No functional benefit of larger femoral heads and alternative bearings at 6 months following primary hip replacement

Simon S JAMESON[1,2], James M MASON[2], Paul N BAKER[1], Paul J GREGG[1], David J DEEHAN[3], and Mike R REED[4]

[1]The National Joint Registry for England and Wales; [2]School of Medicine, Pharmacy and Health, Durham University, Queen's Campus, Stockton-on-Tees; [3]Newcastle Hospitals NHS Foundation Trust, Newcastle upon Tyne; [4]Northumbria Healthcare NHS Foundation Trust, Ashington, Northumberland, UK.
Correspondence: simonjameson@doctors.org.uk
Submitted 2013-12-10. Accepted 2014-05-19

Background and purpose — There has been a recent trend towards the use of greater femoral head sizes in an attempt to improve function and enhance stability after primary hip replacement. This has been associated with the use of alternative bearings, theoretically to reduce wear and improve implant longevity.

Methods — We examined the influence of these variables on patient-reported outcome measures (PROMs) for a consecutive series of primary hip replacements using National Joint Registry (NJR) and PROMs-linked data. To minimize the confounding influence of implant design factors, the single most commonly used brand in England and Wales (DePuy Corail Pinnacle) was examined. Improvement in patient hip-specific outcomes (Oxford hip score, OHS), general health outcomes (Euroqol, EQ-5D), and rates of self-reported complications (bleeding, wound problems, readmission, and reoperation) were compared for different head sizes (28-mm, 32-mm, and 36-mm) and bearings (metal-on-polyethylene (MoP), ceramic-on-polyethylene (CoP), and ceramic-on-ceramic (CoC)), adjusting for differences in case mix.

Results — At a mean follow-up of 7 months, improvements in OHS and EQ5D index were similar for 28-mm and 36-mm heads. A 32-mm head was associated with poorer function (OHS: 20, 99% CI: 19–21, p = 0.002; EQ5D index: 0.39, 99% CI: 0.36–0.42, p = 0.004), although these small differences may not be of clinical importance. There were no statistically significant benefits of either CoP or CoC bearings compared to a MoP bearing. Complication rates were similar within comparisons of head sizes or bearings.

Interpretation — In this short-term study, we did not find any functional benefits of larger head sizes or alternative bearings, after adjusting for other influences. We question their use in routine primary hip replacement given the lack of evidence of improved long-term survival in the literature.

Greater femoral head size may improve function and enhance stability after primary total hip replacement (THR) (Bartz et al. 2000, Cuckler et al. 2004, Hummel et al. 2009). Previous studies have shown a greater range of movement with increasing head size (Amstutz et al. 1975, Matsushita et al. 2009). Use of a larger head size is an attractive option in younger patients who require stability at higher levels of function, and in older patients in order to reduce dislocation risk. However, greater surface area may also increase wear rates, irrespective of bearing materials (Charnley et al. 1969, Dowling et al. 1978, Livermore et al. 1990, Bragdon et al. 2013, Jack et al. 2013), and there have been reports of excessive taper load with large-diameter bearings (Langton et al. 2012, Meyer et al. 2012). Larger heads have been associated with the use of alternative bearings, in order to reduce wear and improve implant longevity. The National Joint Registry (NJR) in England and Wales has described an increase in the use of larger femoral head sizes (over 28 mm)—from 5% to 50% between 2005 and 2010 (England and Wales National Joint Registry 2012). Over the same period, the use of ceramic-on-polyethylene (CoP) and ceramic-on-ceramic (CoC) bearings has increased.

Medium-term revision rates are higher with CoC bearings than with metal-on-polyethylene (MoP) bearings generally, across registry data, and specifically, when the most commonly used implant in England and Wales (Corail stem/Pinnacle cup; DePuy Ltd., Leeds, UK) was analyzed (Sexton et al. 2009, Jameson et al. 2013). Larger femoral sizes using hard-bearing technology did not give any functional improvement over 28-mm MoP (Hanna et al. 2012) in a small randomized trial, and larger head sizes have not been found to offer any gait-related benefits (Zagra et al. 2013). The functional benefits of increasing head size and alternative bearings have yet to fully assessed.

DOI 10.3109/17453674.2014.972259

Patient-reported outcome measures (PROMs) supplement revision risk in the assessment of success after joint replacement (Devlin et al. 2010). PROMs are routinely collected on National Health Service (NHS) patients undergoing THR in England. Data on hip replacement patients, their surgeons, and the implants used are collected by the NJR. These datasets can be linked in order to compare early outcomes for specific patient and implant groups at the national level. The present analysis explored the effect of bearing surface and femoral head size on PROMs and complications following THR. We hypothesized that larger heads and alternative bearings would have no functional benefit over standard (28-mm MoP) bearings.

