and Reported?: An Update on Current Practice in the Published Literature between 2005 and 2008.** *The patient* 2010, **3**(4):249-256.
41. Orme B: **Sample size issues for conjoint analysis studies.** *Sawtooth Software Research paper Series Squim*, WA, USA: Sawtooth Software Inc 1998.
42. Bellamy N: **WOMAC osteoarthritis index: user guide IX:** Nicholas Bellamy; 2008.
43. StataCorp: **Stata Statistical Software: Release 15.** In. College Station, TX, USA; 2017.
44. Miguel FS, Ryan M, Amaya-Amaya M: **'Irrational' stated preferences: a quantitative and qualitative investigation.** *Health economics* 2005, **14**(3):307-322.
45. Swait J, Adamowicz W: **Choice environment, market complexity, and consumer behavior: a theoretical and empirical approach for incorporating decision complexity into models of consumer choice.** *Organizational Behavior and Human Decision Processes* 2001, **86**(2):141-167.
46. Veldwijk J, Determann D, Lambooy MS, van Til JA, Korffage IJ, de Bekker-Grob EW, de Wit GA: **Exploring how individuals complete the choice tasks in a discrete choice experiment: an interview study.** *BMC medical research methodology* 2016, **16**:45.
47. Culliford D, Maskell J, Judge A, Cooper C, Prieto-Alhambra D, Arden NK: **Future projections of total hip and knee arthroplasty in the UK: results from the UK Clinical Practice Research Datalink.** *Osteoarthritis and cartilage* 2015, **23**(4):594-600.
48. Yu D, Jordan KP, Bedson J, Englund M, Blyth F, Turkiewicz A, Prieto-Alhambra D, Peat G: **Population trends in the incidence and initial management of osteoarthritis: age-period-cohort analysis of the Clinical Practice Research Datalink, 1992-2013.** *Rheumatology (Oxford, England)* 2017, **56**(11):1902-1917.

- 1 49. Lin J, Zhang W, Jones A, Doherty M: **Efficacy of topical non-steroidal anti-inflammatory**
2 **drugs in the treatment of osteoarthritis: meta-analysis of randomised controlled trials.**
3 *BMJ (Clinical research ed)* 2004, **329**(7461):324.
- 4 50. Jackson H, Barnett LA, Jordan KP, Dziedzic KS, Cottrell E, Finney AG, Paskins Z, Edwards JJ:
5 **Patterns of routine primary care for osteoarthritis in the UK: a cross-sectional electronic**
6 **health records study.** *BMJ open* 2017, **7**(12):e019694.
- 7 51. Martel-Pelletier J, Barr AJ, Cicuttini FM, Conaghan PG, Cooper C, Goldring MB, Goldring SR,
8 Jones G, Teichtahl AJ, Pelletier JP: **Osteoarthritis.** *Nature reviews Disease primers* 2016,
9 **2**:16072.
- 10 52. Raja R, Dube B, Hensor EM, Hogg SF, Conaghan PG, Kingsbury SR: **The clinical characteristics**
11 **of older people with chronic multiple-site joint pains and their utilisation of therapeutic**
12 **interventions: data from a prospective cohort study.** *BMC musculoskeletal disorders* 2016,
13 **17**:194.
- 14 53. Shelbaya A, Solem CT, Walker C, Wan Y, Johnson C, Cappelleri JC: **The economic and clinical**
15 **burden of early versus late initiation of celecoxib among patients with osteoarthritis.**
16 *ClinicoEconomics and outcomes research : CEOR* 2018, **10**:213-222.
- 17 54. Cordero-Ampuero J, Darder A, Santillana J, Caloto MT, Nocea G: **Evaluation of patients' and**
18 **physicians' expectations and attributes of osteoarthritis treatment using Kano**
19 **methodology.** *Quality of life research : an international journal of quality of life aspects of*
20 *treatment, care and rehabilitation* 2012, **21**(8):1391-1404.
- 21 55. Llewellyn-Bennett R, Edwards D, Roberts N, Hainsworth AH, Bulbulia R, Bowman L: **Post-trial**
22 **follow-up methodology in large randomised controlled trials: a systematic review.** *Trials*
23 2018, **19**(1):298.
- 24
- 25
- 26