The trainer, the trainee and the surgeons’ assistant

CLINICAL OUTCOMES FOLLOWING TOTAL HIP REPLACEMENT

Balancing service provision and surgical training is a challenging issue that affects all healthcare systems. A multicentre prospective study of 1501 total hip replacements was undertaken to investigate whether there is an association between surgical outcome and the grade of the operating surgeon, and whether there is any difference in outcome if surgeons’ assistants assist with the operation, rather than orthopaedic trainees. The primary outcome measure was the change in the Oxford hip score (OHS) at five years. Secondary outcomes included the rate of revision and dislocation, operating time, and length of hospital stay.

There was no significant difference in ΔOHS or complication rates between operations undertaken by trainers and trainees, or those at which surgeons’ assistants and trainees were the assistant. However, there was a significant difference in the duration of surgery, with a mean reduction of 28 minutes in those in which a surgeons’ assistant was the assistant.

This study provides evidence that total hip replacements can be performed safely and effectively by appropriately trained surgeons in training, and that there are potential benefits of using surgeons’ assistants in orthopaedic surgery.

The number of total hip replacements (THRs) being performed is rising, and the increasing pressures on hospitals to reduce waiting times for elective surgery has resulted in greater difficulties in balancing adequate learning opportunities for surgeons in training, and service provision. Thus, trainee surgeons may not be adequately equipped to perform THRs independently, and may achieve inferior clinical outcomes compared with their consultant colleagues.

There is a paucity of research specifically assessing outcome measures between trainees and consultants, especially in orthopaedics, and it is difficult to give an evidence-based opinion. The few studies that exist involve a single centre and relatively small numbers of patients. Previous comparisons have also been between fully supervised trainees and consultants, rather than independently operating trainees.

An inverse relationship between the surgical volume and mortality and complication rates in joint arthroplasty has been frequently reported. Marston et al prospectively compared 413 Stanmore and Charnley THRs and noted that trainees had a significantly higher revision rate than consultants. Unwin and Thomas showed that dislocation rates were higher when hip hemiarthroplasties were undertaken by unsupervised trainees using a posterior approach. Increased complication rates have also been reported in other surgical specialties, although some studies in cardiothoracic, general and orthopaedic surgery have shown that the results obtained by trainees are comparable with those obtained by consultants. Moran et al reported no difference in the Harris hip scores between THRs performed by supervised specialist registrars and consultants at up to 18 months.

In a related but somewhat separate area, in 1999 the Royal College of Surgeons of England published a discussion document examining the future role of non-medically qualified assistants in the surgical team, using the term ‘surgeons’ assistants’ to describe such individuals. The role of the surgeons’ assistant has subsequently expanded especially in specialties such as cardiothoracic surgery. This trend looks set to continue. Together with the implementation of the European Working Time Directive, National Enquiry into Patient Outcome and Death and Modernising Medical Careers, the use of surgeons’ assistants has the potential to improve service provision but also to affect the training requirements of...
junior orthopaedic surgeons. There is little information on the impact of surgeons’ assistants on clinical outcomes in orthopaedic surgery.

Therefore we undertook a multicentre prospective cohort study involving a large sample of patients undergoing THR to investigate whether there is an association between surgical outcome and the grade of the operating surgeon (trainees vs trainers), and whether there is any difference in outcome if surgeon’s assistants assist with the operation rather than trainees (surgeons’ assistants vs trainees).

### Patients and Methods

The study involved 1501 THRs undertaken in seven centres between January 1999 and January 2002, of which 1367 were unilateral and 67 were bilateral. In all cases a cemented Exeter femoral component (Stryker Howmedica Osteonics, Mahwah, New Jersey) was used with various acetabular implants, both cemented and cementless, the most common of which were the Exeter (Stryker Howmedica) (n = 637, 42%) and the Charnley (DePuy, Warsaw, Indiana) (n = 398, 26%).

