Fracture of the upper femur is a common injury in the elderly. The reported hospital mortality following its surgical correction has varied between 2.7% and 28% (McLaren, Stockwell and Reid, 1978; Goucke, 1985). It is more than 20 years since the effect of anaesthesia on mortality in these patients was first compared prospectively (Gauthier and Hamelberg, 1963). However, it was McLaren, Stockwell and Reid (1978) who first suggested that subarachnoid block (SAB) was associated with a reduced early mortality compared with general anaesthesia (GA) for internal fixation procedures. Subsequently, neither McKenzie and colleagues (1980) nor Davis and Laurenson (1981) could demonstrate a statistically significant difference, although in both studies mortality was less following SAB. In our opinion at that time, none of these studies was sufficiently large epidemiologically. Therefore, we planned a large multi-centre prospective randomized trial to re-examine the effect of anaesthesia on mortality in these elderly patients. This paper presents the results of that investigation.

SUMMARY
In a prospective randomized multi-centre study, the mortality following internal fixation surgery for fracture of the upper femur was investigated in 538 elderly patients allocated to receive subarachnoid blockade or general (narcotic-relaxant) anaesthesia. The 28-day mortality was 6.6% with subarachnoid, and 5.9% with general anaesthesia. The difference was not significant (95% confidence limits: -3.5 to +4.8). At 1 year following surgery, the mortality was 20.4%. Increasing age, ischaemic heart disease, cardiac failure, preoperative arrhythmias and poor ASA status were all associated with increases in early and long term mortality. A delay to surgery of more than 24 h from admission was also associated with an increased 28-day mortality. Senile dementia and admission other than from the patient's own home, were factors associated with a poorer long term outcome. From the point of view of mortality, subarachnoid anaesthesia did not appear to confer any advantages over general anaesthesia in non-prosthetic surgery for hip fracture in the elderly.

PATIENTS AND METHODS
The study was formulated on a multi-centre basis and co-ordinated from Christchurch. It was approved by the Ethics Committee of each of the five participating hospitals. For all or part of the period from June 1982 to September 1984, patients undergoing internal fixation procedures for trochanteric or subcapital fracture of the femur were considered for the trial unless one or more of the following were present: age less than 55 yr;
multiple trauma; possible replacement hemi- or total arthroplasty; pathological fracture.

Patients could also be excluded from the trial, but only before randomization, at the discretion of the anaesthetist if a specific contraindication to either SAB or GA were present. However, information on these, and any other patients excluded from the trial for whatever reason, including refusal, was documented in the postoperative period—to assess the rate of inclusion and any bias to patient selection.

Once they had been assessed for anaesthesia, patients suitable for inclusion in the trial were randomly allocated to receive either SAB or GA, the sexes being randomized separately within each hospital. In all other respects the patients were treated according to current practice in the respective hospitals. The design of the study was broad and allowed some latitude in the precise details of the two anaesthetic techniques, while requiring adherence to certain specific principles of management (table I). Patients in whom satisfactory SAB could not be established received GA. However, for statistical evaluation they were retained in the SAB group as advocated by Peto and co-workers (1976).

The anaesthetist completed pre- and intraoperative data sheets for each patient. Information included demographic details, past medical history, ASA status, (ASA, 1963) and concomitant disease, social status and details of the course of anaesthesia.

The postoperative follow-up (at 28 days) of each patient was the responsibility of one anaesthetist at each hospital. Postoperative complications, as recorded in the medical notes, and the duration of postoperative stay in the acute orthopaedic ward were recorded. All deaths during that time were documented. Postmortem confirmation of the cause of death was sought, whenever possible.

Long term follow up of the New Zealand patients (89% of the total) was based on the New Zealand Register of Births, Deaths and Marriages of the Justice Department as at December 31, 1984. Postoperative follow up therefore varied from 3 to 30 months, depending on the date of surgery.

