The Bone & Joint Journal

Supplementary Material

10.1302/0301-620X.106B12.BJJ-2023-1266.R2

The DCE development process

Attribute and level identification and selection

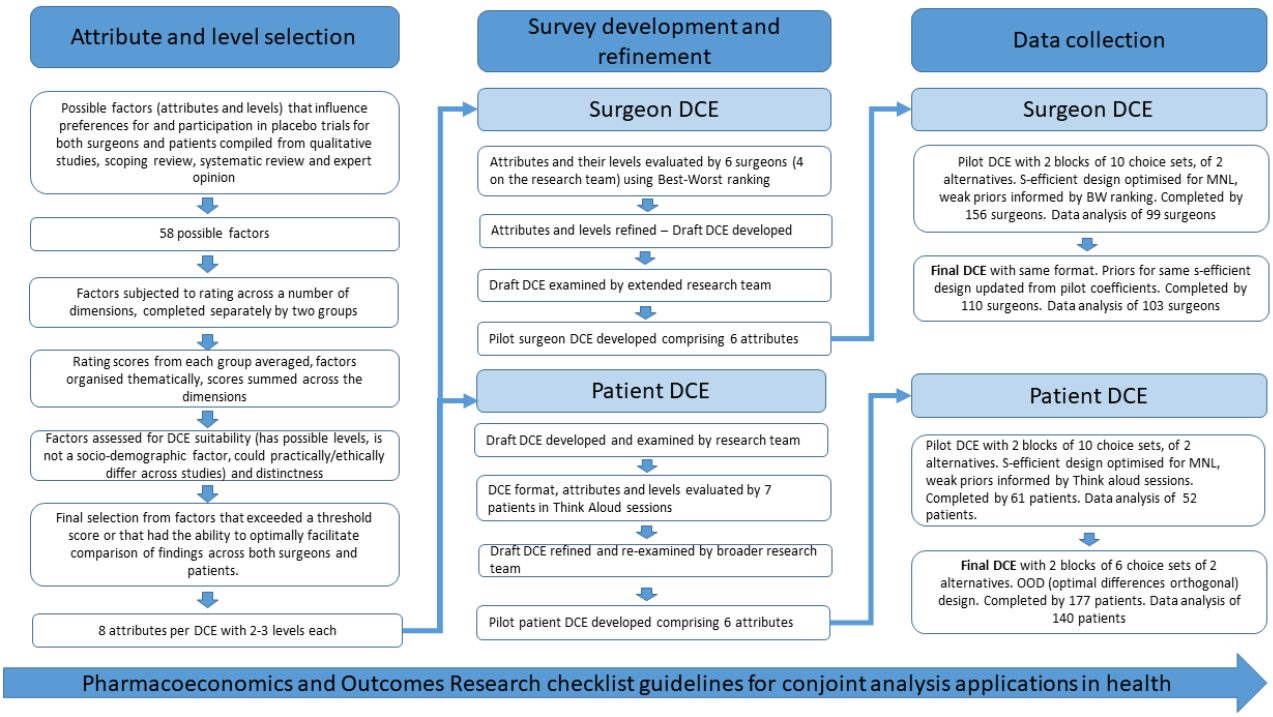


Fig a. The survey development process

A rigorous and extensive multi-methods approach was used to determine the attributes and levels of the discrete choice experiments (DCEs). An initial list of all possible factors that influence preferences for and participation in placebo trials for both patients and surgeons was sourced from a qualitative study,[1] scoping review,[2,3] and systematic review,[4] as well as expert opinion among the research team. This extensive list of 58 possible factors was then subjected to a comprehensive rating table analysis which was completed separately by two groups comprising researchers and clinicians across health economics, epidemiology, surgery, nursing, physiotherapy, and psychology. A final selection of eight attributes was compiled from those that either had a score greater than a pre-determined criterion, or were suitable for both patients and surgeons to more optimally facilitate the comparison of findings across both surgeons and patients.

For the surgeon group, the reduced selection of eight attributes was developed into a best-worst ranking analysis, which was completed by six surgeons, who indicated which attributes of participating in a placebo surgery trial were the most and least important to them, and which levels were their most and least preferred. The best-worst ranking exercise was chosen for surgeons to minimize the required response time and facilitate completion within their busy schedules.[5] The surgeons were provided with the opportunity to make comments and add additional attributes and levels. A draft DCE was created from these results and thoroughly discussed among the extended research team.