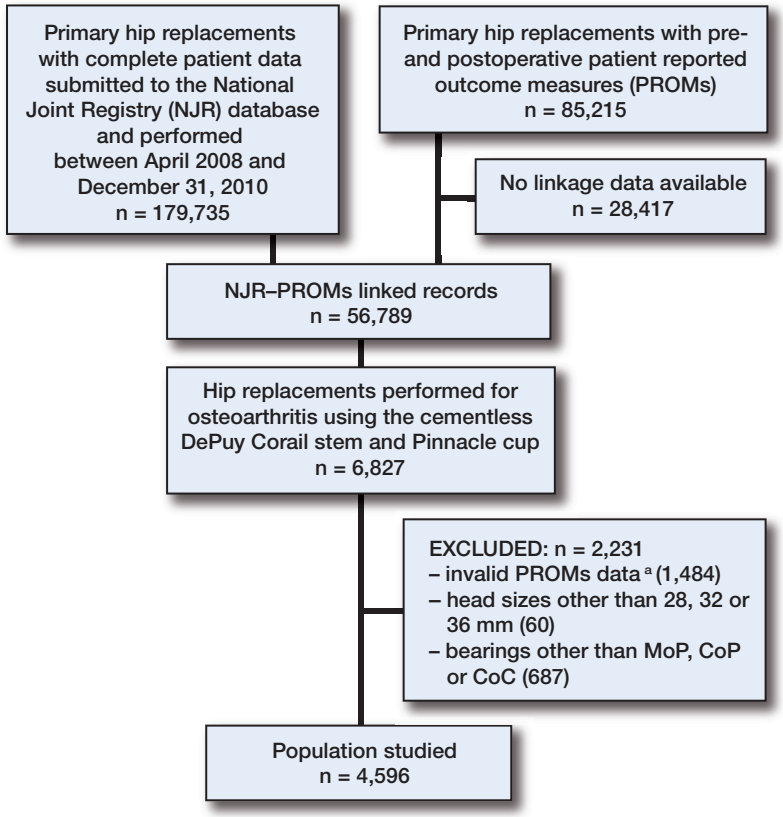

Figure 1. Flow chart describing the study cohort. MoP – metal-on-polyethylene, CoP – ceramic-on-polyethylene, CoC –ceramic-on-ceramic. [a] Invalid PROMs data includes records with missing outcome score rccords, preoperative scores dated more than 12 months prior to operation, postoperative records without a date or dated < 6 months or > 12 months following the primary hip replacement, non-identical duplicates (all excluded) and identical duplicates (only one record retained).

Material and methods

Design

We conducted a cohort study using prospectively collected patient-level NJR and PROMs-linked data to compare outcome scores and self-reported complications after primary THR for different head sizes and different bearings.

Data

The national PROMs study collects joint-specific and general health scores preoperatively and around 6 months postoperatively, and self-reported complications. By linking databases at the patient level, PROMs data can be combined with the corresponding demographic and operative details held in the NJR. To link them, a number of criteria were used: firstly, to ensure correct matching, 2 unique identifiers (NJR and procedure numbers) recorded in both datasets were used; secondly, the operation date recorded by the patient in the PROMs data had to be within ± 30 days of the operation date recorded on the NJR record, to ensure that the patient was scoring the same procedure.

We chose to perform the analysis using the single most commonly used brand of THR used in England and Wales (Corail stem/Pinnacle cup; DePuy Ltd., Leeds, UK), in order to control for any implant-related influences (England and Wales National Joint Registry 2012).

There were a number of exclusion criteria. For the NJR data, these were: all procedures with an indication other than OA, procedures with missing implant or patient data, head sizes smaller than 28 mm and larger than 36 mm, and rarely used bearings (ceramic-on-metal and metal-on-ceramic). Metal-on-metal bearings were also excluded, as few of these are now implanted (England and Wales National Joint Registry 2012). Procedures with PROMs data that were missing, undated, dated more than 12 months prior to or following the operation, or that had non-identical duplicates were excluded; for identical duplicates, the first record was retained for analysis. Where the presence of a co-morbidity or complication was asked for in the questionnaire but left blank by the patient, it was assumed to be absent. The study population is summarized in Figure 1. The PROMs project was introduced in April 2008. Linkage of NJR data to PROMs was possible between this date and December 2010 (the limit of our access to NJR data). Details regarding the delivery and return of PROMs questionnaires were not available for this study. The demographic, surgical, and implant-related variables available for analysis are listed in Table 1 (see Supplementary data).

The national PROMs project uses validated measures of hip-specific outcomes (Oxford hip score (OHS)) (Dawson et al. 1996) and general health outcomes (EuroQol (EQ-5D)) (EuroQol group 2009). For this analysis, the outcomes of interest were improvements between preoperative and postoperative scores (the "change scores") and self-reported postoperative complications (bleeding, wound problems, re-admission, and reoperation). Change scores, being approximately normally distributed, are analytically preferable to postoperative scores (Browne et al. 2007). The OHS (score 0–48) has

previously been shown to be a reliable, valid, and responsive outcome measure (Murray et al. 2007). A threshold of 3 points has been proposed to demonstrate a clinically important difference (Murray et al. 2007). The EQ-5D index (score –0.59 to 1.00) is a generic measure of health used for clinical and economic appraisal. It evaluates 5 different aspects of general health (mobility, self-care, usual activities, pain/discomfort, and anxiety/depression) that are scored and combined using population weightings to produce a single index value for health status (group 2009). Patients are also asked about comorbidities, general health, and self-reported disability as part of the preoperative PROMs questionnaire. These can be used to adjust for differences in health status between patient groups. Sample sizes for all the head-size and bearing groups were in excess of the minimum numbers identified in the PROMs feasibility pilot for identification of meaningful differences (more than 150 per group) (Browne et al. 2007).

Patients are also asked to indicate their satisfaction with the outcome following surgery (excellent, very good, good, fair, or poor), and whether they deem surgery to have been a success (much better, a little better, about the same, a little worse, or much worse). While unadjusted values have been provided for information, no attempt was made to adjust for baseline differences in either success or satisfaction, as previous analyses have demonstrated that the variables available within the NJR and PROMs databases are insufficient to explain these differences (i.e. the influence of unmeasured variables has a greater effect than the effect of the measured variables) (Browne et al. 2007, Hamilton et al. 2013).

Statistics

The variables available for the analyses are shown in Table 1 (see Supplementary data). The head sizes analyzed were 28 mm, 32 mm, and 36 mm. The bearings were MoP, CoP, and CoC. Differences in baseline characteristics across the groups would be a source of confounding in any comparative analysis. Thus, to test the hypothesis that there were no differences between groups, the following tests were employed: 1-way analysis of variance (ANOVA, parametric continuous data variables), the Kruskal-Wallis test (non-parametric continuous data variables), or the chi-square test (categorical data variables).