**Statistical Methods**

The study was formulated using the recommendations in the report to the Medical Research Council’s Leukaemia Steering Committee (Peto et al., 1976) for the design and analysis of multicentre randomized clinical trials requiring prolonged observation of the patient.

The planned sample size ($n = 506$) was calculated on the basis of previous mortality figures as sufficient to provide an 80% chance of detecting a halving of the 28-day mortality from 16% in the GA group to 8% in the SAB group (Cohen, 1977). The trial was stopped just after this planned sample size (actual $n = 549$). Statistical analysis was carried out using the BMDP statistical package (Dixon, 1981).

The two anaesthetic groups were compared on the preoperative descriptor and prognostic vari-
### RESULTS

**Patient selection**

A total of 549 patients were entered into the trial. Eleven were excluded because one or more of the exclusion criteria were present, leaving 538 patients for analysis (table II). The relatively large input from Christchurch represents case collection in a busy general hospital over the whole period of the trial. The contributions from the other centres were drawn over shorter periods, or from hospitals with a lower throughput of upper femoral fractures.

The ages, weights and ASA gradings (table III) were similar in each hospital, with the exception of Nelson. The Nelson patients were older (mean 81.5 yr, SD 7) and sicker (30% of patients ASA grade IV) than those in the other centres. Twenty percent of the included patients had subcapital and 80% trochanteric fractures. The patients with subcapital fracture were equally distributed between the sexes and the two anaesthetic groups and there were no statistical differences in preoperative status or intraoperative course between them and the patients with trochanteric fractures.

### Anaesthetic management

In the 259 patients in the SAB group, attempts at lumbar puncture were unsuccessful in 30 patients (11.5%) and an incomplete block requiring additional analgesia of some form occurred in 14 patients (5.4%).

Hypotension, defined as a decrease in systolic

---

**Table II. The numbers of patients from each centre, female/male ratios and anesthetic techniques used**

<table>
<thead>
<tr>
<th>Hospital (period of study)</th>
<th>Patients</th>
<th>Female/male ratio</th>
<th>General anaesthesia</th>
<th>Subarachnoid blockade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auckland (7/82-2/84)</td>
<td>38</td>
<td>3.2:1</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>Christchurch (6/82-9/84)</td>
<td>313</td>
<td>2.9:1</td>
<td>164</td>
<td>149</td>
</tr>
<tr>
<td>Dunedin (6/82-9/84)</td>
<td>97</td>
<td>4.7:1</td>
<td>46</td>
<td>51</td>
</tr>
<tr>
<td>Nelson (6/82-9/84)</td>
<td>33</td>
<td>16:1</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>Perth (2/84-9/84)</td>
<td>57</td>
<td>3.4:1</td>
<td>33</td>
<td>24</td>
</tr>
<tr>
<td>Total</td>
<td>538</td>
<td>3.5:1</td>
<td>279</td>
<td>259</td>
</tr>
</tbody>
</table>

---

**Table III. Mean (SD) age, weight and ASA grading (expressed as % patient total) for each sex in 538 elderly patients with hip fracture**

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
<th>All patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr)</td>
<td>77.1 (9.8)</td>
<td>80.2 (8.4)</td>
<td>79.5 (8.8)</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>66.0 (10.1)</td>
<td>57.1 (11.4)</td>
<td>59.2 (11.7)</td>
</tr>
<tr>
<td>ASA Grade</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I or II</td>
<td>32%</td>
<td>46%</td>
<td>43%</td>
</tr>
<tr>
<td>III</td>
<td>54%</td>
<td>46%</td>
<td>48%</td>
</tr>
<tr>
<td>IV</td>
<td>14%</td>
<td>8%</td>
<td>9%</td>
</tr>
</tbody>
</table>
arterial pressure of more than 20% of the pre-induction value persisting for more than 10 min, occurred in 38% of the SAB group and 24% of the GA group. Of the SAB patients who exhibited hypotension, 82% received a sympathomimetic drug whereas only 2.5% of the hypotensive patients in the GA group did so.

