Long-term Cognitive and Social Sequelae of General Versus Regional Anesthesia during Arthroplasty in the Elderly

Warren R. Nielson, Ph.D.,* Adrian W. Gelb, M.B.Ch.B.,† Joseph E. Casey, M.A.,‡ Foster J. Penny, M.D.,§ Richard N. Merchant, M.D.,§ Pirjo H. Manninen, M.D.†

This study compared the effects of general and regional anesthesia on cognitive and psychosocial functioning in elderly persons. Sixty-four patients between 60 and 86 yr of age undergoing knee arthroplasty were randomly assigned to receive either general or regional anesthesia. A battery of psychometric tests, including the Satz-Mogel form of the Wechsler Adult Intelligence Scale—Revised, the Wechsler Memory Scale—Revised, and the Sickness Impact Profile, and various neuropsychological measures were administered by a blinded observer just before surgery and again 3 months later. Analyses of covariance revealed improvements in most measures that were equivalent between groups. The results indicated that there were no cognitive or psychosocial effects of general or regional anesthesia after 3 months in elderly persons undergoing knee arthroplasty. In this patient population, general anesthesia poses no more risk to long-term mental function than regional anesthesia. (Key words: Age factors: elderly. Anesthetic techniques: general; regional. Brain: cognition; memory; psychomotor function. Postoperative period.)

There has been a dramatic change in demography in the latter half of this century: approximately 11% of the population is now over 65 yr of age. This percentage is expected to increase to 12.5% by the turn of the century. It has been estimated that after the age of 65, over half of the population will require surgical intervention at least once during the remainder of their lives. Despite these facts, the contribution of anesthetic technique to postoperative morbidity remains unresolved.

The relative advantages and disadvantages of general anesthesia and regional anesthesia for major surgery in geriatric patients remains controversial. One important concern is that general anesthesia may result in more postoperative mental dysfunction. The current study addresses the issue of whether regional anesthesia offers advantages over general anesthesia for elderly patients in terms of both neuropsychological functioning and the ability to perform activities of daily living. Because the objective of the current study was to evaluate possible long-term effects of anesthesia, patient assessments were made preoperatively and 3 months postoperatively.

Materials and Methods

Participants

Ninety-eight individuals between 60 and 86 yr of age who were admitted to the hospital for elective knee arthroplasty participated in the study. This group excluded persons with a medical contraindication to general or regional anesthesia and those with a neurologic or psychiatric disorder. After admission to the hospital and once consent to participate was obtained, the patient was then randomized to receive spinal or general anesthetic. The attending anesthesiologist was informed as to which group the patient had been randomized but the actual choice of drugs was left to the anesthesiologist.

Psychologic Measurement

A research associate/psychometrist who was trained in administration of the tests and procedures that were used conducted the assessments unaware of the experimental condition to which participants were assigned. Each patient underwent a 3-h assessment within 24 h before surgery and again 3 months after surgery. The following psychometric instruments were used to assess impact of illness and various neuropsychological functions.

Wechsler Adult Intelligence Scale—Revised

The most commonly used, individually administered test of general intelligence, this restandardization of the Wechsler Adult Intelligence Scale—Revised (WAIS-R) is composed of 11 subtests, six verbal and five nonverbal. It measures a broad spectrum of mental abilities such as wealth of general knowledge, common sense, judgment, verbal concept formation, and visuospatial, visuoconstructional, and problem-solving skills. In addition to scores for each of the subtests, three age-adjusted intelligence quotients are calculated that provide an overall measure of verbal abilities (VIQ), performance or visuospatial/constructional abilities (PIQ), and a summary or full scale score (FSIQ). A Satz-Mogel short form of the WAIS-R was used in the study. This abbreviated version requires about one-half the time of the full WAIS-R to administer (i.e., 30–45 min), uses items from all subtests, and is highly correlated with the full WAIS-R (r = 0.98).
Wechsler Memory Scale–Revised

The Wechsler Memory Scale–Revised (WMS-R)\textsuperscript{6} is an instrument that measures memory function in adolescents and adults. It is composed of a series of subtests that measure various facets of attention and memory including mental control, figurative memory, logical memory, visual paired associates, verbal paired associates, visual reproduction, digit span, and visual memory span. These subtests are grouped into two composite scores: general memory and attention/concentration. General memory is subdivided into verbal memory and visual memory. A delayed recall composite is an index of how much verbal and figurative information is retained over a half-hour period. In addition, a global memory quotient is obtained. The full examination requires 45 min to 1 h to administer. The WMS-R and its predecessor, the WMS,\textsuperscript{6} are perhaps the most widely used and acknowledged tests of memory function.

