JBJA Journal of Bone and Joint Surgery - American 1996 - 1998
June 1997, Volume 79-A, Number 6
858 Demographic Biases of Scoring Instruments for the Results of Total Knee Arthroplasty*
Article

AUTHOR(S): BRINKER, MARK R., M.D.†, HOUSTON, TEXAS; LUND, PETER J., M.D.‡; BARRACK, ROBERT L., M.D.‡, NEW ORLEANS, LOUISIANA

Investigation performed at the Department of Orthopaedic Surgery, Tulane University School of Medicine, New Orleans, and the Fondren Orthopedic Group, Texas Orthopedic Hospital, Houston

J Bone Joint Surg [Am] 1997; 79-A; 858-65

ABSTRACT: **Four knee-scoring systems were used to evaluate 200 adult subjects who had no history of injury, abnormality, or treatment of the knees, hips, lower extremities, or spine. All subjects were in the age-range (fifty to 100 years; average, 65.5 years) typical of candidates for total knee replacement. In addition to a physical examination, complete demographic data were collected for each subject. The knee scores were normalized by dividing the observed score by the maximum possible score. The average normalized total knee score was 91 per cent (range, 22 to 100 per cent) according to the knee score of The Hospital for Special Surgery, 95 per cent (range, 10 to 100 per cent) according to the system of Hungerford and Kenna, 89 per cent (range, -7.75 to 100 per cent) according to a modification of the scoring system of The Knee Society, and 95 per cent (range, 26.5 to 100 per cent) according to the system of Hofmann et al. Demographic variables that had a significant negative correlation with the knee scores included advanced age (particularly of eighty-five years or more), a family income below the poverty level, and two major medical conditions or more.**

Observed differences in knee scores between different study groups that have not been matched for various clinically relevant factors are at least as likely to represent differences in the patient populations as they are to represent differences in the operative technique or the design of the implant.

Numerous scoring systems have been devised to evaluate patients who have symptoms related to the knee. Many of these systems have been used to assess the status of individuals before and after a total knee arthroplasty by assigning a numerical score for a variety of factors, including pain, function, range of motion, muscle strength, stability, deformity, and contracture. The use of a numerical score as a measure of outcome after total knee arthroplasty has had widespread appeal for the orthopaedic community and, as a result, several scoring systems have been introduced. Recently, however, hip and knee-scoring systems have been criticized and their validity has been questioned[2,3,5,6,10,13-15,17,24]. The lack of uniformity among the systems and the use of dissimilar terminology have been widely recognized[2,3,5,6,10,13,14,24]. A number of studies designed to analyze variations in scoring systems have shown measurable discrepancies in total scores[2,5,6].

With the recent growth of interest in outcome studies in orthopaedics, the need for validated scoring instruments has been recognized. Validity, however, can be difficult to define. According to Gross, a scoring system is internally valid if there is small observer variability within a given population and it is externally valid only if it controls for differences between study populations.

Differences between study populations might include risk factors (confounding variables) that have an unrecognized effect on the scores of subjects enrolled in a study. Greenfield used the term case mix to describe differences between study populations. He stated: “The term `case mix' refers to the features that increase the risk of a bad outcome or influence the choice of treatment. The purpose of case-mix adjustment is to separate the effects of the care given from those of the preexisting health status and other factors (such as age and socioeconomic status) that affect outcome measures.” While this phenomenon has been recognized, to our knowledge the effect of clinically relevant factors on numerically based knee-scoring systems has not been delineated.

The purpose of the current investigation was to establish normative data for four commonly used knee-scoring systems[18,19,21,22] and to define the effect, if any, of clinically relevant factors — that is, age, gender, race, relative body weight, socioeconomic status, and number of major medical conditions — on the knee scores. We did this by examining a group of subjects who had no history of injury, problems, or operations related to the knee, in order to eliminate the effects of abnormal conditions or treatments on the outcome measures (instruments) to be studied.