There was one non-fatal anaphylactoid reaction at induction of general anaesthesia. No intraoperative death occurred.

**Anaesthetic technique and 28-day mortality**

The overall 28-day mortality was 6.3% (table IV); 6.6% in the SAB group and 5.9% in the GA group. The difference in mortality was not statistically significant (95% confidence limits for the difference between the groups: −3.5% to +4.8%).

The individual hospital figures, except for those from Christchurch, appear to show differences in mortality; however, these did not reach statistical significance at the 5% level, and did not consistently favour either anaesthetic technique. The overall mortality from Nelson, both 28-day and long term, was significantly higher than that from the other hospitals, but this was related to the poorer preoperative clinical state, particularly the presence of pneumonia, in this group and not to differences in management.

In the "failed spinal" patients the mortality rate did not differ from that of the series as a whole. There was no difference in the 28-day mortality between patients with trochanteric fractures and those with subcapital fractures.

**Anaesthetic technique and long-term mortality**

The mortality at 1 year was 20.4% and there was no difference in long-term mortality between the two anaesthetic groups. The relationship between the observed and the expected mortality rates is shown in table V. There was a significantly higher mortality during at least the first 6 months following surgery when compared with the general population matched for age and sex. This was most marked in the first month when the observed mortality was nine times that expected. There was no difference in long-term outcome between trochanteric and subcapital fractures.

**Factors related to mortality**

Several preoperative factors including the age of the patient and ASA grading were found to relate to increased mortality both within 28 days of surgery and long-term. These are listed in table VI. Several postoperative complications were also associated with a higher mortality (table VI). When the preoperative high risk factors indicated by an asterisk in table VI were included in the mortality analysis, there was still no difference in either 28-day or long-term mortality between the two anaesthetic groups.

**Factors not related to mortality**

Factors that appeared unrelated to either 28-day or long-term mortality included a history of hypertension or CNS disease other than dementia. Although the mortality was slightly higher in

| Table V. Observed and expected numbers of deaths in elderly patients following emergency hip surgery for time periods up to 2.5 years after operation (n = 479, New Zealand only). * P < 0.05; ** P < 0.01; *** P < 0.002 |
|-----------------|-----------------|-----------------|-----------------|
| Period          | Observed no. of deaths | Expected no. of deaths | Observed: expected ratio | Cumulative mortality |
| 1st month       | 31               | 3.5             | 9.0***           | 6.3%             |
| 2nd month       | 19               | 3.6             | 5.2***           | 10.4%            |
| 3rd month       | 10               | 3.4             | 2.9**            | 12.5%            |
| 4th-6th month   | 17               | 9.1             | 1.9*             | 16.0%            |
| 7th-12th month  | 21               | 15.8            | 1.3              | 20.4%            |
| Over 12 months  | 22               | 19.3            | 1.1              | 25.1%            |
| Total           | 120              | 54.6            | 2.2              |                 |
men, the difference between men and women was not statistically significant either in terms of crude survival or in terms of deaths relative to those expected from population life-tables taking into account age and sex. Neither the type of fracture nor the occurrence of intraoperative hypotension (as defined above) appeared to influence the outcome. The diagnosis of postoperative thromboembolic disease was based on clinical signs only and, therefore, its contribution to mortality could not be assessed. Pulmonary embolism was clinically diagnosed in 1.5% of patients.

**Cause of death**

There were 33 deaths within 28 days in the whole group, but only 10 patients underwent postmortem examination. In 26 cases the cause of death could be ascertained either from postmortem or the patient's medical records. These are shown in table VII.

**Duration of hospital stay**

In 524 patients the duration of postoperative stay in the acute hospital was recorded. The mean (SD) hospital stay was 15.7 (9.6) days for the SAB group, and 15.9 (9.0) days for the GA group. There was no significant difference between the two groups.