Trail-making Test

A sensitive psychometric indicator of brain dysfunction, the Trail-making Test is given in two parts, A and B, and provides a measure of visual conceptual abilities and visuomotor tracking.\textsuperscript{7,8} Part A requires the subject to connect in sequence 25 numbered circles located randomly on a page (i.e., 1-2-3...25). Part B requires the subject to connect the same number of circles numbered 1-13 with the remaining lettered A–L by alternating between the two sequences (i.e., 1-A-2-B-3-C...25).

Controlled Oral Word-association Test

This test of verbal fluency requires the subject to report in three separate 1-min trials all the words he or she can think of beginning with the letters C, F, and L.\textsuperscript{9} The individual’s score is the average of the three trials plus an adjustment for age and education. Damage to any part of the brain tends to lower fluency scores, with lesions of the frontal lobes having the greatest effect on performance.\textsuperscript{10}

Finger Oscillation Test

This test measures motor speed and coordination of the dominant and nondominant hands.\textsuperscript{11,12} Six trials are administered on each hand. The individual is required to tap a lever with their index finger as fast as possible for a 10-s interval. The number of taps for each trial is recorded on a mechanical counter. If, after rank-ordering the scores, either the highest or lowest is more than five points removed from its adjacent score, it is not included in the calculation of the mean. Thus, the score for each hand is the mean of five or six trials. A depressed score on one hand implies contralateral brain impairment, particularly of the perisylvian region.

Two-point Discrimination

With eyes closed, the patient is asked to indicate whether he or she has been touched on the hand with either one or two points of anesthesiometer.\textsuperscript{13} A two-point discrimination threshold is the least distance between the points that an individual can detect. This test is sensitive to dysfunction of the cerebral hemisphere contralateral to the affected hand, particularly in the parietal lobe.

Hand Preference Questionnaire

The patient is asked to demonstrate seven activities, such as throwing a ball, hammering a nail, writing, etc. This test provides an objective measure of handedness.\textsuperscript{13} The examiner than makes note of all apraxic errors (i.e., use of a body part as the object).

Sickness Impact Profile

The impact of illness on the activity patterns of each patient was measured using the Sickness Impact Profile.\textsuperscript{14} The SIP is an instrument that assesses the impact of illness using a structured interview format. The 136 items evaluate the following 12 areas: 1) ambulation, 2) mobility, 3) body care and movement, 4) social interaction, 5) communication, 6) alertness behavior, 7) emotional behavior, 8) sleep and rest, 9) eating, 10) work, 11) home management, and 12) recreation and pastimes. The first three scales comprise the Physical Dimension Score; the second four scales make up the Psychosocial Dimension Score.

RESEARCH DESIGN AND DATA ANALYSIS

A split-plot factorial design with one between (regional versus general anesthesia) and one within-group factor (presurgery versus postsurgery) formed the basis for statistical analysis of the data collected during the course of the study. The data were analyzed using analysis of covariance with presurgical test scores as covariates. This statistical technique was chosen to increase the power of the statistical test.\textsuperscript{15} The data sets were as follows: 1) the WAIS-R scale scores, 2) the WMS-R index scores (verbal memory, visual memory, attention/concentration, and delayed recall), 3) the two SIP dimension scores, and 4) the other sensorimotor and cognitive test scores. All results are reported as mean ± SD.

Results

Of the initial sample of 98 patients, 64 completed a preoperative assessment and a 3-month postoperative assessment and agreed to randomization (39 general and 25 regional anesthetization). Twenty-two persons did not return for the postanesthesia assessment and 12 were excluded because they did not receive the anesthetic condition to which they had been randomized. In addition, four regional anesthesia patients required general anes-
thmia due to the inadequate effect of spinal anesthesia. In accordance with the "intact to treat" principle, these patients were included in the regional anesthesia group. However, even when the data were analyzed with these four patients included in the general anesthesia group, the results were unchanged. Using the IQ as a sensitive index of overall change and assuming a ten-point decrement to be clinically meaningful, our study had a power of 0.90.

The average age in the general anesthesia group (70.08 ± 6.20 yr) did not differ from that found in the regional anesthesia group (68.04 ± 6.00 yr; P > 0.10). These groups were also similar in terms of average number of years of education (general group, 10.38 ± 2.82 yr; regional group, 11.60 ± 3.98 yr; P > 0.10) and ASA physical status (general group, 2.3 ± 0.5; range, 1–3; regional group, 2.1 ± 0.7 [range 1–3]).