For the patient group, a draft patient DCE was developed from the reduced selection of eight attributes and evaluated by seven patients using a 'think aloud' format. The think-aloud format was used for patients to provide richer data and ensure the DCE was comprehensible and appropriate for patients.[6,7] Think-aloud sessions were conducted with individual patients via video conferencing or phone interview, and included questions about the attributes, levels, format, length, use of images and overall comprehension of the DCE. An updated patient DCE was developed from talk-aloud session findings and developed into a pilot DCE completed by 52 patients, where there was an opportunity for feedback. The preferred attributes and levels were determined through pilot DCE analysis with a baseline category multinomial logit model. To assess the level of cognitive burden, data quality and bias tests across the choice-sets were conducted, revealing a decline in quality in the final choice-sets, consistent with patient feedback.

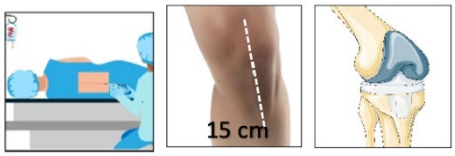
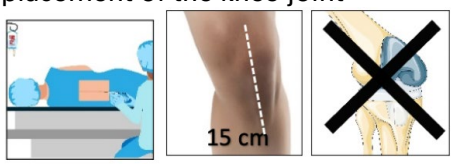
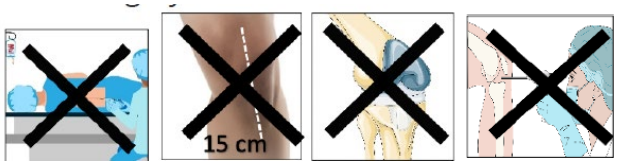
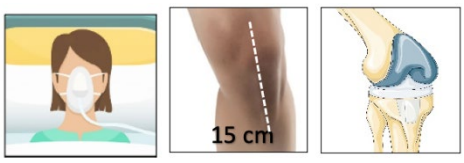
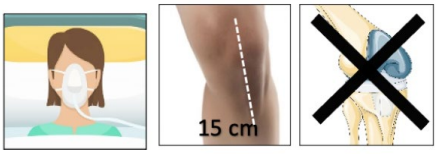
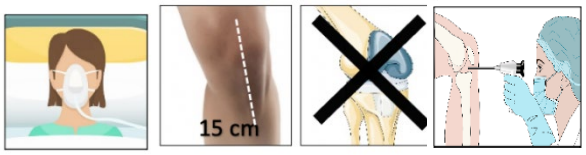
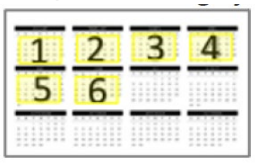
DCE construction and format

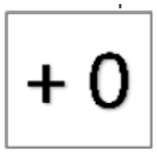
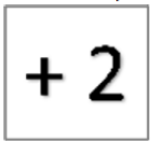
As seen in Table i, the surgeon DCE was presented in the standard table format, with a row for attributes and with the key words which differed across levels highlighted to assist comprehension if read quickly. As seen in Table ii, the patient DCE utilized pictures as well as bold font, removed the attribute label column, and combined the delivery of some attributes (the anaesthetic, the arms of the study and the terminology used) into the same row and sentence. This was done to make the DCE more ecologically valid and better reflect the way a non-hypothetical study may be presented to patients. These decisions were interrogated through the think-aloud interviewing and found to be appropriate. The order of the attributes and alternatives were fixed across participants for each of the unique 20 (surgeon DCE) and 12 (patient DCE) choice-sets. Respondents had the option to return to previous questions and alter responses prior to survey completion.

Table i. Example choice set for the surgeon discrete choice experiment (DCE). Above each choice set was the following information: "Below are two hypothetical 12-month placebo surgery studies. As a surgeon, participation would involve your patients (who consent) being randomised to the study and you performing the placebo/surgery on your patients based on their randomization. For the purposes of these questions, a high-fidelity placebo is one which closely mimics the index procedure but leaves out the supposed therapeutic component. A low-fidelity placebo is one which involves an anaesthetic and skin incision only. Indicate which study design you would prefer and if you would want to participate in that study." Below each choice-set were two questions, each with radio buttons. They were: "I prefer Clinical Study A / I prefer Clinical Study B" and "I **would** participate in the study I preferred / I **would not** participate in the study I preferred".