**Patients missed from the trial**

There were 57 patients from the New Zealand centres (11.9% of the total included) who met all the inclusion criteria but, for a variety of reasons, were excluded from the trial. Four patients (7%) died within 28 days. Of this group of 57 patients, 40 received GA of whom two died, and 17 received SAB of whom two died. These mortality rates do not differ statistically from each other or from the 538 patients included in the trial, confirming that no undue bias in patient selection had been introduced to the study.

**DISCUSSION**

Since McLaren, Stockwell and Reid (1978) reported a significant reduction in the 28-day mortality following SAB (7%) compared with GA (28%), the influence of anaesthesia on mortality following hip fracture surgery has been much debated. McKenzie and colleagues (1980) in a study of 100 patients found a 28-day mortality of 15.7% following GA compared with 10.2%.

**Table VI. The pre-and postoperative factors associated with increased 28-day and long-term mortality in elderly patients with hip fracture.* Inclusion in mortality analysis did not produce a difference in either mortality between groups**

<table>
<thead>
<tr>
<th>Mortality</th>
<th>28-Day</th>
<th>Long-term</th>
</tr>
</thead>
</table>
| **Preoperative factors** | 1. Ischaemic heart disease*  
2. Cardiac failure*  
3. Cardiac arrhythmias on ECG*  
4. ASA grade III or IV*  
5. Age; especially > 85 yr | 1. ASA grade III or IV*  
2. Cardiac arrhythmias on ECG*  
3. Ischaemic heart disease*  
4. Dementia  
5. Cardiac failure  
6. Admission other than from patient's own home  
7. Surgical delay over 24 h from admission  
8. Age |
| **Postoperative factors** | 1. Cerebrovascular accident  
2. Pneumonia  
3. Renal failure | 1. Cerebrovascular accident  
2. Pneumonia  
3. Renal failure |

**Table VII. Recorded cause of death in 26 of the 33 patients who died within 28 days after operation. No data were obtainable for the other seven patients**

<table>
<thead>
<tr>
<th>Cause of death</th>
<th>General anaesthesia</th>
<th>Subarachnoid blockade</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pneumonia</td>
<td>4</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Myocardial infarction</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Ischaemic heart disease</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>and cardiac failure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cerebrovascular accident</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Renal failure</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Pulmonary embolus</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Peptic ulcer</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Carcinoma</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>10</td>
<td>16</td>
<td>26</td>
</tr>
</tbody>
</table>
following SAB. Davis and Laurenson (1981), reporting on 132 patients, observed a 28-day mortality following GA of 13.2%, compared with 4.6% following SAB. In neither study did the mortality difference reach statistical significance at the 5% level. McKenzie, Wishart and Smith (1984) studied a further 150 patients and observed that SAB significantly decreased the 14-day mortality, compared with GA following femoral neck surgery, but that by 2 months after the operation no difference in mortality was evident between the groups.

More recently, Valentin and associates (1986) have presented data from 578 patients randomly allocated to SAB or GA in whom the overall mortality at 30 days was 7% and in whom no difference in mortality was present between the two anaesthetics. In the present study of 538 patients, the 28-day mortality was virtually the same at 6.3% and we have also been unable to demonstrate any difference in acute mortality related to anaesthetic technique.

The appropriate sample size for a study depends on the expected mortality rates (Cohen, 1977). The present study was planned using 28-day mortality rates suggested by the literature and our previous work. However, the actual mortality in the GA group was less than half that expected (6% v. 16%). The planned sample size gave power of 80% to detect (two-tailed) a reduction in mortality from 16% (GA) to 8% (SAB), but power of only 50% to detect a reduction from 8% to 4%. Over 1000 patients would be necessary to achieve power of 80% for that difference. Thus, as overall mortality decreases it is increasingly difficult to detect differences resulting from anaesthetic technique, even though there may be an effect in conditions of higher mortality.