Intraoperative Data

Eighty percent of the general anesthesia group and 79% of the regional anesthesia group received no anxiolytic preanesthetic medication; however, the remaining participants received either diazepam or lorazepam orally. Induction of general anesthesia was with thiopental (278.9 ± 108.9 mg), and tracheal intubation was facilitated with succinylcholine (112.3 ± 20.7 mg). Twenty-one patients received lidocaine (79.5 ± 20.1 mg), and 14 received a defasciculating dose of d-tubocurarine (4.5 ± 3.7 mg). Anesthesia was maintained with N2O in oxygen, isoflurane, and fentanyl (0.19 ± 0.11 µg). The muscle relaxants used were either pancuronium bromide or atracurium, and their effects were reversed with atropine and neostigmine at the end of the procedure. In the spinal anesthesia group, 22 patients received tetracaine (15.6 ± 3.4 mg) and the remainder bupivacaine (16.4 ± 4.3 mg). Dural puncture was performed at the L3–4 or L4–5 level. Four of the patients receiving general anesthesia received intravenous ephedrine (8.8 ± 2.5 mg), and five of those receiving spinal anesthesia received ephedrine (7.7 ± 9.8 mg). There were no differences between the groups in intraoperative blood pressure (general group, 130.4 ± 13.2 mmHg; spinal group, 127.4 ± 15.2 mmHg). The duration of surgery was similar (range, 85–200 min.). No patient required a blood transfusion, and the amount of crystalloid administered was similar in the two groups (general group, 1.5 ± 0.8 l; spinal group, 1.7 ± 0.6 l). All spinal anesthetic patients received intraoperative sedation with either diazepam (n = 18; 10.3 ± 4.1 mg) or lorazepam (n = 7; 2.4 ± 0.4 mg). There were no differences between the groups in terms of the amount or choice of postoperative opioid use.

Returners versus Nonreturners

Although it was expected that in this population, given the length and difficulty of the assessment procedure, a relatively large number of individuals would decline follow-up testing, we were concerned that this might introduce a selection bias. To assess this possibility, the 22 nonreturners were compared with the returners on a variety of measures, including age, medications, amount and type of anesthesia, vasopressors, antiemetics, and intraoperative times. In addition, these groups were compared on the full battery of preoperative intellectual, memory, and functional measures described above. The only variable for which these groups differed was an age by spinal anesthesia interaction, with spinal anesthesia nonreturners being slightly older (72.2 ± 4.7 yr) than spinal anesthesia returners (68.0 ± 5.8 yr; P < 0.03). There were no differences either between the two general anesthesia groups (returners vs. nonreturners) or collapsing across anesthesia type. A higher percentage of regional (28.7%) compared to general (17.7%) anesthetic participants did not return for the second session. Analysis among these subgroups (i.e., return/nonreturn × anesthetic type) again failed to identify differences in the dependent variables described above.

Effect of Age

Although analyses of covariance with the presurgical test scores as covariates statistically controlled for the effects of age and presurgical (baseline) levels, it is possible that individuals at the more extreme end of the age continuum may have been differentially affected by general versus regional anesthesia. To assess this possibility, the sample was divided into two groups: those under 70 and those over 70 yr of age. The data for these groups were analyzed using analyses of variance (ANOVA). The age × anesthetic type × time (pre/post) interactions for these analyses were nonsignificant for all dependent variables, indicating that the older group did not perform more poorly following general as compared to regional anesthesia.

Intelligence

The general and regional anesthesia groups did not differ in the extent to which their IQs changed from preto post surgery (fig. 1). Thus, although there was a significant increase in Full Scale IQ scores from the first to the second assessment (P < 0.001), there were no between-group effects for type of anesthesia. Similarly, VIQ and PIQ showed increases over time (P < 0.005 and P < 0.001, respectively) but no between-group effects.

Memory

Analyses of covariance were also conducted on the four WMS-R index scores: verbal memory, visual memory, attention/concentration, and delayed recall. Consistent with the IQ results, there were no differences between the two
anesthesia groups for verbal memory (P > 0.10), visual memory (P > 0.05), attention/concentration (P > 0.10), or delayed recall (P > 0.10). Table 1 illustrates the relevant means and standard deviations; however, main effects for time (pre- vs. postsurgery) emerged such that both groups had higher scores following surgery (verbal memory, P < 0.001; visual memory, P < 0.01; delayed recall, P < 0.0001; attention/concentration, P < 0.05).

