Appendix 1. MOOSE Checklist

Reporting of background should include	
Problem definition	Introduction
Hypothesis statement	Introduction paragraph 4. "Reporting of pain outcomes in the orthopaedic literature frequently emphasises improvement in mean scores. To advise both patients and their healthcare professionals, it is important to have a clear understanding of the frequency and extent of pain following total hip or knee replacement. In the absence of appropriate clinical trials, the best way to explore this is the prospective study of unselected patients"
Description of study outcome(s)	Background paragraph 4 Methods/ Data sources and searches: disease specific patient reported outcome measures described Data synthesis and analysis
Type of exposure or intervention used	Background. Total hip or knee replacement
Type of study designs used	Methods/ Study selection. Prospective studies in consecutive/ unselected populations
Study population	Methods/ Study selection. Prospective studies in consecutive/ unselected populations
Reporting of search strategy should include	
Qualifications of searchers (eg, librarians and investigators)	Methods/ Study selection. Researchers experienced in systematic reviews and rheumatology
Search strategy, including time period included in the synthesis and keywords	Methods/ Data sources and searches, and Appendix 2
Effort to include all available studies, including contact with authors	Methods/ Data extraction and Quality assessment. We did not contact authors. Potentially, data is available not just from published studies with mean pain outcome scores. It is also available as routinely collected data. We included only published studies in representative populations with appropriate outcome data. Also considered in Discussion Methods/ Study selection.
Databases and registries searched	Methods/ Data sources and searches
Search software used, name and version, including special features used (eg, explosion)	Methods/ Data sources and searches.
Use of hand searching (eg, reference lists of obtained articles)	Methods/ Data sources and searches.
List of citations located and those excluded,	PRISMA style flow diagram shown in

including justification	Figure 1
Method of addressing articles published in languages other than English	Methods/ Data sources and searches. No exclusions on basis of language. No studies were identified that were not published in English
Method of handling abstracts and unpublished studies	Methods/ Data sources and searches. We did not include studies only published as abstracts
Description of any contact with authors	Methods/ Data extraction and Quality assessment/Discussion. We did not approach authors of studies with pain measured at follow up but not reported as proportions with degrees of pain. In recent reviews (Beswick et al. Lancet 2008, Beswick et al. Reviews in Clinical Gerontology 2010) we had additional data provided by under half of authors. Recent review by Mullan et al. 2009 suggests this is a common issue in reviews. This is considered in Discussion. Authors of studies with appropriate data but with specific missing information were contacted by email.
Reporting of methods should include	
Description of relevance or appropriateness of studies assembled for assessing the hypothesis to be tested	Results
Rationale for the selection and coding of data (eg, sound clinical principles or convenience)	Results/ Data synthesis and analysis
Documentation of how data were classified and coded (eg, multiple raters, blinding, and interrater reliability)	Results/ Study selection/ Data extraction/ and Quality assessment
Assessment of confounding (eg, comparability of cases and controls in studies where appropriate)	We identified only studies where populations were representative of the population receiving joint replacement
Assessment of study quality, including blinding of quality assessors; stratification or regression on possible predictors of study results	To assess whether -studies were representative of the joint replacement population we assessed quality of studies based on: blind outcome assessment, incompleteness of outcome data collection, and other sources of bias (representativeness of study population). These are describe in Methods/ Study quality, Appendix 3, and throughout the Results section
Assessment of heterogeneity	In Results/ Overview we have considered quality of studies as a source of heterogeneity. In Discussion paragraph 7

	we explain why the dataset is limited with regard to heterogeneity analyses.
Description of statistical methods (eg, complete description of fixed or random effects models, justification of whether the chosen models account for predictors of study results, dose-response models, or cumulative meta-analysis) in sufficient detail to be replicated	No analysis with combination was possible as described in Discussion paragraph 2.
Provision of appropriate tables and graphics	Results summarised in Figure 2 and Table 1. Also Study flow diagram in Figure1, Search strategy in Appendix 2, Quality assessments in Appendix 3 and Pain outcomes in Appendix 4.
Reporting of results should include	
Graphic summarizing individual study estimates and overall estimate	Figure 2 and Results section
Table giving descriptive information for each study included	Table 1
Results of sensitivity testing (eg, subgroup analysis)	Not possible due to range of outcome measures.
Indication of statistical uncertainty of findings	Discussed in detail in Results section and Discussion
Reporting of discussion should include	
Quantitative assessment of bias (eg, publication bias)	Risk of bias table showing quality/ representativeness of studies included as Appendix 3. Considered extensively in Results sections: we used number of study centres and losses to follow up as markers of representativeness.
Justification for exclusion (eg, exclusion of non–English-language citations)	No exclusions on the basis of language of publication.
Assessment of quality of included studies	As described in Methods/ Quality assessment we used relevant issues from te Cochrane risk of bias table. Specifically these were: blind outcome assessment, incompleteness of outcome data collection, and representativeness of the study cohort. These are then applied in detail in the Results section.
Reporting of conclusions should include	
Consideration of alternative explanations for observed results	In the Introduction paragraph 5 and Discussion paragraph 11 we consider the possibility that patients lost to follow up have different pain outcomes than those followed up.
Generalisation of the conclusions (ie, appropriate for the data presented and within the domain of the literature review)	We think that reporting the proportion of people with a poor pain outcome across the studies is the best approach. A measured speculation on outcomes of

	those lost to follow up seems appropriate in Results/ Overview.
Guidelines for future research	Discussion paragraph 12 and 13 discuss possible interventions based on determinants of good and bad outcomes.
Disclosure of funding source	Funding described