## Quantitative Analysis of the Feasibility of a Cataract Decision Aid RCT

**Quantitative Analysis for Involve-CAT Feasibility of a Cataract Decision Aid Randomised Controlled Trial**

Involve-CAT

A feasibility assessment for a possible future fully powered Randomised Controlled Trial of the use of a Cataract Decision Aid providing information on cataract surgery, including personalised risks and benefits.

A quantitative analysis as part of Work Package 4 of a NIHR funded Cataract Research Programme.

Analyses and first draft by Mariusz Grzeda

Writing completed by John SparrowBackground

The overarching aim of the cataract research programme was to investigate possible ways to improve decision making processes for people approaching cataract surgery. In relation to this element of the work, the grant application assumed earlier development of statistical models predicting the outcome of surgery including occurrence of an operative surgical complication (significant breach of the lens-zonule barrier referred to as Posterior Capsule Rupture – PCR), vision loss related to surgery (Visual Acuity Loss or VA Loss) and self-reported benefit from surgery based on Cat-PROM5 scores (a Patient Reported Outcome Measure or PROM). Measures thus included both subjective and objective indices. During the first part of the programme (Work Package 1 or WP1) the PROM, Cat-PROM5 was developed and validated. The predictive models for PCR and VA Loss were developed and validated as a part of WP2. In the third work package (WP3) factors predicting self-reported benefits from a cataract operation were explored with regards to a change in the Cat-PROM5 measure between pre- and post-operation time-points. The aim of the final work package (WP4) was to incorporate the predictive models for risks and benefits into a decision aid and to assess how this influenced, and was perceived by patients approaching cataract surgery, and the clinicians delivering their care. Specifically, it was designed to reduce uncertainty and confusion in relation to a decision about whether or not to proceed with a cataract operation. Current practice is such that the possible benefits and risks of potential adverse outcomes are presented to patients in vague terms such as ‘likely to see better after the operation’ or as a list of possible complications that might happen along with average rates.

Aims of Involve-CAT (WP4)

The study took the form of a feasibility study exploring the possibility of establishing a future randomised controlled trial (RCT) using a cataract decision support aid as an intervention. The development of the Cataract Decision Aid (CDA) and the qualitative analysis of its performance are described in separate reports, this report will cover the analysis of the quantitative data arising from the feasibility trial. The key hypothesis considered in the study was that the quality of the process of patient-clinician Shared Decision Making (SDM) in cataract surgery is improved through use of a Cataract Decision Aid (CDA) because it improves patient knowledge, encourages the patient’s deliberation process and increases patient’s ‘readiness’ for making an informed decision about the treatment. Information contained in the CDA (see Appendix 8) included general information about cataract surgery as well as specific individualized risk and benefit predictions based on the patient’s eye and general health. The quantitative outcomes presented here were extensively supported by qualitative analyses of a Shared Decision Making (SDM) approach based on in-depth interviews reported separately.

Specific quantitative issues explored in WP4 included:

1. Examination of the suitability of candidate quantitative outcome metrics for use in a possible future fully powered RCT, these metrics together forming the Cataract Decision Quality Measure, (CDQM - developed as a part of WP3).
2. Consideration of the feasibility RCT effect sizes to inform sample size estimation for a possible future fully powered RCT designed to fully test possible benefits of using a CDA in clinical practice. The approach assumed standard levels of alpha (type-I) and beta (type-II) errors as described in the power study section of this report.
3. A validation exercise for of the model predicting self-reported benefits from cataract surgery expressed as a change in the self-reported Cat-PROM5 measure between pre-operative and post-operative assessments. *Model development is the subject of a separate report – see Appendix 6.*
4. The estimation of possible costs arising from potential wide scale implementation of the decision aid. *This is the subject of a separate report – see Appendix 15.*

Study design

The reported feasibility study took a form of a two-arm RCT with the CDA as an intervention. The intervention group was defined as a group of patients in which the CDA was used while patients in the control group underwent standard NHS care. The allocation of patients to groups was conducted through a 1:1 block randomisation process by centre. It was assumed that within each centre 5-6 participants should be allocated within each arm (receiving the CDA intervention vs. not receiving the intervention).

The research process was multistage, starting with pre-screening and proceeding through assessment of patient eligibility for the study, recruitment, obtaining consent, randomisation, baseline clinical and self-reported pre-operative vision assessment with Cat‑PROM5, applying either the CDA intervention or defaulting to standard care, making a shared decision about surgery, and finally documenting the outcome of the operation, including a post-operative self-reported vision difficulty assessment with Cat-PROM5.

Participants

The current feasibility study assumed recruitment of 40 participants from 4 cataract research centres (Bristol, Torbay, Brighton, Cheltenham), 10 patients each per centre. During the study however it became clear that Cheltenham would be unable to join the study due to local capacity issues and Torbay only able to join late due to staff illness. This required over-recruitment by Bristol and Brighton, with full recruitment of 42 patients none-the-less being successfully achieved.