Univariable analysis was performed initially to identify variables that possibly influenced each outcome, based on statistical rejection criteria of $p > 0.1$; these variables were then included in the multivariable models. Analysis of covariance (ANCOVA) was used for testing of differences in OHS and EQ5D index change scores across head size and bearing groups. Multivariable logistic regression was used to analyze differences in the risk of each of the complications across groups. Time from implantation to questionnaire completion was included in models to evaluate whether differences in duration of follow-up influenced findings. Preoperative scores were included in all models, as recommended by the designers of the Oxford hip score (Murray et al. 2007).

Reflecting analysis of a large dataset, statistical models for the change scores were evaluated with the margins function in STATA in order to provide predicted values (including 99% confidence intervals (CIs)) separately for each of the head-size and bearing categories. p-values are provided as statistical evaluation of the differences between the reference (head size of 28 mm and MoP bearing) and other variables within the category. For complication risks, results are presented as odds ratios (ORs) with CIs: ratios greater than 1 indicate that risk is higher when compared with the reference category. Due to the statistical methods employed and the large population size, only covariates fitting models with $p < 0.01$ were retained in final models as significant influences, to reduce the risk of type-1 error. All models were fitted using STATA 12.

For more information on the statistical techniques used, see Supplementary data.

Ethics

Explicit patient consent was taken for both the NJR and PROMs data collection. Further ethics approval is not required for registry studies in the UK.

Results

There were 4,596 NJR-PROMs linked primary procedures. MoP accounted for 47% (2,171), CoC for 45% (2,064), and CoP for 7.9% (361). A standard (28-mm) head size was used in 41% of procedures (1,864), the 36-mm in 41% (1,863), and the 32-mm in 19% (869). When the demographics were compared across bearing groups, patients with a CoC bearing were generally younger and in better health, but there were comparable numbers of women in each group and mean BMI was similar. Patients who had a larger head size implanted were generally younger and in better health (Table 2). Patients fitted with a 32-mm head generally had a higher ASA grade and higher BMI, and poorer general health. Although there were statistically significant differences between categories of bearing and head size for each of the surgical covariates as a result of the large study population, the groups were qualitatively similar and therefore broadly comparable (Tables 2 and 3).

In the unadjusted PROMs data, patients with a CoC bearing and a larger head size generally had higher preoperative and postoperative OHS and EQ5D indices. Patient-reported levels of satisfaction and success were similar across the bearing groups and the 28-mm and 36-mm head size groups, although the 32-mm head size group had poorer scores (Tables 4 and 5).

Improvement in outcome score

Improvements in OHS and EQ5D index were similar for MoP, CoP, and CoC bearings (Table 6, see Supplementary data).

Improvements in OHS and EQ5D index were similar for a 28-mm and a 36-mm head. Although the difference was small,

Table 2. Patient demographics for the population studied, by head size

	All patients	Head size 28 mm (Reference)	32 mm	36 mm	p-value [a]
Number (%)	4,596	1,864 (40.6)	869 (18.9)	1,863 (40.5)	
Patient factors					
Age, mean (SD) range	66.0 (9.7) 25.2–95.1	68.0 (8.9) 31.0–95.1	68.2 (10.1) 26.2– 94.1	62.9 (9.5) 25.2–90.2	< 0.001
Females, n (%)	2,620 (57.0)	1,303 (69.9)	535 (61.6)	782 (42.0)	< 0.001
ASA, n (%)					< 0.001
1	804 (17.5)	286 (15.3)	93 (10.7)	425 (22.8)	
2	3,360 (73.1)	1,410 (75.6)	644 (74.1)	1,306 (70.1)	
3+	432 (9.4)	168 (9.0)	132 (15.2)	132 (7.1)	
BMI (SD) range [b]	29.0 (5.4) 15–65	28.7 (5.2) 15–56	29.1 (5.6) 16–50	29.5 (5.4) 16–65	0.03
Comorbidities, n (%)					
Heart disease	388 (8.4)	152 (8.2)	91 (10.5)	145 (7.8)	0.05
Stroke	67 (1.5)	29 (1.6)	13 (1.5)	25 (1.3)	0.9
Diabetes	355 (7.7)	158 (8.5)	70 (8.1)	127 (6.8)	0.2
Hypertension	1,764 (38.4)	739 (39.7)	366 (42.1)	659 (35.4)	0.001
Circulation	221 (4.8)	102 (5.5)	45 (5.2)	74 (4.0)	0.09
Lung	270 (5.9)	112 (6.0)	42 (4.8)	116 (6.2)	0.3
Depression	309 (6.7)	126 (6.8)	44 (5.1)	139 (6.7)	0.07
Preoperative general health					0.02
Excellent	265 (5.8)	100 (5.4)	53 (6.1)	114 (6.1)	
Very good	1,413 (30.7)	566 (30.4)	235 (27.0)	611 (32.8)	
Good	2,008 (43.7)	862 (46.2)	394 (45.3)	755 (40.5)	
Fair	781 (17.0)	285 (15.3)	165 (19.0)	330 (17.7)	
Poor	129 (2.9)	51 (2.7)	22 (2.5)	53 (2.8)	
Preoperative disability	2,229 (48.5)	928 (50.0)	448 (51.6)	853 (45.8)	0.007
Surgical factors					
Chemical VTE prophylaxis					< 0.001
LMWH only	2,583 (56.2)	1,115 (59.8)	518 (59.6)	950 (51.0)	
Aspirin only	415 (9.0)	146 (7.8)	50 (5.8)	219 (11.8)	
Other	544 (11.8)	163 (8.7)	98 (11.3)	283 (15.2)	
None	304 (6.6)	133 (7.1)	36 (4.1)	135 (7.3)	
Not recorded	750 (16.3)	307 (16.5)	167 (19.2)	276 (14.8)	
Mechanical VTE prophylaxis					< 0.001
Compression stockings (CS)	1,249 (27.2)	500 (26.8)	222 (25.6)	527 (28.3)	
CS + mechanical pump	918 (20.0)	296 (15.9)	134 (15.4)	488 (26.2)	
Foot pump only	523 (11.4)	278 (14.9)	86 (9.9)	159 (8.5)	
Mechanical calf pump only	756 (16.5)	312 (16.7)	173 (19.9)	271 (14.6)	
Other	22 (0.5)	7 (0.4)	0 (0.0)	15 (1.0)	
None	378 (8.2)	164 (8.8)	87 (6.8)	127 (6.8)	
Not recorded	750 (16.3)	307 (16.5)	167 (19.2)	276 (14.8)	
Bearing					< 0.001
Metal-on-polyethylene	2,171 (47.2)	1,423 (76.3)	510 (58.7)	238 (12.8)	
Ceramic-on-polyethylene	361 (7.9)	198 (10.6)	100 (11.5)	63 (3.4)	
Ceramic-on-ceramic	2,064 (44.9)	243 (13.0)	259 (29.8)	1,562 (83.8)	
Anesthesia					< 0.001
Regional	2,077 (45.2)	871 (46.7)	303 (34.9)	903 (48.5)	
General	827 (18.0)	312 (16.7)	168 (19.3)	347 (18.6)	
Regional and general	906 (19.7)	357 (19.2)	224 (25.8)	325 (17.4)	
Not recorded	786 (17.1)	324 (17.4)	174 (20.0)	288 (15.5)	
Lead surgeon grade					< 0.001
Consultant	3,475 (75.6)	1,484 (79.6)	618 (71.1)	1,373 (73.7)	
Other	1,121 (24.4)	380 (20.4)	251 (28.9)	490 (26.3)	
Position					0.02
Lateral	3,598 (78.3)	1,449 (77.7)	666 (76.6)	1,483 (79.6)	
Supine	248 (5.4)	108 (5.8)	36 (4.1)	104 (5.6)	
Not recorded	750 (16.3)	307 (16.5)	167 (19.2)	276 (14.8)	