Study characteristic	**Clinical Study A**	**Clinical Study B**
The procedure being tested	Knee **replacement** surgery	Knee **arthroscopy**
Who approaches the patients and describes the study to them	You, the **surgeon**	A **research nurse**
The arms of the study	A **two** arm study Arm 1) the standard surgery Arm 2) a **low fidelity** placebo	A **three** arm study Arm 1) the standard surgery Arm 2) a **low fidelity** placebo Arm 3) a **high fidelity** placebo
The anaesthetic used in the study	**Regional** anaesthetic	**General** anaesthetic
The conditions under which patients can crossover	After **6 months, before** results are known	After **12 months, before** results are known
Role in the trial	In addition to performing the surgeries, you **would be invited** to have input into the trial design and to join the authorship team	You would only perform the surgeries, you **would not have** input into the trial design and **would not be invited** to join the authorship team

Table ii. Example choice set for the patient discrete choice experiment (DCE). Above each choice set was the following information: "You are waiting on a knee replacement surgery. This involves an anaesthetic, a 15 cm skin cut and knee joint replacement. A study coordinator asks if you would like to participate in a research study testing the effectiveness of knee replacement surgery. The study coordinator describes two possible clinical studies to you." Below each choice set were two questions, each with radio buttons. They were: "Select here if you would prefer Clinical Study A/ Select here if you would prefer Clinical Study B" and "Select here if you **would** participate in the study you preferred / Select here if you **would not** participate in this study and proceed with surgery as planned" The procedure attribute (i.e. knee replacement surgery; knee arthroscopy) only varied across choice sets so that the decisions more closely resembled the kind of decision patients may make about participating in a clinical study.

Clinical Study A	**Clinical Study B**
In this study, there are three treatment arms; 1) the standard knee replacement surgery with a regional/spinal anaesthetic  2) a similar **placebo** surgery with a regional/spinal anaesthetic and skin cut but no replacement of the knee joint  3) no surgery and continue as usual. 	In this study, there are three treatment arms; 1) the standard knee replacement surgery with a general anaesthetic  2) a similar **sham** surgery with a general anaesthetic and skin cut but no replacement of the knee joint  3) another similar surgery which involves a general anaesthetic, skin cut and inspection,  but no replacement of the knee joint.
If you are not satisfied with your symptoms, you have the option of undergoing the standard surgery 6 months after your surgery, during the study, **before results are known**. 	If you are not satisfied with your symptoms, you have the option of undergoing the standard surgery 12 months after your surgery, after the study has ended, **if the standard surgery is found to be more effective**.' 

You would not have any extra study appointments in addition to your usual follow up appointments. 	You would have a couple of extra study appointments in addition to your usual follow up appointments. 

The preference elicitation question

Two preference elicitation questions were asked for each choice-set presented. The first, a forced choice: "Which study design do you prefer?" and the second, an opt-in: "Would you participate in the study you preferred?". Both required a response before progressing to the next choice-set. Although the participation (opt-in) question alone addresses the key research question, this dual-question design was favoured as it was anticipated that many respondents would have strong views regarding participation in placebo-controlled surgical trials and select 'yes' or 'no' to all questions. Including only the opt-in question would, therefore, likely limit the scope and richness of the results.

Pilot DCE design

Two pilot DCEs were implemented, one for surgeons (n = 99) and one for patients (n = 52). Both pilot DCEs used an S-efficient design, optimized for a multinomial logit model (MNL) model, with two blocks of 10 choice-sets, each with two alternatives. This number of choice-sets was consistent with the average number of choice-sets for blocked designs.[8] When detailed prior information about the attributes is unavailable from previous studies or a pilot, point estimates close to zero are recommended.[9] For both pilot DCEs, the overall size of the coefficients was therefore estimated at around 0.5 for preferred levels (consistently ranked/described as important) and 0.1 for levels with more ambiguous preferences.

DCE design decisions

The continued use of the S-efficient design was considered the most appropriate design to meet the sample size limitations of the surgeon cohort; there is a fixed number of orthopaedic surgeons in Australia from which to be sampled.