The reported mortality rate following femoral neck fracture has varied widely (Goucke, 1985). However, our overall 28-day mortality at 6.3% is very similar to that reported in recent major epidemiological studies (Dahl, 1980; Jensen, 1981; Kenzora et al., 1984; Valentin et al., 1986). Whilst marked variations may be seen when mortality at only one point in time is measured, these tend to disappear when long-term outcome is considered. For instance, the very low 28-day mortality (2.7%) reported by Goucke (1985) reaches 20.8% 1 year after the operation—a value which is the same as that for both our patients (20.4%) and those of Valentin and associates (1986) (19%).

As can be seen from figure 1, the cumulative mortality followed very similar trends in the two anaesthetic groups throughout the follow-up period. At no stage was there evidence of a differing mortality rate between the two as reported by McKenzie, Wishart and Smith (1984) in their patients. Both ours and the Hellerup study (Valentin et al., 1986) appear to illustrate the widely recognized phenomenon in prospective studies of mortality of the observed mortality rates being lower than predicted.

The results from the individual hospitals taking part in this trial illustrate how small series can
produce results suggestive of a major effect when this is not confirmed in the series as a whole. In multicentre trials there is no assumption that the type of patient or quality of management is comparable in all centres. It is only assumed that if a treatment is of medical value, then this will be true at every centre and, although the treatment difference actually observed in an individual centre may be randomly obscured or even reversed, the overall sum of all the separate treatment differences, one from each centre, should point in the right direction and the separate differences should thus reinforce each other, even if most, or perhaps all, of the separate differences are not individually significant (Peto et al., 1976).

Despite the very similar results there are certain methodological differences between our own and the Hellerup studies. Whilst the Cox proportional hazards model was used in both, we attempted to take into account well-defined risk factors in the analysis. We specifically excluded patients undergoing prosthetic surgery, although many of our patients with subcapital fractures would probably have undergone Moore's hemiarthroplasty at Hellerup. Overall, the balance of patients is not quite the same in the two studies.

A much more important difference relates to the handling of “failed” SAB. In the present study 11.5% of the patients were originally allocated to receive SAB, but required GA because of the failure to establish adequate blockade. These were retained in the SAB group for statistical analysis. In the Hellerup group the failure rate was also approximately 10%, but they chose to exclude these patients from their analysis after randomization. We consider it correct statistically to retain them in the original random allocation group for mortality analysis on an “intention to treat” basis as advocated by Peto and co-workers (1976). The outcome in these patients must be included in the SAB group if we are to have a realistic appraisal of the effect on mortality were subarachnoid anaesthesia to be routinely used for this form of surgery. In addition, because their follow-up was not reported, the Hellerup group make assumptions about their excluded patients not unduly biasing their study, which may or may not be true. Since we followed the 57 patients who met all the inclusion criteria for the study but were excluded for various reasons before randomization, we can be fairly certain that no undue bias resulted in the present study.

From a practical viewpoint, it would appear from these two large studies that there is a quite high failure rate with spinal anaesthesia in elderly patients with hip fracture when the clinical practice of a large number of anaesthetists with a wide range of experience (including many junior trainees) is reviewed in this way. This contrasts with studies in which spinal anaesthesia has been performed by only one anaesthetist, or a few experienced anaesthetists, when a failure rate of less than 5% has usually been the case. Such practical differences must always be borne in mind when advocating general policies for patient management based on tightly defined and carefully controlled experimental work.

We believe that the present study and that of the Hellerup group (Valentin et al., 1986) realistically represent the clinical situation if a general policy of SAB for femoral neck fracture surgery were implemented in a busy hospital or hospital group. In these circumstances anaesthetists differ in details of anaesthetic management and the patients may be under the care of one of several surgical teams. Having said this, we found from examination of the anaesthetic records that there was a reassuring uniformity of management and that the design of the study was faithfully adhered to by our colleagues.