ASA: American Society of Anaesthesiologists score; PROMs: patient-reported outcomes measures.
[a] Difference between groups with 1-way ANOVA (continuous data variables) or chi-squared test (categorical data variables).
[b] BMI (body mass index) data on 2,726.

the 32-mm head was associated with a poorer outcome (OHS: 20, 99% CI: 19–21, p = 0.002; EQ5D index: 0.39, 99% CI: 0.36–0.42, p = 0.004) (Table 7, see Supplementary data).

Risk of complications

There were no statistically significant differences in risk of

Table 3. Patient demographics for the population studied, by bearing

	All patients	Bearing: Metal-on-polyethylene (Reference)	Ceramic-on-polyethylene	Ceramic-on-ceramic	p-value [a]
Number (%)	4,596	2,171 (47.2)	361 (7.9)	2,064 (44.9)	
Patient factors					
Age, mean (SD) range	66.0 (9.7) 25.2–95.1	70.8 (8.0) 31.0–95.1	64.6 (8.3) 39.0–87.2	61.1 (9.1) 25.2–91.5	< 0.001
Females, n (%)	2,620 (57.0)	1,259 (58.0)	220 (60.9)	1,141 (55.3)	0.06
ASA, n (%)					< 0.001
1	804 (17.5)	246 (11.3)	64 (17.7)	494 (23.9)	
2	3,360 (73.1)	1,673 (77.1)	249 (69.0)	1,438 (69.7)	
3+	432 (9.4)	252 (11.6)	48 (13.3)	132 (6.4)	
BMI (SD) range [b]	29.0 (5.4) 15–65	28.7 (5.2) 15–56	29.3 (5.3) 18–54	29.2 (5.6) 18–65	0.09
Comorbidities, n (%)					
Heart disease	388 (8.4)	233 (10.7)	28 (7.8)	127 (6.2)	< 0.001
Stroke	67 (1.5)	42 (1.9)	4 (1.1)	21 (1.0)	0.04
Diabetes	355 (7.7)	201 (9.3)	27 (7.5)	127 (6.2)	0.001
Hypertension	1,764 (38.4)	960 (44.2)	136 (37.7)	668 (32.4)	< 0.001
Circulation	221 (4.8)	132 (6.1)	25 (6.9)	64 (3.1)	< 0.001
Lung	270 (5.9)	122 (5.6)	21 (5.8)	127 (6.2)	0.8
Depression	309 (6.7)	123 (5.7)	25 (6.9)	161 (7.8)	0.02
Preoperative general health					0.001
Excellent	265 (5.8)	104 (4.8)	20 (5.5)	137 (6.6)	
Very good	1,413 (30.7)	649 (29.9)	93 (25.8)	669 (32.4)	
Good	2,008 (43.7)	1,004 (46.2)	184 (51.0)	819 (39.7)	
Fair	781 (17.0)	363 (16.7)	53 (14.7)	380 (18.4)	
Poor	129 (2.9)	51 (2.3)	11 (3.0)	59 (2.9)	
Preoperative disability	2,229 (48.5)	1,117 (51.5)	192 (53.2)	920 (44.6)	< 0.001
Surgical factors					
Chemical VTE prophylaxis					< 0.001
LMWH only	2,583 (56.2)	1,322 (60.9)	211 (49.0)	1,050 (50.9)	
Aspirin only	415 (9.0)	179 (8.3)	21 (5.8)	215 (10.4)	
Other	544 (11.8)	189 (8.7)	45 (12.5)	310 (15.0)	
None	304 (6.6)	115 (5.3)	27 (7.5)	62 (7.8)	
Not recorded	750 (16.3)	366 (16.9)	57 (15.8)	327 (15.8)	
Mechanical VTE prophylaxis					< 0.001
Compression stockings (CS)	1,249 (27.2)	644 (29.7)	98 (27.2)	507 (24.6)	
CS + mechanical pump	918 (20.0)	342 (15.8)	79 (21.9)	497 (24.1)	
Foot pump only	523 (11.4)	324 (14.9)	10 (2.8)	189 (9.2)	
Mechanical calf pump only	756 (16.5)	311 (14.3)	70 (19.4)	375 (18.2)	
Other	22 (0.5)	4 (0.2)	1 (0.3)	17 (0.8)	
None	378 (8.2)	180 (8.3)	46 (12.7)	152 (7.4)	
Not recorded	750 (16.3)	366 (16.9)	57 (15.8)	327 (15.4)	
Head size					< 0.001
28 mm	1,864 (40.6)	1,423 (65.6)	198 (54.9)	243 (11.8)	
32 mm	869 (18.9)	510 (23.5)	100 (27.7)	869 (18.9)	
36 mm	1,863 (40.5)	238 (11.0)	63 (17.5)	1,562 (40.5)	
Anesthesia					0.01
Regional	2,077 (45.2)	987 (45.5)	153 (42.4)	937 (45.4)	
General	827 (18.0)	344 (15.9)	72 (19.9)	411 (19.9)	
Regional and general	906 (19.7)	454 (20.9)	77 (18.2)	375 (18.2)	
Not recorded	786 (17.1)	386 (17.8)	59 (16.3)	341 (16.5)	
Lead surgeon grade					0.002
Consultant	3,475 (75.6)	1,600 (73.7)	294 (81.4)	1,581 (76.6)	
Other	1,121 (24.4)	571 (26.3)	67 (18.6)	483 (23.4)	
Position					0.1
Lateral	3,598 (78.3)	1,708 (78.7)	284 (78.7)	1,606 (77.8)	
Supine	248 (5.4)	97 (4.5)	20 (5.5)	131 (6.4)	
Not recorded	750 (16.3)	366 (16.9)	57 (15.8)	327 (15.8)	