Outcome measures

The Cataract Decision Quality Measure (CDQM) developed as a part of WP3 was used to assess patients’ decision quality. The CDQM is a measure intended to capture patient’s knowledge about options, preferences and readiness to make a decision about the treatment. It was treated as a primary outcome in this study. The CDQM questionnaire was completed twice, first before the consultation at the baseline visit and then immediately following the consultation. It included several items that were grouped in four sections: A assessed knowledge about cataracts, C readiness to make a decision, and B&D functioned together as a tool in which patients first indicated what was important to them (B) and then actually decided on the treatment (D).

A secondary quantitative outcome was Cat-PROM5, a self-reported measure of vision quality developed and validated in WP1 of the grant programme. The Cat-PROM5 questionnaire was completed by patients twice, initially at the baseline pre-operative time-point and then at the post-operative follow-up visit.

Statistical analyses

Data completeness was good for the feasibility study primary and secondary variables both before and after intervention / surgery (range 100% to 88% complete). Since the study was a feasibility study performed on a limited sample of 42 patients, statistical analyses were kept simple. Analysis included descriptive statistics with frequency analyses with chi-square tests. The effects of the intervention were analysed in two ways. For the summary scores expressed by a single value (knowledge about cataract (Section A), readiness for decision making (Section C) and Cat-PROM5) t-tests for both dependent (paired) samples and independent samples were undertaken. For the linked sections B&D, Spearman’s Rho assessed concordance between what was reported as being important and what was subsequently chosen.

To inform a sample size estimate a power study was undertaken providing calculations of sample sizes needed for a possible future fully powered RCT to investigate the impact of the CDA on the quality of patient decisions. The magnitudes of effect sizes were chosen according to Cohen’s classification of standardised effect sizes (standard deviation of unity).

Results

The intervention and standard care control groups were evenly matched at baseline.

*Primary and secondary outcomes*

This section summarises results of comparisons of primary and secondary outcomes across intervention and standard care groups, for before and after consultation / operation.

Independent t-tests showed no important differences between groups at follow up for knowledge or readiness to decide. Paired t-tests between baseline and post-consultation showed no change for knowledge in either group. Unexpectedly, readiness to make a decision declined after the consultation in the intervention group.

No significant differences were observed between the intervention and control groups at either baseline or post-operative points for Cat-PROM5 scores. As expected, significant improvements in Cat-PROM5 scores were observed between baseline and post-operative completions for both groups (paired t-tests). Despite there being no statistically significant differences in Cat-PROM5 scores post-operatively for this small sample, the score improvement in the CDA intervention group (3.40) was almost half a logit greater than in the control group (2.96).

Basic psychometric analyses were undertaken for questions in section A and C using classical test theory (CTT) to detect possibly malfunctioning questions (item to total correlations and Cronbach’s alpha). These indicated that the scale of knowledge questions (Section A) would benefit from review and further refinement (low Cronbach’s alphas and item to total correlations). Similar analyses on readiness to make a decision questions (Section C) were however encouraging.

A further aspect of the CDQM was investigation of whether using the CDA during the consultation improves the level of concordance between what is important for a patient and what they then actually choose in terms of their treatment decision. Spearman’s rho computed for indicators from sections B (4 items) and D (1 item - Choice of the treatment) Were close to zero

Power study

One of the aims of this feasibility study was to provide sample size calculations for a possible future fully powered RCT of the decision aid. The analyses of statistical power presented were conducted by the analytical approach. All computions were performed in G\*Power and summarized below in a form of a series of graphics.

For t-test for dependent samples (matched pairs) the standard deviation of the difference between measures at pre-consultation and post-consultation was standardised to 1.00 and effect sizes were set at 0.20SD as a minimum and increased in steps of 0.30SD. This enabled checking the assumed levels: minimal (0.20SD), moderate (0.50SD), and large effects (0.80SD) as recommended by Cohen, with an additional very large effect (1.10SD).

The graphic in Figure 1 indicates that for detection of a small effect (0.2SD - red line) from pre- to post-operatively a sample of 400 would provide >95% power for a two sided alpha of p=0.05. And sample 200 is needed to obtain 80% power for detection of this small effect.

The graphic in Figure 2 indicates that for t-tests for independent groups, a small effect size (0.2SD) would be detectable with 80% power by a sample size of 800 (1:1 allocation, 400 in each group). For 90% power 1050 cases in all would be needed. Larger effect sizes would be detectable with power >90% at a sample size of 200 in all.

Figure 1. Power for various sample sizes for differences between two dependent group means



Figure 2. Power for various sample sizes for differences between two independent group means.



Based on a difference of around 0.72 Logits or 0.36SD, as observed for the secondary Cat‑PROM5 outcome, the graphic in Figure 3 illustrates that for differences of this magnitude between two independent group means a total sample of 250 would be required for detection of this effect with 80% power and 325 needed for 90% power.

Figure 3. Power for various sample sizes for small to medium differences of 0.36SD between two independent group means.



Conclusions

In conclusion this study has illustrated that a fully powered randomised controlled trial of the Cataract Decision Aid (CDA) would be feasible in terms of recruitment of centres, recruitment of patient participants and sample size. The primary outcome measure, the Cataract Decision Quality Measure (CDQM) would however require further refinement in advance of a full trial.