Materials and Methods

Three hundred and seventy-three volunteers who were fifty years old or more were interviewed and screened by one of us (M. R. B.) before inclusion in the study group. The volunteers were recruited (by an offer of a free examination) through community groups, churches and synagogues, and bulk mailers distributed throughout Louisiana. No information was provided regarding the purpose of the study or the type of history to be elicited or physical examination to be performed. One hundred and seventy-three volunteers were excluded from the study because they were unable to walk; had sought treatment for pain in the hip or knee in the past; had a known disease of the spine; had had an operation on the spine, hip, or knee; or had had a previous injury or abnormal condition of the lower extremities (other than peripheral vascular disease). Thus, 200 of the 373 volunteers met the eligibility criteria.

A detailed medical history was obtained, by one of

us (M. R. B.), from all 200 patients, and a physical examination was performed by a team of nine orthopaedic surgery residents, each of whom had attended a pre-study workshop where all terms, techniques, and methods of examination had been agreed on.

The average age of the subjects was 65.5 years (range, fifty to 100 years); there were fifty-seven men and 143 women. The average age and the gender distribution of our subjects very closely approximates that reported by Callahan et al., who performed a meta-analysis of 130 studies that included 9879 patients who had had a tricompartmental total knee replacement. In the present study, there were 154 white subjects and forty-three black subjects. This distribution approximates that for the state of Louisiana. Two of our subjects were Hispanic and one was Asian. Because of the small number of Hispanic and Asian subjects, no specific conclusion could be made regarding these ethnic groups.

The relative body weight was classified on the basis of height with the method of Stern and Insall. None of the subjects were underweight, 120 (60 per cent) were of normal weight, fifty-six (28 per cent) were mildly obese, sixteen (8 per cent) were moderately obese, and three (2 per cent) were severely obese. Five (3 per cent) of our subjects could not be classified because they were of such a short stature that their height was not found on the 1983 Metropolitan Life Insurance Company Table.

Medical conditions were recorded with use of our previous criteria[4]. A major medical condition was any condition that could potentially limit daily activities or walking, as defined by the consensus of a committee consisting of ten orthopaedic surgeons. Over-all, sixty-one subjects (31 per cent) had one major medical condition, forty (20 per cent) had two, ten (5 per cent) had three, nine (5 per cent) had four, and three (2 per cent) had five. The most prevalent major medical conditions were hypertension (seventy subjects; 35 per cent), a history of a malignant tumor (twenty-seven subjects; 14 per cent), angina pectoris (twenty-four subjects; 12 per cent), coronary artery disease (twenty-two subjects; 11 per cent), and diabetes mellitus that necessitated daily medication (twenty subjects; 10 per cent). One hundred and forty-seven subjects (74 per cent) took prescription medication for their medical conditions.

The average yearly family income of the 193 subjects who responded to this question was $23,800 (range, $0 to $111,000). Of the 181 subjects for whom complete data regarding the size of the family and its income were available, forty-six (25 per cent) had a family income that was below the poverty level[12]. This approximates the distribution of incomes for the state of Louisiana. One hundred and forty-three subjects (72 per cent) had some form of medical insurance.

The physical examinations were performed at three examination-specific stations. Each of the nine residents (three residents at each examination station) performed the same unique specific tasks for each of the subjects. For example, the arc of passive motion of both lower extremities was measured, by a team of three residents: one resident manipulated the lower extremity, one made the measurements with the goniometer, and one recorded the data to the nearest 5 degrees. Quadriceps muscle strength was tested and was recorded according to the guidelines of each of four knee-scoring systems[18,19,21,22].

Data were recorded on standardized flow sheets so that four commonly used knee scores could be calculated: the knee score of The Hospital for Special Surgery[22], the system of Hungerford and Kenna, a modification of the scoring system of The Knee Society[21], and the system of Hofmann et al. A score was calculated for both the right and the left knee of each subject; we then averaged the two scores together to obtain a single score for statistical analysis.