ASA: American Society of Anaesthesiologists score; PROMs: patient-reported outcomes measures.
[a] Difference between groups with 1-way ANOVA (continuous data variables) or chi-squared test (categorical data variables).
[b] BMI (body mass index) data on 2,726.

bleeding, wound complications, re-admission, or reoperation after case-mix adjustment when the bearing groups were compared (Table 8, see Supplementary data). Bleeding risk was higher in the 32-mm head-size group (OR = 1.8, 99% CI: 1.2–2.8, $p < 0.001$) but not in the 36-mm head-size group (OR 1.3, CI 0.9 to 2.0, 0.063) when compared with the 28-mm group.

Table 4. Patient-reported outcomes for populations studied, by bearing

	Metal-on-polyethylene	Ceramic-on-polyethylene	Ceramic-on-ceramic	p-value [a]
Number (%)	2,171 (47.2)	361 (7.9)	2,064 (44.9)	
Oxford hip score				
Preoperative, mean (SD) range	18.4 (8.0) 1–43	18.5 (8.4) 0–43	19.2 (8.1) 1–46	0.003
Postoperative, median (range)	42 (0–48)	41 (7–48)	43 (2–48)	< 0.001
EQ5D index				
Preoperative, mean (SD)	0.368 (0.314)	0.375 (0.331)	0.394 (0.311)	0.01
range	-0.349 to 1	-0.484 to 0.883	-0.594 to 1	
Postoperative, median (range)	0.815 (-0.319 to 1)	0.796 (-0.319 to 1)	0.848 (-0.594 to 1)	0.008
Satisfaction, n (%)				0.3
Excellent	882 (40.6)	144 (39.9)	887 (43.0)	
Very good	804 (37.0)	123 (34.1)	728 (35.3)	
Good	364 (16.8)	66 (18.3)	321 (15.6)	
Fair	90 (4.2)	23 (6.4)	89 (4.3)	
Poor	31 (1.4)	5 (1.4)	39 (1.9)	
Success, n (%)				0.05
Much better	1,936 (89.2)	303 (83.9)	1,828 (88.6)	
A little better	157 (7.2)	46 (12.7)	158 (7.7)	
About the same	33 (1.5)	3 (0.8)	35 (1.7)	
A little worse	26 (1.2)	6 (1.7)	29 (1.4)	
Much worse	19 (0.9)	3 (0.8)	14 (0.7)	
Time from operation to PROMs completion, mean days (SD) range	208 (28) 183–363	207 (23) 183–333	210 (29) 183–361	0.002

PROMs: patient-reported outcome measures.
[a] 1-way ANOVA was used for parametric data, Kruskal-Wallis test for non-parametric data, and chi-squared test for proportions.