### Outcome measures

The Oxford hip score (OHS) was recorded pre-operatively, at three months, and at one, two, three, four and five years post-operatively by postal questionnaire. The OHS has been validated as a reliable measure of symptoms both pre- and post-operatively. It is a patient-centred 12-item health questionnaire designed to assess function and pain, scored 0 to 48, with 48 being the best score and 0 the worst.

The primary outcome measure was the change between the pre-operative and five-year post-operative scores (ΔOHS). Secondary outcome measures included complications such as dislocation and revision surgery. The operating time and the length of hospital stay were also recorded and compared between the groups. Standard practice was evaluated in each centre and, although patients were not randomised to operator grade, there was no reason why selection bias should have occurred between groups. The suggestion may have been that consultants would operate on more disabled or challenging cases, but pre-operative status data did not bear this out (Table I).

### Table I. The pre-operative and change in Oxford hip score (ΔOHS) at three months, and one, two, three, four and five years in the trainer and trainee groups

<table>
<thead>
<tr>
<th></th>
<th>Trainer</th>
<th>Trainees</th>
<th>p-value (ANOVA)†</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total THRs*</td>
<td>Number</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>Pre-operative OHS</td>
<td>1501</td>
<td>973</td>
<td>16.6 (8.0)</td>
</tr>
<tr>
<td>Mean post-operative OHS</td>
<td>1105</td>
<td>720</td>
<td>40.5 (8.5)</td>
</tr>
<tr>
<td>Mean ΔOHS</td>
<td>1300</td>
<td>809</td>
<td>18.3 (9.5)</td>
</tr>
<tr>
<td>3 months</td>
<td>1244</td>
<td>795</td>
<td>23.3 (9.2)</td>
</tr>
<tr>
<td>1 year</td>
<td>1166</td>
<td>775</td>
<td>23.1 (9.7)</td>
</tr>
<tr>
<td>2 years</td>
<td>1143</td>
<td>751</td>
<td>23.4 (9.5)</td>
</tr>
<tr>
<td>3 years</td>
<td>1061</td>
<td>699</td>
<td>23.9 (9.2)</td>
</tr>
<tr>
<td>4 years</td>
<td>1105</td>
<td>720</td>
<td>23.9 (10.1)</td>
</tr>
<tr>
<td>5 years</td>
<td>1300</td>
<td>809</td>
<td>18.3 (9.5)</td>
</tr>
</tbody>
</table>

* THRs, total hip replacements
† ANOVA, analysis of variance

### Table II. Age and surgical approach baseline characteristics by surgeon operator grade

<table>
<thead>
<tr>
<th></th>
<th>Trainers</th>
<th>Trainees</th>
<th>p-value (chi-squared test)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age at operation in years (n = 1497)*</td>
<td>68.0 (21 to 93)</td>
<td>68.8 (26 to 94)</td>
<td>0.28</td>
</tr>
<tr>
<td>Surgical approach (n = 1082)*</td>
<td>703</td>
<td>379</td>
<td>N/A‡</td>
</tr>
<tr>
<td>Anterolateral (n = 693)†</td>
<td>402</td>
<td>291</td>
<td>N/A‡</td>
</tr>
<tr>
<td>Posterior (n = 389)†</td>
<td>301</td>
<td>88</td>
<td>N/A‡</td>
</tr>
</tbody>
</table>

* data available on this number of total hip replacements
† subsample for comparisons
‡ N/A, not applicable

The primary outcome measure was the change between the pre-operative and five-year post-operative scores (ΔOHS). Secondary outcome measures included complications such as dislocation and revision surgery. The operating time and the length of hospital stay were also recorded and compared between the groups. Standard practice was evaluated in each centre and, although patients were not randomised to operator grade, there was no reason why selection bias should have occurred between groups. The suggestion may have been that consultants would operate on more disabled or challenging cases, but pre-operative status data did not bear this out (Table I).