Tests of data quality and patient feedback indicated that a design that allowed fewer choice-sets while maximizing differences between attributes would be more appropriate for the patient DCE to balance patient comprehension and statistical power. As a result, the final patient DCE utilized an optimal differences orthogonal (OOD) design with fewer choice-sets (from ten to six) and a reduction in levels for three of the attributes (number of additional appointments, conditions under which patients can crossover and terminology used to describe the placebo).

It is considered good practice to include a dominant alternative, or a repeated choice set in the DCE design to assess its validity.[10] We did not include either of these options for the following reasons. There was no clearly dominant alternative that we could include as we were uncertain of the direction of preference for the levels of most attributes so could not include a dominant alternative that effectively functioned as a validity check. We could have included a repeated option, however, one of the themes that emerged from the feedback from both groups from the pilot testing was that face validity was important, and that two ways of achieving that were to ensure there were clear differences between the options, with as few repeated options as possible was important. These pieces of feedback, such as "I do not understand what it is supposed

to achieve" and it was "too repetitive", showed that in this engaged participant group, who have strong opinions, face validity (participants perceiving they knew what the study was testing) was important in ensuring that participants took the study seriously, and that adding a repeated option (and maintaining a design where there was a lot of repeated levels across the two options) might weaken the face validity of the tool. Another related important piece of information from the pilot testing was that we needed to ensure we had as minimum number of choice-sets as possible to reduce fatigue (for patients) and minimize time away from work (surgeons). Further, as it was anticipated that it would be difficult to recruit a large sample, and because the number of 'true' choice-sets is an important factor along with sample size, it was decided therefore to maximize the difference between the options and omit a repeated option to ensure the face validity and maximize the number of 'true' choice-sets while minimizing the time taken to complete the DCE.

Additional statistical analysis

Sample size planning

Sample size planning was conducted using Louviere and Swait's[11] parametric method (as described in equation 1 below) and Ngene estimates (for the efficient design) to determine the minimum sample size required.

Louviere's method provides a minimum sample size required to run an estimable model, and a recommended sample size to estimate the two block designs. For the surgeon group, given a baseline choice probability of 50% (1/2), accuracy level of 90%, confidence level of 95% and 10 choice tasks (observations) per respondent, we required minimum sample size of 39 surgeons. For the patient group, given a baseline choice probability of 50% (1/2), accuracy level of 90%, confidence level of 95% and 6 choice tasks (observations) per respondent, we required minimum sample size of 65 patients. The recommended number of respondents for the two-block design were 78 surgeons and 130 patients.

$$N \geq \frac{1-P}{TPJ(\beta^2)}\left[\phi^{-1}\left(\frac{1+\alpha}{2}\right)\right]^2 \qquad (1)$$

where T is the number of choice tasks per respondent; J is the number of alternatives per task; β accuracy level; α confidence level and P is the expected choice probability (when no prior information about respondents' preferences is known – this corresponds to 1/J); and ϕ is the normal distribution (CDF).

It should be noted that these estimates of minimum samples required to run an estimable model do not take into consideration the number of attributes and levels, nor the efficiency of the design. The patient DCE included fewer choice sets, which results in a greater minimum sample size according to Louviere's method (compared to the surgeon), however the patient DCE utilized a design which maximized the differences between the levels of each attribute for each choice set, which mediates the need for as many choice-sets for a minimum sample size. Similarly, the patient DCE had fewer levels for each attribute than the surgeon DCE, which also assists in the estimation of the model, as there are fewer factors that need estimating from the same number of choices. Therefore, these minimum sample sizes are purely a guide, and cannot be exclusively used to determine if the results are or are not valid if they are not met.

The Ngene estimated sample size recommended minimum sample size for the surgeon S-efficient design was 131. The number of patients and surgeons who completed at least one question from the survey instrument was 177 and 110, respectively.