Table 5. Patient-reported outcomes for populations studied, by head size

	28 mm	32 mm	36 mm	p-value [a]
Number (%)	1,864 (40.6)	869 (18.9)	1,863 (40.5)	
Oxford hip score				
Preoperative, mean (SD) range	18.2 (8.0) 1–44	18.3 (7.9) 2–43	19.6 (8.2) 0–46	< 0.001
Postoperative, median (range)	43 (0–48)	41 (4–48)	48 (2–48)	< 0.001
EQ5D index				
Preoperative, mean (SD)	0.366 (0.315)	0.365 (0.316)	0.401 (0.311)	0.001
range	-0.484 to 1	-0.319 to 1	-0.594 to 1	
Postoperative, median (range)	0.848 (-0.319 to 1)	0.796 (-0.126 to 1)	0.850 (-0.594 to 1)	< 0.001
Satisfaction, n (%)				< 0.001
Excellent	791 (42.4)	319 (36.7)	803 (43.1)	
Very good	683 (36.6)	299 (34.4)	673 (36.1)	
Good	288 (15.5)	190 (21.9)	273 (14.7)	
Fair	77 (4.1)	46 (5.3)	79 (4.2)	
Poor	25 (1.3)	15 (1.7)	35 (1.9)	
Success, n (%)				0.001
Much better	1,673 (89.8)	736 (84.7)	1,658 (89.0)	
A little better	124 (6.7)	103 (11.9)	134 (7.2)	
About the same	27 (1.5)	14 (1.6)	30 (1.6)	
A little worse	22 (1.2)	10 (1.2)	29 (1.6)	
Much worse	18 (1.0)	6 (0.7)	12 (0.6)	
Time from operation to PROMs completion, mean days (SD) range	207 (26) 183–361	209 (28) 184–357	210 (29) 183–363	0.002

PROMs: patient-reported outcome measures.
[a] 1-way ANOVA was used for parametric data, Kruskal-Wallis test for non-parametric data, and chi-squared test for proportions.

Risk of wound complications was higher in the 36-mm group (OR = 1.7, 99% CI: 1.1–2.6, p = 0.002). Re-admission and reoperation risks were similar when the different head sizes were compared (Table 9, see Supplementary data).

Discussion

This large cohort study using NJR-PROMs linked data from a single hip system showed no functional benefit when femoral

head sizes greater than 28 mm and bearings other than MoP were used for primary THR. These findings are important for clinicians attempting to determine the most suitable hip implants for patients with osteoarthritis.

While this is the largest study to date to report the effects of head size and bearing type on functional outcome for a single hip replacement brand, there are some potential limitations to our findings. The study design was observational and thus vulnerable to omitted variables, which may have confounded the findings. Some data were unavailable for analysis; for example, radiological data on cup positioning. There were also large numbers of procedures that could not be analyzed, either because the datasets could not be linked (for example, due to missing patient identification details) or because there were specific data fields with incomplete data. Moreover, any bias in completion and return of PROMs cannot be assessed, as no details were available regarding the number of questionnaires sent out or returned. Non-responders are more likely to be younger patients (and therefore more likely to have larger head sizes and alternative bearings). However, previous PROMs analyses have demonstrated that non-responders generally perform more poorly (Ostendorf et al. 2004). Irrespective of these issues, similarities between the unadjusted and adjusted models and robustness under different model fitting assumptions support the stability of estimates.

By restricting the implants to the most commonly used brand only, we were able to remove the problems associated with adjusting for multiple brands (differing implant design characteristics and bearing surface manufacturing processes). The Corail Pinnacle constitutes 14% of all hip implants used in England and Wales since 2003. Cementless components now predominate, and around 40% of those implanted in 2011 were Corail Pinnacle (England-and-Wales-National-Joint-Registry 2012). [Ann-Britt: v.v. ta bort bindestrecken] This implant combination also offers a wide range of bearing options and head sizes, making it a good choice for analysis, and sample sizes of each of the comparison groups were in excess of the minimum numbers required to identify meaningful differences (Browne et al. 2007). False-negative results are therefore unlikely, especially as OHS and EQ5D indices for each of the groups were qualitatively similar, irrespective of confidence intervals.

As with all NJR-PROMs studies, the design of the study was constrained by the data available. In particular, this limited the length of follow-up available and the measures of hip function used. Although functional outcome at 6 months can be described as early, the follow-up in this study was adequate for group comparisons. The literature shows little improvement in change in OHS between 6 months and 5 years following hip replacement, suggesting that the results in our short-term study are a reliable indication of longer-term outcome (Andrew et al. 2008, Judge et al. 2013).

Greater femoral head size may reduce dislocation risk. In a randomized trial of patients undergoing metal-on-highly-crosslinked polyethylene hip replacements, Howie et al. (2012) found that a head size of 36 mm reduced the 1-year dislocation risk compared to 28 mm in 533 primary procedures—from 4.4% to 0.8%. In addition, a large cohort study of over 240,000 THRs performed in England and Wales found a reduction in 1-year dislocation risk from 1.4% to 1.1% over a period of 5 years, during which the use of large femoral head sizes increased. However, there was no change in the 12-month revision rate (Jameson et al. 2011). Lower dislocation rates with larger femoral head sizes have also been found in Australian and Norwegian registry data (Bystrom et al. 2003, ANJRR 2013). Without radiographic data, dislocation risk is difficult to analyze. The 1-year risk in the study by Howie et al. (2012) is high compared to others; Stroh et al. (2013) reported dislocation with head sizes less than 36 mm of only 1.8% (10 in 559) at 5 years, which is similar to the English NHS data (taking into account that most dislocations occur in the first year after surgery) (Jameson et al. 2011). Before larger head sizes are recommended, the incidence of late dislocation, wear, periprosthetic osteolysis, and liner fracture should also be established (Howie et al. 2012). Importantly, revision risk did not decrease in the English NHS data, suggesting that most early dislocations are relatively benign and do not require a revision procedure. Although dislocation data were unavailable in the current study, other outcomes were used as surrogate endpoints. The majority of dislocations following primary THR occur within the first 5 weeks after surgery (Bourne and Mehin 2004); therefore, patient-reported complications associated with a dislocation, such as re-admission and reoperation, should demonstrate a difference if, for example, one head size had a lower risk than another. Increasing femoral head size may also increase range of movement, but several studies have shown that it is the patient's bony anatomy that causes impingement, and therefore limits this range irrespective of head size (Bunn et al. 2012, Klingenstein et al. 2012). Interestingly, PROMs were poorer and bleeding risk was higher in patients implanted with a 32-mm head. This may be explained by the patient demographics, which showed a slightly greater mean BMI and relatively more patients with higher ASA grade and poorer health. In these patients, a surgeon may choose to increase head size from 28 mm in an attempt to increase stability, thereby potentially reducing a need for further revision surgery. Differences in unmeasured patient variables across groups (which cannot be controlled) may also account for this finding. However, it must be stressed that the statistically significant OHS change difference of 1.4 between 28-mm and 32-mm femoral heads is unlikely to be clinically significant.