Statistical analysis was performed with the Student t test, the chi-square test, and analysis of variance with use of the SAS statistical package (SAS Institute, Cary, North Carolina). The result was considered significant if the p value was 0.05 or less. We have included p values or upper bounds whenever available. When multiple comparisons were made, we considered the results from the Tukey Studentized range test (a variant of the Newman-Keuls test) and the Bonferroni correction. The two methods led to identical conclusions in all cases at the 0.05 level of significance.

For the purpose of statistical analysis, the total knee scores and the component scores were normalized by dividing the observed score by the maximum possible score. For example, a subject who has a total score of 160 points and a pain score of 45 points, according to the system of The Knee Society, would have a normalized total score of 80 per cent (160 divided by 200) and a normalized pain score of 90 per cent (45 divided by 50). The data were analyzed to determine the effect of clinically relevant factors, such as age, gender, race, relative body weight, socioeconomic status, and number of major medical conditions, on the normalized total and component scores. The component scores were for pain; function; range of motion; muscle strength; and stability, deformity, and contracture. These clinically relevant factors were selected because it was the consensus of a research committee at our institution that they were the most likely to have an effect on the knee score. We acknowledge that a variety of other variables could have been studied and, in this regard, our study should be considered preliminary.

Results

The normalized total scores according to the systems of Hungerford and Kenna (95 per cent) and Hofmann et al. (95 per cent) were significantly higher than those according to the knee score of The Hospital for Special Surgery (91 per cent) and the system of The Knee Society (89 per cent) ($p < 0.05$) (Table I). The normalized scores for range of motion and for stability, deformity, and contracture were highest with the system of Hungerford and Kenna and that of Hofmann et al. One hundred and seventy-four (87 per cent) of the subjects had a normalized total knee score of 90 to 100 per cent with those systems,

compared with 139 (70 per cent) and 126 (63 per cent) with the systems of The Hospital for Special Surgery and The Knee Society, respectively (Table II).

Forty-six subjects (23 per cent) reported that their ability to walk was limited by at least one factor (Table III). Nineteen subjects (10 per cent) had a medical condition that was manifested as shortness of breath, chest pain, generalized weakness, or other symptoms; twenty-three (12 per cent) had a musculoskeletal condition such as pain, stiffness, muscle weakness, or fatigue in the hip or knee; and four (2.0 per cent) had both.

Significant differences in normalized total scores were seen among age-groups with each of the four scoring systems, with older subjects having lower scores ($p < 0.0001$). The subjects who were eighty-five years old or more had significantly lower scores than all of the other age-groups, according to all four scoring systems ($p < 0.05$). No significant differences were observed among the age-groups of less than eighty-five years except that, according to the system of The Knee Society, the subjects who were eighty to eighty-four years old had significantly lower scores than those who were fifty to seventy-four years old ($p < 0.05$). There were no significant differences in the normalized pain scores between the age-groups according to the knee score of The Hospital for Special Surgery ($p = 0.39$) or the systems of The Knee Society ($p = 0.29$) and Hofmann et al. ($p = 0.14$); however, older subjects had significantly lower scores with the system of Hungerford and Kenna ($p = 0.04$). This effect was noted when the subjects who were eighty-five years old or more were compared with those who were less than eighty-five years old ($p < 0.05$). The subjects who were eighty-five years old or more had significantly lower normalized function scores with the knee score of The Hospital for Special Surgery and the scoring system of The Knee Society ($p < 0.0001$). The system of Hungerford and Kenna and that of Hofmann et al. do not have a function component. With all four scoring systems, there were significant differences in the normalized scores for range of motion and for stability, deformity, and contracture among the age-groups, with older subjects having lower scores ($p < 0.0005$). Similarly, older subjects also had significantly lower normalized scores for muscle strength according to the knee score of The Hospital for Special Surgery, the system of Hungerford and Kenna, and that of Hofmann et al. ($p < 0.0001$) (Table IV). The system of The Knee Society does not have a muscle strength component.

Over-all, there were no significant differences in the normalized total scores between men and women with any of the four scoring systems ($p > 0.40$). The situation was similar regarding the normalized component scores ($p > 0.10$). Similarly, there were no significant differences among the normalized total or component scores among the five classes of relative body weight ($p > 0.10$) (Table IV).