Data cleaning

Respondents who did not complete the entire survey, or completed it in fewer than three minutes, were excluded from the analysis (surgeons = 2/110, patients = 5/177). As the levels were balanced across the alternatives, participants should have an equivalent preference for alternatives A and B and select each for approximately 50% of the choices. Respondents who selected one generic alternative (A or B) as the preferred alternative for at least 80% of the choice-sets were thus further excluded from analysis (surgeons = 5/108, patients = 32/172), as this was deemed indicative of either a strong response bias or protest responding due to the forced preference choice. The sample sizes used in the analysis of the preference question were, therefore, 103 surgeons and 140 patients. Similarly, respondents who selected either 'yes' or 'no' to 80% of the participation questions were removed from the analysis of the participation question (surgeons = 61/103, patients = 114/140). While strong preferences for the 'yes' or 'no' response to the participation question are reflective of true preferences, the data from a respondent's choice tasks are not useful if their responses of 'yes' or 'no' were influenced by their overall feelings towards placebo surgery trials, and not the levels presented in each choice. The sample sizes used in the analysis of the participation question were, therefore, 42 surgeons and 26 patients. The smaller sample sizes for the participation questions therefore resulted from a deliberate filtering process designed to exclude participants prone to straight-lining behaviour - namely, those who indicated either a willingness or unwillingness to participate in all types of trials. This decision was made on the understanding that including such participants in the participation question would introduce biases in our estimations, thereby compromising the participation question would introduce biases in our estimations, thereby compromising the validity of our results.

Referring to the minimum required samples of 39 (surgeons) and 65 (patients) noted in the section above, both surgeon and patient samples substantially exceeded these minimums for the preference question, however, for the participation question, only the surgeon sample exceeded the minimum sample required. While this smaller sample size for patients for the participation question does introduce concerns around the validity of the results, it was a result of addressing another validity concern. Further, as noted in the above section, the minimum sample sizes are a guide only and are not in themselves an indication of a lack of validity of the patient data. Nonetheless, it does require caution in interpretation of the effects.

Econometric analysis of DCE data

Respondents' preferences and participation choices for a clinical study were modelled using a MNL. This models the utility of participating in clinical study *j* in choice *t* for respondent *n* as:

$$U_{ntj} = \beta \boldsymbol{X_{ntj}} + \varepsilon_{ntj} \quad (2)$$

$\boldsymbol{X_{ntj}}$ is a vector representing the clinical study or participation attributes of alternative *j* presented to respondent *n* (patient or surgeon) in choice *t* and β is the marginal utility of each attribute and ε_{ntj} is the error term following a Gumbel distribution. This analysis assumes that respondents gain utility (or satisfaction) from participation in a clinical study, the utility gained depends on the clinical study participation attributes, and respondents choose the clinical study that would bring them the highest utility such that:

$$U_{ntj}^{SURGEONS} = \beta_1 PROCEDURE + \beta_2 WHO\ APPROACHES + \beta_3 STUDY\ ARMS + \beta_4 ANAESTHETIC + \beta_5 CROSSOVER\ CONDITIONS + \beta_6 ROLE\ IN\ TRIAL + \varepsilon_{ntj} \quad (3)$$

$$U_{ntj}^{PATIENTS} = \beta_1 TERMINOLOGY + \beta_2 PROCEDURE + \beta_3 ANAESTHETIC + \beta_4 STUDY\ ARMS + \beta_5 APPOINTMENTS + \beta_5 CROSSOVER\ CONDITIONS + \varepsilon_{ntj} \quad (4)$$

The interpretation of coefficients depends on the attributes' unit of measurement. The signs (+/–) of the coefficients indicates if a unit change in the attribute increases or decreases the likelihood of choosing a clinical study or participation in a clinical study."

Additional Discussion

Discussion of non-significant (participation) results

Although studies have found that one of the barriers to patient participation in research trials is the extra time commitment,[12-14] consultation with consumer representatives and design experts during the attribute selection stage of this DCE suggested that some patients might view the extra follow up appointments in a clinical trial as an incentive to participation. Our results suggest that when choosing between two placebo surgery trials, it made no difference if the study had no extra appointments or a couple of extra appointments. Similarly, studies have reported that the use of the word 'sham' compared with 'placebo' could lead to different perceptions,[15,16] and amongst the pre-DCE consultations with surgeons, there was a strong belief that the term 'sham' is aversive and could likely contribute to participant's lack of engagement with such studies. Our results suggest that if any influence of the word exists, it is overshadowed by other characteristics, as there was no effect of the word being sham or placebo on people's preference or willingness to participate. While these would have been easily modifiable factors in designing a new placebo surgery trial, there is little evidence from this study to suggest they would have made a difference to participation.