Medium- to long-term risk of revision is lowest in MoP bearings across worldwide registries and published trials, and no functional benefit of alternative bearings has so far been found (Sedrakyan et al. 2011). When implant survival of 35,386 Corail Pinnacle THRs was examined in the NJR cohort, hard bearings had a higher revision risk than MoP (MoM: HR =

1.9, p < 0.001; CoC: HR = 1.6, p = 0.003) (Jameson et al. 2013). In a randomized trial of 49 patients who received either large-head MoM articulations or standard (28- to 32-mm) MoP THRs, no statistically significant benefits were found in function, dislocation, or implant failure (Hanna et al. 2012), although the numbers are small and this could be a consequence of type-II statistical error.

Metal-on-metal bearings were not included in this study. Although a large number of Corail Pinnacle THRs have been implanted in England and Wales with these bearings, their use is now low due to high revision rates and concerns regarding the systemic effects of excessive metal wear debris.

In summary, no functional benefits of larger head sizes or alternative bearings were found in this national analysis of NJR-PROMs linked THRs, after adjusting for other influences. Given the lack of evidence in the literature of improved implant survival with these implant options, we question their use in routine primary THR.

Supplementary data

Tables 1 and 6–9 are available at Acta's website (www.actaorthop.org), identification number 7059.

We thank the patients and staff of all the hospitals in England and Wales who have contributed data to the National Joint Registry. We are grateful to the Healthcare Quality Improvement Partnership (HQIP), the NJR Steering Committee, and the staff of the NJR center for facilitating this work.

The National Joint Registry for England and Wales is funded through a levy raised on the sale of hip and knee replacement implants. The cost of the levy is set by the NJR Steering Committee. The NJR Steering Committee is responsible for data collection. This work was funded by a fellowship from the National Joint Registry. The authors have conformed to the standard protocol of the NJR regarding data access and publication. The views expressed are those of the authors and do not necessarily reflect those of the National Joint Register Steering Committee or the Health Quality Improvement Partnership (HQIP), who do not vouch for how the information is presented.

No benefits in any form have been received or will be received from a commercial party related directly or indirectly to the subject of this article.

SJ developed the idea, analyzed the data, and mainly wrote the paper. JM provided extensive advice on data analysis and co-wrote the paper. PB co-developed the idea, assisted with statistical issues, and edited the paper. PG developed the idea and edited the paper. DD co-developed the idea and edited the paper. MR co-developed the idea, co-wrote the paper, and supervised the project.

Amstutz H C, Lodwig R M, Schurman D J, Hodgson A G. Range of motion studies for total hip replacements. A comparative study with a new experimental apparatus. Clin Orthop Relat Res 1975; (111): 124-30.

Andrew J G, Palan J, Kurup H V, Gibson P, Murray D W, Beard D J. Obesity in total hip replacement. J Bone Joint Surg Br 2008; 90 (4): 424-9.

Bartz R L, Nobel P C, Kadakia N R,Tullos H S. The effect of femoral component head size on posterior dislocation of the artificial hip joint. J Bone Joint Surg Am 2000; 82 (9): 1300-7.

Bourne R B, Mehin R. The dislocating hip: what to do, what to do. J Arthroplasty 2004; 19 (4 Suppl 1): 111-4.

Bragdon C R, Doerner M, Martell J, Jarrett B, Palm H, Malchau H. The 2012 John Charnley Award: Clinical multicenter studies of the wear performance of highly crosslinked remelted polyethylene in THA. Clin Orthop Relat Res 2013; 471 (2): 393-402.

Browne J, Jamieson L, Lewsey J. "Patient reported outcome measures (PROMs) in elective surgery: report to the Department of Health, 2007." 2007. Retrieved 29/08/2012, from http:// www.lshtm.ac.uk/php/hsrp/research/proms_report_12_dec_07.pdf.

Bunn A, Colwell C W, Jr., D'Lima D D. Bony impingement limits design-related increases in hip range of motion. Clin Orthop Relat Res 2012; 470 (2): 418-27.

Bystrom S, Espehaug B, Furnes O, Havelin L I. Femoral head size is a risk factor for total hip luxation: a study of 42,987 primary hip arthroplasties from the Norwegian Arthroplasty Register. Acta Orthop Scand 2003; 74 (5): 514-24.

Charnley J, Kamangar A, Longfield M D. The optimum size of prosthetic heads in relation to the wear of plastic sockets in total replacement of the hip. Med Biol Eng 1969; 7 (1): 31-9.

Cuckler J M, Moore K D, Lombardi A V, Jr., McPherson E, Emerson R. Large versus small femoral heads in metal-on-metal total hip arthroplasty. J Arthroplasty 2004; 19 (8 Suppl 3): 41-4.

Dawson J, Fitzpatrick R, Carr A, Murray D. Questionnaire on the perceptions of patients about total hip replacement. J Bone Joint Surg Br 1996; 78 (2): 185-90.