Trainers vs trainees. Pre-operative OHS scores were available for all THRs. Two groups were assessed: trainers or associate specialist consultants (n = 973) and trainees including all levels of registrar grade (n = 528). For the primary outcome, OHS data at five years’ follow-up was available for 1105 THRs (74%), 720 performed by trainers and 385 by trainees. Data were available for both the posterior (n = 389) and the anterolateral (n = 693) approach. Trainers operated on patients with a mean age of 68.0 years.
(21 to 93), compared with 68.8 years (26 to 94) for the trainees \( (p = 0.28) \) (Table II).

**Senior vs junior registrars.** Out of 528 THRs initially undertaken by trainees, 313 were performed by senior registrars and 215 by junior registrars. The absolute pre-operative and \( \Delta \)OHS at three months and one, two, three, four and five years’ follow-up for senior registrars were compared with those of junior registrars. Although classification on the clinical report form consisted only of ‘senior’ and ‘junior’ registrar for these data, it is reasonable to assume that a senior registrar is more likely to have passed the fellowship examination (FRCS Trauma and Orthopaedics) than a junior registrar.

**Surgeons’ assistants vs trainees.** In orthopaedics, a surgeon’s assistant as defined by the Royal College of Surgeons of England is ‘any non-medically qualified assistant in surgery who performs pre-operative orthopaedic assessment in some teams; post-operative evaluation of range of movement is also a function; as is assisting at operation’. A surgical trainee is by definition medically qualified at either registrar or senior house officer grade.

Out of all THRs performed by consultants and associate specialists, 967 were grouped according to the assistant; group A, surgeons’ assistants \( (n = 262) \) and group B trainees, \( (n = 665) \). A subgroup analysis was also performed evaluating the outcomes when consultants assisted other consultants (group C, \( n = 40 \)).

The mean age of patients in group A was 72.1 years (32 to 91), in group B 66.6 years (21 to 93), and in group C 64.7 years (26 to 91).

In group A there were 79 men and 183 women; in group B, 267 men and 398 women; and in group C, 15 men and 25 women. The mean body mass index of the patients was similar in all three groups.

**Statistical analysis.** For the outcome measures, an analysis of variance (ANOVA) and appropriate Tukey’s post hoc tests were used to compare differences in data between the groups. Categorical and frequency data were analysed with chi-squared and Fisher’s exact tests, and the \( \alpha \) level of significance was defined as < 5\% \( (p \leq 0.05) \). SPSS 12.0.1 for Windows (SPSS Inc., Chicago, Illinois) software was used for statistical analysis.

**Results**

A total of 377 patients were removed during the study; 46 (3\%) were lost to follow-up without their implant status or outcome being known, 25 (2\%) had removal of all components or a revision, 149 (10\%) died, 92 (6\%) refused further participation, and 65 (4\%) were removed for other reasons. These included dementia, the development of cancer, the patient moving away and the patient being too ill to attend. In addition, there were some incomplete data and this was dependent on the type of outcome being assessed. For example, the OHS data for 19 THRs, although not lost to follow-up, were missing or corrupt. The results identify the individual sample size for each separate analysis, usually in respective tables.

**Trainees vs trainees.** Trainers operated on patients with a marginally better pre-operative functional status \( \text{mean OHS} = 16.6 \text{ (SD 8.0)} \) than did trainees \( \text{mean} = 15.2 \)
(SD 7.7), p = 0.001) and obtained a relatively superior post-operative score (mean five-year OHS, 40.5 (SD 8.5) vs 39.2 (SD 9.2) for trainees, p = 0.02). The ΔOHS between trainer and trainee groups was not significantly different at three months, one, two, three and five years after the surgery (Table I).

The complication rates were similar in the two groups (Table III). There were 15 dislocations (relative frequency, 1.5%) in the trainer group and nine (relative frequency, 1.7%) in the trainee group (p = 0.48). The overall incidence of dislocation was 1.6% (24 of 1501 THRs). The relative frequency of revision in the trainer group was 1.4% compared with 1.3% in the trainee group (p = 0.59), with an overall incidence in the study of 1.3% (20 of 1501 cases).