There were also a number of attributes where the levels were significantly different from each other when selecting the preferred study, but not so when selecting if they would participate in this study. For the patient group, there was significant preference away from a three-arm study, where the third arm is no surgery, compared to a two-arm study with the intervention and placebo control (skin incision and anaesthetic only). However, whilst the direction of the preference remained for the participation question the finding was no longer significant. Given the small sample for the participation data, there is a possibility that this effect simply failed to reach significance due to a lack of power. Interestingly, there was no significant difference in preference or participation for the surgeons among the different study arm options. It is possible that if asked about this attribute alone surgeons may indicate different preferences between the different options, as seen in,[1] but when considered among other features in the DCE format, it is not an important enough difference to influence their choice between the different trials.

Whilst previous research has noted the general concerns around the use of a general anaesthetic,[1,16] the option of a regional anaesthetic did not lead to studies being more selected as preferred or more participated in for either group. In fact, for the preference question for the surgeons there was actually a significant preference for general anaesthetic. This could reflect surgeons selecting attributes which improve the scientific rigor of the study as general anaesthetic can assist with issues of blinding in placebo surgery trials.[16] However, the direction of preference changed for surgeons for the participation question. Although this effect was not significant, it could reflect the fact that general anaesthetic comes with a greater risk, and so when considering whether they would wish to participate in such a study they may not wish

to personally perform a procedure under general anaesthetic that could be safely performed under a regional anaesthetic. This opens the possibility that surgeons may answer the two questions differently; this is explored further in the limitation section.

For both the patient and surgeon groups there was an effect on preferences for the crossover attribute that also did not translate into participation decisions. Patients significantly preferred hypothetical studies where they could crossover to the intervention arm at six months, before the study results were known, compared to studies where they would have to wait 12 months and only be offered the option to crossover if the study found the intervention to be effective. This is consistent with literature that has found some patients enrol in placebo surgery trials with the hope that they will receive the intervention procedure, thus questioning the equipoise of patients and the possibility of therapeutic misconception.[14] Surgeons were provided with more crossover options than patients but similarly preferred the six-month crossover option, and additionally preferred that patients be allowed to crossover at 12 months, before the results are known, compared to the 12-month crossover contingent on intervention effectiveness. The three-month crossover option was not significantly different from the 12-month contingent option for surgeons. While the direction of preference for the different crossover options was consistent across the preference and participation questions for patients, it reversed for surgeons. This is discussed in the limitation section below.

The sample sizes for the preference question were consistent with those estimated for adequate power. However, for the participation question, there was a necessary removal of many surgeons and patients with 'non-modifiable views' (yes or no to all questions). The smaller sample sizes for the participation questions thus resulted from a deliberate filtering process designed to exclude participants prone to straight-lining behaviour - namely, those who indicated either a willingness or unwillingness to participate in all types of trials. This decision was made on the understanding that including such participants in the participation question would introduce biases in our estimations, thereby compromising the participation question would introduce biases in our estimations, thereby compromising the validity of our results. However, it should be noted that this approach comes with inherent trade-offs, notably the reduction in sample size and the consequent impact on the statistical power to detect smaller or more variable effects. This limitation was observed in the variations between the preference and participation questions within the surgeon group, where the direction of preference for crossover and anaesthetic attributes appeared inconsistent.

Despite these limitations, it is important to note that our findings, particularly those attributes with statistically significant effects, remain robust. The consistency in the direction of preference between the preference and participation questions for these attributes with significant effects, and the consistency of preferences between the surgeons and patients, underscores the validity of our findings, even within a smaller sample size constraint. This suggests that, for key attributes, our study can provide reliable estimates that contribute meaningful insights into the preferences of surgeons and patients towards placebo surgery trials.