Devlin N J, Parkin D, Browne J. Patient-reported outcome measures in the NHS: new methods for analysing and reporting EQ-5D data. Health Econ 2010; 19 (8): 886-905.

Dowling J M, Atkinson J R, Dowson D, Charnley J. The characteristics of acetabular cups worn in the human body. J Bone Joint Surg Br 1978; 60-B (3): 375-82.

England and Wales National Joint Registry. "National Joint Registry for England and Wales 9th Annual Report." 2012. from http://www.njrcentre.org.uk/njrcentre/Portals/0/Documents/England/Reports/9th_annual_report/NJR%209th%20Annual%20Report%202012.pdf.

group E. "EuroQol (EQ5D Score)." 2009. Retrieved 29/08/2012, from http://www.euroqol.org/.

Hamilton D F, Lane J V, Gaston P, Patton J T, Macdonald D, Simpson A H, Howie C R. What determines patient satisfaction with surgery? A prospective cohort study of 4709 patients following total joint replacement. BMJ Open 2013; 3 (4): e002525.

Hanna S A, Sewell M D, Sri-Ram K, Miles J, Aston W J, Pollock R C, Carrington R W, Briggs T W. The effect of femoral head size on functional outcome in primary total hip arthroplasty: a single-blinded randomised controlled trial. Hip Int 2012; 22 (6): 592-7.

Howie D W, Holubowycz O T, Middleton R. Large femoral heads decrease the incidence of dislocation after total hip arthroplasty: a randomized controlled trial. J Bone Joint Surg Am 2012; 94 (12): 1095-102.

Hummel M T, Malkani A L, Yakkanti M R, Baker D L. Decreased dislocation after revision total hip arthroplasty using larger femoral head size and posterior capsular repair. J Arthroplasty 2009; 24 (6 Suppl): 73-6.

Jack C M, Walter W L, Shimmin A J, Cashman K, de Steiger R N. Large diameter metal on metal articulations. Comparison of total hip arthroplasty and hip resurfacing arthroplasty. J Arthroplasty 2013; 28 (4): 650-3.

Jameson S S, Lees D, James P, Serrano-Pedraza I, Partington P F, Muller S D, Meek R M, Reed M R. Lower rates of dislocation with increased femoral head size after primary total hip replacement: a five-year analysis of NHS patients in England. J Bone Joint Surg Br 2011; 93 (7): 876-80.

Jameson S S, Baker P N, Mason J M, Rymaszewska M, Gregg P J, Deehan D J, Reed M R. Independent predictors of failure up to 7.5 years after 35 386 single-brand cementless total hip replacements. Bone Joint J 2013; 95-B(6): 747-57.

Judge A, Arden N K, Batra R N, Thomas G, Beard D, Javaid M K, Cooper C, Murray D. The association of patient characteristics and surgical variables on symptoms of pain and function over 5 years following primary hip-replacement surgery: a prospective cohort study. BMJ Open 2013; 3 (3): e002453.

Klingenstein G G, Yeager A M, Lipman J D, Westrich G H. Increased range of motion to impingement with large head total hip arthroplasty: point of diminishing returns. Hip Int 2012; 22 (3): 261-5.

Langton D J, Sidaginamale R, Lord J K, Nargol A V, Joyce T J. Taper junction failure in large-diameter metal-on-metal bearings. Bone Joint Res 2012; 1 (4): 56-63.

Livermore J, Ilstrup D, Morrey B. Effect of femoral head size on wear of the polyethylene acetabular component. J Bone Joint Surg Am 1990; 72 (4): 518-28.

Matsushita A, Nakashima Y, Jingushi S, Yamamoto T, Kuraoka A, Iwamoto Y. Effects of the femoral offset and the head size on the safe range of motion in total hip arthroplasty. J Arthroplasty 2009; 24 (4): 646-51.

Meyer H, Mueller T, Goldau G, Chamaon K, Ruetschi M, Lohmann C H. Corrosion at the cone/taper interface leads to failure of large-diameter metal-on-metal total hip arthroplasties. Clin Orthop Relat Res 2012; 470 (11): 3101-8.

Murray D W, Fitzpatrick R, Rogers K, Pandit H, Beard D J, Carr A J, Dawson J. The use of the Oxford hip and knee scores. J Bone Joint Surg Br 2007; 89 (8): 1010-4.

ANJRR. "Australian National Joint Replacement Registry." 2013. Retrieved 08/03/2014, from https://aoanjrr.dmac.adelaide.edu.au/documents/10180/127202/Annual%20Report%202013?version=1.2&t=1385685288617.

Ostendorf M, van Stel H F, Buskens E, Schrijvers A J, Marting L N, Verbout A J, Dhert W J. Patient-reported outcome in total hip replacement. A comparison of five instruments of health status. J Bone Joint Surg Br 2004; 86 (6): 801-8.

Sedrakyan A, Normand S L, Dabic S, Jacobs S, Graves S, Marinac-Dabic D. Comparative assessment of implantable hip devices with different bearing surfaces: systematic appraisal of evidence. BMJ 2011; 343: d7434.

Sexton S A, Walter W L, Jackson M P, De Steiger R, Stanford T. Ceramic-on-ceramic bearing surface and risk of revision due to dislocation after primary total hip replacement. J Bone Joint Surg Br 2009; 91 (11): 1448-53.

Stroh D A, Issa K, Johnson A J, Delanois R E, Mont M A. Reduced dislocation rates and excellent functional outcomes with large-diameter femoral heads. J Arthroplasty 2013; 28(8): 1415-20.

Zagra L, Anasetti F, Bianchi L, Licari V, Giacometti Ceroni R. No difference in gait recovery after THA with different head diameters: a prospective randomized study. Clin Orthop Relat Res 2013; 471(12): 3830-7.