Analysis of the duration of surgery showed a significant difference in the mean operating time for trainees (104 minutes, 40 to 240) compared with trainers (85 minutes, 28 to 254) (p < 0.001). The mean length of hospital stay in the trainer group was 9.8 days (SD 10.2) and in the trainee group, 11 days (SD 27.7) (p = 0.22).

Senior vs junior registrars. The mean ΔOHS at five years follow-up in the senior registrar group was 25.2 (SD 9.4), which was significantly higher (p = 0.001) than that in the junior registrar group (ΔOHS = 21.8 (SD 10.5)) (Table IV). However, there was no difference in the incidence of dislocation (p = 0.46) or the revision rate (p = 0.30) between the groups. The mean duration of surgery for the senior registrar group was 106 minutes (48 to 240) compared with 102 (50 to 225) minutes for the junior registrar group (p = 0.09).

Surgeons’ assistants vs trainees. There was no significant difference in the mean ΔOHS post-operatively between the groups at any time (Table V). There were no significant differences in the incidence of dislocation or the rate of revision between the groups (Table VI). Patients in group A had a significantly shorter mean operating time (65 minutes, 35 to 187) than in group B (93 minutes, 28 to 254) and group C (91 minutes, 40 to 240) (p < 0.001), but the mean hospital stay was similar in all groups (group A, 8.9 days (SD 3.8); group B, 10.1 days (SD 11.9), group C, 10.8 days (SD 12.5) (p = 0.23).

Discussion
To our knowledge, this is the largest study of its kind in a Western healthcare system with the follow-up period of five years, comparing outcomes between trainers and trainees, senior and junior registrars, and also addressing the role of surgeons’ assistants in the delivery of orthopaedic care. The findings are potentially contentious and should therefore be interpreted with caution.

First, this study provides reassurance that trainees are capable of achieving comparable clinical results to consultants when performing THR. Although patients in the trainer group had a better OHS post-operatively, they also had better scores pre-operatively. The ΔOHS (clinical benefit) following surgery was similar in the trainee and trainer groups in the longer term at five years. This reassurance is most important for both patient populations and healthcare providers. It demonstrates that the system used for providing THR, in its present form using a variety of personnel, appears to be working. A patient undergoing THR by a suitable trainee surgeon will have an equally satisfactory outcome to that performed by a
consultant. Although there is potential for concern among the senior experienced colleagues who would like to think that they perform better operations the study should not be viewed negatively, but rather as an indication of the quality of training and education of trainee surgeons. It may be that another factor contributing to the competence of trainees is their ability to learn certain operations well when trained in high-volume units. Each of the centres in the study could be classified as a high-volume unit.

The use of the ΔOHS as the primary outcome measure for such a study is open to question. A more comprehensive picture would perhaps be derived from survival data, particularly at ten-year follow-up. The lack of radiological data also compromises the conclusion and limits long-term prediction, especially as some patients may have an asymptomatic radiological failure. However, as a defence, we believe that any impending failure or difference between the groups would have become apparent in the OHS data after five years. The Oxford self-reported scoring system has been shown to be a good marker of failure, albeit in the knee joint.

In terms of dislocations and revision rates, there was again no significant difference between the trainer and trainee groups. In 1996, Hedlundh et al published data on 4230 THRs showing that inexperienced surgeons had twice the number of dislocations as more experienced colleagues. Our findings contradict this, and indeed support the findings of Moran et al, who found no difference in dislocation rates between trainees and consultants. There have been advances in surgical technique over the last decade, including better awareness of the factors that contribute to dislocation, and this may help explain the improvement in dislocation rates in the trainee group.