The lack of consistency between the preference and participation questions is both a limitation and an interesting opportunity for future research. Given the small sample sizes, noted above, the lack of significance in the participation question seen for attributes with significant differences for the preference question (the anaesthetic and crossover attributes) is likely explained by a lack of power. There is a curious difference seen in the surgeon DCE, however, that is worth considering. For the

surgeons, while the three levels of the crossover attribute were not significantly different from the baseline level for the participation question, they all had effects in the opposite direction to those seen in the preference question, two of which were significant. The same reversal was found for the anaesthetic attribute. While this could well be a statistical artifact not to be trusted because of the low power, the consistency of the difference across the crossover attribute poses the question - is there in fact a psychological difference between choosing which study is better, and which you would participate in? For example, surgeons may have a research understanding that premature crossover to the intervention arm (for example at six months) can lead to a potential source of bias, and thus not want to participate in such a study, leading to a negative coefficient for this level in the participation question. They may also however believe as a clinician that the surgery in question is effective and patients should receive it as soon as possible, and thus consider studies with a shorter crossover time as the better study, leading to a positive coefficient for the same level in the preference question. Potentially, in answering which is the better study, surgeons consider the question from a clinical perspective, but when answering the participation question, they consider it as a researcher. If true clinical equipoise existed, as it should for any surgeon participating in a trial, these conflicts may not arise. However, the surgeons completing this survey were not surgeons who have already committed to participation, and thus such conflicts may be more present in hypothetical questions as these. Future research into the potential trade-offs and different decisions surgeons make as clinician vs researcher is certainly worthwhile, but so is future methodological research examining consistency between forced choice preference style questions and opt in/out participation or purchasing style questions.

Table iii. Clinical characteristics, views of placebo surgery, and overall willingness to participate for surgeons.

Variable	Total
Total, n	103
Percentage of consultants	92%
Median years practising (range)	14 (0 to 54)
Percentage practising only in the public health system	11%
Percentage practising only in private practice	18%
Percentage who have been previously involved in research	
CI	59%
Surgeon	59%
Participant	32%
Have not been involved in research	3%
Mean risk score (1 to 5) (SD)	
Financial risk-taking	2.5 (1.2)
Clinical risk-taking	2.5 (1.1)
Career risk-taking	2.6 (1.1)
Percentage who stated that having an orthopaedic placebo surgery trial occurring in their hospital would have a negative; positive; neutral impact on their case load	29%; 16%; 55%
Percentage who stated that being involved in an orthopaedic placebo surgery trial would have a negative; positive; neutral impact on their case load	37%; 15%; 49%
Percentage who stated that being an investigator in an orthopaedic placebo surgery trial would have a negative; positive; neutral impact on their case load	33%; 22%; 45%
Percentage who were, in principle, willing to be involved (perform the randomized surgeries) in a placebo surgery trial.	63%
Percentage of DCE responses selected as willing to participate (across participants and alternatives)	46%
Percentage of surgeons who always selected in the DCE they would not participate (always selected they would)	31% (18%)

Table iv. Sociodemographic characteristics and overall willingness to participate of patients. Not all percentages = 100% for all items, e.g. the procedure participants were waitlisted for, as participants could select multiple options.

Variable	Total
Total, n	140
Hospital recruited from, n (%)	
Hospital 1	41 (29)
Hospital 2	99 (71)
Mean age, yrs (range)	59 (19 to 83)
Sex, n (%)	
Female	76 (54)
Male	64 (46)
Employment status, n (%)	
Casual	10 (7)
Full-time	29 (21)
Part-time	2 (1)
Retired/unemployed not looking for work	82 (59)
Seeking employment	8 (6)
Self-employed	9 (6)
Annual income, n (%)	
$0 to $18,200	36 (25)
$18,201 to $45,000	63 (46)
$45,001 to $120,000	37 (26)
$120,001 to $180,000	3 (2)
$180,001 and over	1 (1)
Procedure waitlisted for, n (%)	
Knee scope	9 (6)
ACL reconstruction	9 (6)
Partial knee replacement	5 (4)
Total knee replacement	58 (41)
Hip scope	1 (1)
Total hip replacement	49 (35)
Shoulder scope	2 (1)
Shoulder replacement	2 (1)
Hand, wrist or elbow surgery	2 (1)
Other	9 (6)
Private health insurance, n (%)	
Yes	16 (11)
No	124 (89)
Previous surgery, n (%)	
Yes	112 (80)
No	28 (20)
Previous research, n (%)	
Yes	20 (14)
No	112 (80)
Unsure	8 (6)
Mean pain score (SD)*	2.8 (0.86)
Mean ability score (SD)†	2.5 (0.93)
Mean health literacy score (SD)	2.3 (3.0)

Mean risk (SD)	5.8 (2.4)
DCE responses selected as willing to participate (across participants and alternatives), %	45
Patients who always selected in the DCE they would not participate (always selected they would), %	40 (30)

*1 to 5, no pain to extreme pain.
†1 to 5, no problems to unable.
‡0 to 16, good to poor.
¶0 to 10, low to high risk seeking.