OTHER SENSORMOTOR AND COGNITIVE TESTS

The Hand Preference test indicated that the groups were equated for right versus left handedness (P > 0.10). A square root transformation was conducted on the number of apraxic errors before the analysis due to the frequency of zeros in participants’ scores. Tables 2 and 3, and figures 2 and 3 portray the means and SD for each group for both pre- and postsurgical assessments. None of the analyses of covariances conducted on these variables indicated between-group effects for anesthesia (P > 0.10). Unlike some of the WAIS-R and WMS-R variables, none of these measures were, collapsing across groups, significantly higher at the postsurgical assessment.

SICKNESS IMPACT PROFILE

Analyses of covariance on the physical and psychosocial dimensions of the SIP indicated large improvements in scores for both groups at the postsurgical assessment (P < 0.001). Figure 4 illustrates the relevant means and SD. There were no between-group effects for type of anesthesia (P > 0.10).

**Discussion**

Bedford was one of the first investigators to discuss the question of iatrogenic anesthesia-induced mental deterioration in the elderly. He reviewed the records of 4,250 patients over the age of 65 yr, 1,193 of whom had received a surgical intervention under general anesthesia. Of the latter group, 120 (10%) were believed to have been adversely affected. A subset of 18 (1%) were described as “human vegetables.” This report went unchallenged until 1961 when Simpson et al. were unable to find differences in mental ability between general and local anesthesia groups but described deterioration in “social integration.” Unfortunately, the psychometric assessment procedures used were unlikely to detect changes in intellectual functioning that were less than extreme (the Tooting-Bec Questionnaire). No statement was made as to whether either the subjects or those who administered the assessment procedure were blind to the experimental protocol.

More recently, Holle et al. compared epidural and general anesthesia in elderly patients (age range, 56–84 yr) undergoing total hip arthroplasty. Although this study was prospective in nature and involved random assignment of subjects to experimental conditions, no standardized methods of assessing cognitive function were included. Mental status was assessed in a general way through unstructured interviews and a mailed questionnaire. Those who conducted the interview were apparently not blind to either the experimental condition of the interviewees or the experimental hypotheses. No reliability or validity data were presented for the question-

---

**Table 1. WMS-R Index Scores**

<table>
<thead>
<tr>
<th>Group and Assessment Time</th>
<th>Verbal</th>
<th>Visual</th>
<th>Attention/Concentration</th>
<th>Delayed</th>
</tr>
</thead>
<tbody>
<tr>
<td>General (n = 39)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative</td>
<td>95.74 ± 13.86</td>
<td>90.23 ± 14.19</td>
<td>92.10 ± 16.05</td>
<td>91.26 ± 11.58</td>
</tr>
<tr>
<td>Postoperative</td>
<td>100.28 ± 14.02</td>
<td>91.51 ± 15.14</td>
<td>95.08 ± 16.13</td>
<td>96.77 ± 13.90</td>
</tr>
<tr>
<td>Regional (n = 25)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preoperative</td>
<td>100.88 ± 16.27</td>
<td>95.28 ± 16.78</td>
<td>98.08 ± 14.66</td>
<td>97.92 ± 13.90</td>
</tr>
<tr>
<td>Postoperative</td>
<td>105.80 ± 16.52</td>
<td>101.80 ± 17.14</td>
<td>101.80 ± 14.64</td>
<td>104.64 ± 18.76</td>
</tr>
</tbody>
</table>
TABLE 3. Trail-making Test

<table>
<thead>
<tr>
<th></th>
<th>Preoperative</th>
<th>Postoperative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>General</td>
<td>Regional</td>
</tr>
<tr>
<td>Trails A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time (s)</td>
<td>46.97 ± 22.96</td>
<td>38.04 ± 12.83</td>
</tr>
<tr>
<td>Errors</td>
<td>.24 ± .49</td>
<td>.32 ± .63</td>
</tr>
<tr>
<td>Trails B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time (s)</td>
<td>148.30 ± 77.65</td>
<td>125.00 ± 80.48</td>
</tr>
<tr>
<td>Errors</td>
<td>1.08 ± 1.26</td>
<td>1.12 ± 1.42</td>
</tr>
</tbody>
</table>

All values are mean ± SD.

naire. These authors reported that while none of the patients in the epidural group had significant postoperative mental changes, seven of 31 (22%) general anesthesia patients did evidence such changes. Five of the latter group were judged to have changes that reduced the quality of their lives at follow-up (range, 4–10 months).