The operating time was significantly increased in the trainee group compared with the consultant group, and this is to be expected as part of the learning curve. A study based on 31,745 primary THRs reported to the Norwegian Arthroplasty Register showed that the mean operating time for cemented hip replacement was 96 minutes. Operating times longer than 90 minutes were associated with an increased risk of revision for aseptic loosening. Operating times longer than 150 minutes were associated with an increased risk of infection, leading to higher revision rates. In contrast, our study suggests that although trainees took longer to perform cemented THR (mean time 104 minutes) compared with trainers (mean time 85 minutes), this did not affect the revision rates.

There is no information available on the exact state of training for each surgeon including consultants, but the trainees could be further subdivided into junior or senior registrars depending on their FRCS(Orth) status. The senior registrars achieved a significantly better outcome at five years than did the junior registrars, suggesting that more junior registrars should perhaps have increased supervision when performing THR.

Junior registrars aside, our study provides supportive evidence that patients’ faith in the abilities of senior trainees to undertake safe and effective THR is fully warranted. This study was undertaken in standard National Health Service practice, which includes a wide range of definitions for ‘supervision’ of the trainee by a trainer, from looking over the shoulder to occasional advice on the telephone. This is consistent with other papers confirming that surgical outcomes were similar between consultants and trainees under appropriate supervision. A survey by Williams, Hedge and Norton of 100 patients showed that a high proportion (74%) would be quite happy to allow trainees to perform part or all of their surgery.

With regard to the somewhat distinct issue of surgeons’ assistants, the traditional responsibilities of the nurse have evolved considerably to now include those that were previously the remit of medically qualified health professionals. This led to the concept and introduction of the physicians’ assistant in the United States in 1965. In the United Kingdom, the surgeons’ assistant fulfils a similar role to that of the physicians’ assistant in surgical disciplines, and, as in the United States, the number of surgeons’ assistants is increasing. Part of their duties may include performing incisions and closures, harvesting of grafts, and fulfilling the role of a first assistant. As a consequence, in 2001 the National Association of Assistants in Surgical Practice was formed following consultations with the Royal College of Surgeons of England and the Royal College of Nursing. In cardiothoracic surgery, surgeons’ assistants have become well established as part of the surgical team, and their responsibilities have overlapped with those of surgical trainees.

To our knowledge, there have been no studies examining the role of surgeons’ assistants in orthopaedic surgery. Furthermore, their effect on clinical outcomes following surgery has not been evaluated. The power of this study was sufficient (90%) to detect a two-point difference in ΔOHS and has shown that at five years there is no difference in ΔOHS, complication rates or length of hospital stay in the surgeons’ assistant group compared with the trainee group. This is somewhat to be expected, as the key stages of the operation would be carried out by the consultant or associate specialist.

However, the study has revealed a few findings of interest. It highlights a significant difference in duration of surgery, with a mean reduction of 28 minutes in the surgeons’ assistant group (mean = 65 minutes) compared with the trainee group (mean = 93 minutes). This could have a major impact on hospital resources and theatre planning, with a subsequent effect on the number of operations being performed and a reduction in waiting times for elective surgery. If four THRs are performed daily with trainees assisting, almost two hours of extra time could be made available if surgeons’ assistants assisted with surgery. Again, there are dangers of misrepresenting this finding. Although the message is indeed that surgeons’ assistants are useful, safe and their use can reduce theatre time, it must be remembered that the responsible consultant performing the
operation was once a trainee himself, and that assisting is a useful way to learn. Limitation or even withdrawal of training opportunities, albeit indirectly, could have a serious impact on the ability to train surgeons.