Many factors have been reported to be associated with increased mortality after surgery for fracture of the femoral neck. These include age and male sex (Gallanaugh, Martin and Millard, 1976), preoperative physical status and activity (Barnes et al., 1976), preoperative nutritional status (Bastow, Rawlings and Allison, 1983), delay between injury and surgery for whatever reason (Aldrete, Davis and Hingson, 1967), fracture type and operative procedure undertaken (Chan and Hoskinson, 1975; Barnes et al., 1976), preoperative cerebral dysfunction (Miller, 1978), and postoperative morbidity (Kenzora et al., 1984).

Our data are consistent with these reports in that the most important factors demonstrated in the present study were age, cardiovascular disease, ASA grading and cerebral dysfunction. We differ from the Hellerup group only in not seeing a difference in mortality related to sex or to the type of fracture in our patients.

Zindrick and co-workers (1985) found, in a careful 10-year retrospective review, that there appeared to be a tendency for older, less healthy patients to have been treated with primary hemiarthroplasty. They suggested that the previously reported increased postoperative morbi-
dity and mortality associated with this procedure (as compared with internal fixation) may be a result of biased patient selection. A similar bias could explain why, in the Hellerup study, a difference in mortality related to the fracture was seen, which was absent from our study, since the Hellerup study contained hemiarthroplasty procedures whilst ours did not.

Postoperative thromboembolism has been widely regarded as the leading cause of death in these patients (Sevitt and Gallagher, 1959). In the present study, which was based only on clinical signs and limited postmortem data, no evidence of a high contribution of thromboembolism to mortality was seen. Nevertheless, the incidence of deep vein thrombosis does appear to be significantly reduced by subarachnoid anaesthesia (Davis, Quince and Laurensen, 1980; McKenzie et al., 1985). It has been suggested that this might explain the increased mortality with general anaesthesia observed by McKenzie, Wishart and Smith (1984) in the first 2 weeks following surgery. Re-analysing their data (see Appendix) we observed that the mortality in their SAB group was significantly higher than in the GA group in the 3rd and 4th weeks after operation. If these mortality patterns are related in any way to thromboembolic disease, as suggested, then the more likely postulate is that spinal anaesthesia, like low-dose heparin in total hip replacement arthroplasty (Sikorski, 1984), merely delays the onset of deep vein thrombosis rather than preventing it. This has yet to be tested.

Our results, like those of Valentin and colleagues (1986), support the conclusion that when general anaesthesia has a low mortality, subarachnoid anaesthesia will probably not offer additional advantages. A difference in either acute (28-day) or long-term survival following spinal or general anaesthesia in elderly patients undergoing internal fixation surgery for fractures of the neck of femur has not been verified in this study.

APPENDIX

McKenzie, Wishart and Smith (1984) analysed differences in mortality rates at different time intervals by repeated applications of Chi-squared tests, thus increasing the likelihood of a Type I statistical error. We re-analysed the data from their figure 1 in intervals of 0-14 days and 15-49 days, these intervals being chosen to provide expected deaths > 5. Pearson's conditional Chi-squared test for truncated data (Elandt-Johnson and Johnson, 1980) was significant over both periods ($\chi^2 = 5.7$ and 5.9 respectively, $P < 0.02$). Thus there was a significantly higher general anaesthetic mortality in the first 2 weeks and higher spinal anaesthesia mortality in the next 4 weeks. Varying the upper end of that interval had little effect. Naturally, the Mantel–Haenszel test over these periods was not significant ($\chi^2 = 0.01$, $P > 0.9$) because it tests for consistent differences (Elandt-Johnson and Johnson, 1980) which in this case were not present.

ACKNOWLEDGEMENTS

This study was funded in part by the Canterbury Medical Research Foundation. Miss Wilkinson was in receipt of a grant-in-aid from Reckitt & Colman (NZ) Ltd. The help and advice of Professor W. J. Gillespie, Dr J. E. Wells, Dr R. H. Acland, E. de Boer and B. Solomon was greatly appreciated.

REFERENCES