In contrast to the data of Hole et al., Riis and co-workers reported no differential cognitive effects of general anesthesia 3 months following hip arthroplasty in a small group of patients.20,21 This study used a limited battery of psychological tests, and the test administrator was blind to treatment condition. In addition, Riis et al. suggested that the results of Hole et al. were most likely the result of "special circumstances [which] were operating in their investigation."

Bigler et al. compared the effects of general and spinal anesthesia on mental function in elderly patients undergoing surgery for hip fractures.22 Although they concluded that there were no differential effects at 3-month follow-up, their spinal anesthesia group showed improved performance at follow-up that was not evident for the general anesthesia group. A weakness of this study was the use of a very brief (ten-item) mental status examination that may have been insensitive to all but gross cognitive impairment.23

Berggren et al. assessed the extent of postoperative confusion in elderly patients having surgery using epidural versus general anesthesia for femoral neck fractures.24 These authors were unable to identify differences between these groups in amount of disorientation or confusion. Again, the indices of cognitive change were fairly gross and unlikely to detect more subtle effects.

Finally, Ghoneim et al. used a comprehensive test battery and assessed patients preoperatively, within 1 week postoperatively, and again 3 months postoperatively.25 As with other recent studies, no effects of general anesthesia were observed on intellectual capabilities, memory, or activities of daily living.

The results of the current study suggest that general anesthesia does not induce iatrogenic cognitive sequelae among elderly patients undergoing knee arthroplasty. None of the intellectual, memory, or other neuropsychological functions measured during the course of the study indicated either a global deterioration across groups or a greater degree of cognitive impairment among those receiving general anesthesia. Knee arthroplasty was the type...
of surgical intervention selected for the current study to avoid the possible confounding effects of heavy intraoperative sedation and hemodynamic instability associated with major blood loss, including the possible effects of multiple blood transfusions. Because the surgical intervention involved was associated with minimal physiologic disturbance, e.g., hemorrhagic hypotension or blood transfusion, this study afforded a reasonably "pure" assessment of the effects of anesthetic management per se.

To ensure that the results of this research were generalizable to the real-life situation of a busy operating room (i.e., were externally valid), choice of drugs was left up to the attending anesthesiologist. All the patients receiving spinal anesthesia received intraoperative benzodiazepine sedation. This may have served to negate any differences between the groups. However, because the effects of these drugs is relatively short lived, they should not have influenced long-term cognitive functioning measured at the postoperative testing.

For many of the variables assessed during this study, there was an overall improvement at the postsurgical assessment. Although this finding indicates that, collapsing across both anesthetic conditions, there was no deterioration in function, this improvement may well have been a result of practice effects. In the context of the current study, although statistically significant, such changes are unlikely to be clinically meaningful. However, it is worth noting that the two groups were similar in terms of the magnitude of these increases.

Although the relatively high dropout rate in this study was a cause for concern, analysis of the various demographic and outcome variables suggested the following: 1) the nonreturners differed from the returners only in terms of age; 2) there was no interaction between anesthetic choice and return/nonreturners; and 3) there was no interaction between return/no return and performance on any of the variables measured at the initial assessment. Thus, the differential dropout rate does not appear to have biased the results of this study. As elderly persons are frequently concerned about their diminishing mental abilities, feel unwell, or do not wish to go through extensive and personally irrelevant mental gymnastics, they may elect not to participate. For this and other reasons (e.g., lack of ready transportation), a high dropout rate may be common when conducting a thorough assessment with an elderly population.

An important aspect of the current study was the inclusion of a psychometrically sound measure of the impact of the two anesthetic conditions on activities of daily living. Although Simpson et al. reported deterioration in social integration, this variable has not been adequately studied. Unlike Simpson, our results suggest that anesthetic regimen did not have a significant impact on functional status as measured by the SIP.

In summary, the results of the current study are consistent with those of the majority of recent investigations that have evaluated the effect of anesthesia on cognitive functioning in elderly persons. General anesthesia does not appear to present an increase in risk in terms of either neuropsychological functioning or impact on daily life experiences. It remains possible, however, that subsets of patients exist for whom this general rule does not hold.

The authors thank Drs. C. Rorabeck and R. Bourne for their helpful cooperation, and Beverly Robb and Carol Pekaki for secretarial assistance. A special debt of gratitude is owed to Shirley Munk for her unfailing labor and extraordinary organizational skills.

References

11. Halstead WC: Brain and Intelligence. Chicago, University of Chicago Press, 1947