ACL, anterior cruciate ligament; DCE, discrete choice experiment.

References

1. **Bunzli S, Nelson E, Wall L, et al.** factors underlying patient and surgeon willingness to participate in a placebo surgery controlled trial: a qualitative investigation. *Ann Surg Open*. 2021 Nov 4;2(4):e104.
2. **Wall L, Hinwood M, Lang D, et al.** Attitudes of patients and surgeons towards sham surgery trials: a protocol for a scoping review of attributes to inform a discrete choice experiment. *BMJ Open*. 2020 Mar 10;10(3):e035870.
3. **Hinwood M, Wall L, Lang D, et al.** Patient and clinician characteristics and preferences for increasing participation in placebo surgery trials: a scoping review of attributes to inform a discrete choice experiment. *Trials*. 2022 Apr 12;23(1):296.
4. **Bunzli S, Choong E, Shadbolt C, et al.** Placebo surgery controlled trials: do they achieve what they set out to do? A systematic review. *Ann Surg*. 2021 Jun 1;273(6):1102-1107.
5. **Yangui A, Akaichi F, Costa-Font M, Gil LM.** Are ranking preferences information methods comparable with the choice experiment information in predicting actual behavior?. 2014 International Congress, August 26-29, 2014, Ljubljana, Slovenia 182672, European Association of Agricultural Economists.
6. **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 Plan*. 2009 Mar;24(2):151-158.
7. **Hollin IL, Craig BM, Coast J, et al**. Reporting formative qualitative research to support the development of quantitative preference study protocols and corresponding survey instruments: guidelines for authors and reviewers. *Patient*. 2020 Feb;13(1):121-136.
8. **Sochacki KR, Mather RC, Nwachukwu BU, et al.** Sham surgery studies in orthopaedic surgery may just be a sham: a systematic review of randomized placebo-controlled trials. *Arthroscopy*. 2020 Oct;36(10):2750-2762.e2.
9. **Szinay D, Cameron R, Naughton F, Whitty JA, Brown J, Jones A.** understanding uptake of digital health products: methodology tutorial for a discrete choice experiment using the bayesian efficient design. *J Med Internet Res*. 2021 Oct 11;23(10):e32365.
10. **Johnson FR, Yang JC, Reed SD.** The internal validity of discrete choice experiment data: a testing tool for quantitative assessments. *Value Health*. 2019 Feb;22(2):157-160.
11. **Louviere J, Hensher DA, Swait JD.** Stated choice methods: analysis and application. Cambridge, UK: Cambridge University Press. 2000.
12. **Ackerman IN, Buchbinder R, Osborne RH.** Factors limiting participation in arthritis self-management programmes: an exploration of barriers and patient preferences within a randomized controlled trial. *Rheumatology (Oxford)*. 2013 Mar;52(3):472-479.
13. **Friedman DB, Foster C, Bergeron CD, Tanner A, Kim SH.** A qualitative study of recruitment barriers, motivators, and community-based strategies for increasing clinical trials participation among rural and urban populations. *Am J Health Promot*. 2015 May-Jun;29(5):332-338.
14. **Kim SY, Wilson RM, Kim HM, et al.** Comparison of enrollees and decliners of Parkinson disease sham surgery trials. *Mov Disord*. 2012 Apr;27(4):506-11.

15. **Campbell MK, Skea ZC, Sutherland AG, et al.** Effectiveness and cost-effectiveness of arthroscopic lavage in the treatment of osteoarthritis of the knee: a mixed methods study of the feasibility of conducting a surgical placebo-controlled trial (the KORAL study). *Health Technol Assess*. 2010 Jan;14(5):1-180.
16. **Campbell MK, Entwistle VA, Cuthbertson BH, et al.** Developing a placebo-controlled trial in surgery: issues of design, acceptability and feasibility. *Trials*. 2011 Feb 21;12:50.