The role of surgeons’ assistant and trainee may occasionally come into conflict, but a survey by Shrivastava et al \(^{16}\) has shown that the presence of surgeons’ assistants does not appear to adversely affect the training of junior trainees in cardiac surgery. Trainees may benefit from the experience of surgeons’ assistants in learning basic surgical skills such as skin closure and acquiring familiarity with surgical instruments.\(^{26}\) The true cost is one of time, which is borne by the trainer who is supervising the trainee. This time will need to be preserved if surgical training is to continue.\(^{37}\) There is clearly a need for an integrated system in which trainees and surgeons’ assistants can work together for improved patient care and to facilitate mutual training.\(^{38,39}\)

There is an interesting subgroup of patients who had two consultants for their operation, one operating and one assisting. The reasons for this cannot be ascertained, but it may be that such cases were flagged as potentially difficult or high demand, although surprisingly, this is not borne out by the pre-operative scores. Alternatively, it may be due to logistics or availability issues. Interestingly, the combined expertise of two consultants did not appear to improve outcome further. Strong conclusions regarding dual consultant data cannot be made because of the small sample size of this group.

There are some more general limitations. Although the study was planned and data were collected prospectively, it was not randomised. In practice, randomisation would be difficult to achieve; some case selection is inevitable, as consultants would feel morally obliged and duty bound to offer help to a trainee in a difficult case. Similarly, other cases would prove easier than expected, allowing the trainee to proceed without the consultant having to intervene. There is some argument that, as the study was not randomised, bias could have existed in terms of choice of patient for each grade of surgeon; thus, trainees could potentially choose less difficult cases, but this is not supported by the data. The mean pre-operative hip scores suggest that, in reality, the trainees operated on patients with a slightly worse pre-operative status than the trainer group.

An important issue, if not a limitation, is the inequality and apparent inconsistency in the overall sample size for different comparisons. This is a manifestation of being as explicit as possible about the number of THRs reviewed. Rather than round sample sizes to a convenient number for all analyses, which would still remain a large and valid sample, we feel that the paper should reflect the fact that large multicentre cohort studies rarely produce tidy, fully complete data. In this paper we have reported the facts with regard to sample at the risk of looking inconsistent, rather than attempting any forced or artificial omission.

There are two separate points relating to this issue. First, there are a number of patients who were truly lost to follow-up. These have been reported. Secondly, there are also patients with incomplete data for specific variables. This incompleteness has no pattern and is most likely due to accidental errors of omission by the assessor at the time of evaluation. Rather than omit these patients, the preference has been to be inclusive, but with a clear notation for sample size for each analysis. For this reason, each analysis should be regarded separately rather than multiple analyses on one global set of 1501 THRs.

Other limitations include the lack of data on the level of experience of both the trainees and the surgeons’ assistants. The only differentiation was between junior and senior registrars. The level of supervision of the trainee during operation is also unknown. Although for all National Health Service operations the consultant maintains a responsibility for care by any trainee under his or her name, it is not known whether the consultant was on hand or even on site. Finally, as previously stated, there were some cases with incomplete data, but the power of the study remained adequate for most comparisons, including those revealing no difference between groups.

There are many avenues for future extension of this study and, in addition to other radiological findings, these could include whether the quality of cementation and orientation of components are related to the grade of a surgeon. The longer ten-year data will also be of interest.

The authors wish to thank Ms K. Honeybill (Clinical Research Associate, Stryker Howmedica Osteonics, Newbury) for her help in collating the EPOS data. The following people are principal investigators of the EPOS group: Mr P. Gibson, Mr J. Nolan, Mr A. Hamer, Mr M. Fordyce and Mr K. Tuson. The following are study coordinators for the EPOS group: A. Potter, A. McGovern, K. Reilly, C. Jenkins, A. Cooper, C. Darrah, L. Lawton, F. Isaphary and C. Pitchfork. Oxford is an NIHR Musculoskeletal Biomedical Research Unit. D. J. Beard is engaged as an occasional and independent consultant for Stryker EMEA. This study has been supported by Stryker Howmedica Osteonics (Newbury, UK).

Although none of the authors has received or will receive benefits for personal or professional use from a commercial party related directly or indirectly to the subject of this article, benefits have been or will be received but will be directed solely to a research fund, foundation, educational institution, or other nonprofit organisation with which one or more of the authors are associated.

References