When socioeconomic factors were analyzed, there was a relationship between race and income. An analysis of the normalized total and component scores according to race and level of income revealed no significant differences on the basis of race alone with any of the four scoring systems ($p > 0.20$). The subjects whose family income was below the poverty level had significantly lower normalized total scores according to the knee score of The Hospital for Special Surgery ($p = 0.022$) and the systems of Hungerford and Kenna ($p = 0.035$) and The Knee Society ($p = 0.023$). Marginally lower scores were observed for impoverished subjects with the system of Hofmann et al. ($p = 0.058$). Subjects who had a family income that was below the poverty level had significantly lower normalized pain scores with the knee score for The Hospital for Special Surgery ($p = 0.039$) but not with the other three scoring systems ($p > 0.10$). Significantly lower normalized function scores were observed for impoverished subjects with the system of The Knee Society ($p = 0.031$). Impoverished subjects had significantly lower normalized scores for range of motion with the knee score of The Hospital for Special Surgery ($p = 0.021$). This effect was not observed with the other three scoring systems ($p = 0.085$, 0.20, and 0.24). Impoverished subjects had significantly lower normalized scores for muscle strength with the system of Hungerford and Kenna ($p = 0.031$) and that of Hofmann et al. ($p = 0.04$). A significant relationship between the normalized scores for stability, deformity, and contracture and family income level was observed only with the scoring system of The Knee Society ($p = 0.038$) (Table IV).

Significant differences in the normalized total scores according to two of the four systems were observed on the basis of the number of major medical conditions. Subjects who had two major medical conditions or more had significantly lower normalized total scores according to the knee score of The Hospital for Special Surgery ($p = 0.012$) and the system of The Knee Society ($p = 0.017$). This relationship persisted when we accounted for income level ($p < 0.05$) and age ($p < 0.05$), suggesting that the number of major medical conditions is an important predictor of the knee score. A significant relationship was observed between the number of major medical conditions and the normalized function scores according to the knee score of The Hospital for Special Surgery and the system of The Knee Society, with subjects who had two major medical conditions or more having lower scores ($p < 0.005$) (Table IV). There were also significant differences in the normalized scores for range of motion, according to these two systems, on the basis of the number of major medical conditions ($p < 0.05$) (Table IV). There were no significant differences in the normalized scores for pain; muscle strength; or stability, deformity, and contracture, with any of the scoring systems, on the basis of the number of major medical conditions ($p > 0.05$) (Table IV).

Discussion

The results of this study suggest that several clinically relevant factors have an effect on the total and component scores of numerically based knee-scoring systems. Factors that have a significant effect

($p \leq 0.05$) include age, family income, and the number of medical conditions. Factors that do not appear to have a significant effect include gender, race, and relative body weight.

In the current study, age had a significant effect on all four normalized total knee scores, one normalized pain score, two normalized function scores, and all other normalized component scores of the four scoring systems. We[4] as well as Ilstrup et al. have reported similar findings of lower hip scores in older patients, and Constant reported poorer function of the shoulder in older patients who were recovering from an injury of the shoulder. While it is clear that none of the subjects in the present study were followed longitudinally (this was a cross-sectional study) to allow observation of individual changes in total and component scores with time, the significant trend of lower scores for our subjects who were eighty-five years old or more suggests a diminution in strength and function with aging. The average normalized total score according to the system of Hungerford and Kenna was 96 per cent for subjects who were fifty-five to fifty-nine years old, compared with only 70 per cent for those who were eighty-five years old or more. Conclusions drawn from studies of patients followed for ten or twenty years after a total knee arthroplasty must be tempered by the fact that an observed decline in the knee scores may represent the natural morbidity of aging in a patient who has an otherwise well functioning replacement.

No significant relationship was observed between gender or relative body weight and the normalized total or component knee scores. Similarly, there was no significant relationship between race and the normalized total or component knee scores; however, we did observe a significant relationship between race and socioeconomic status: a greater relative proportion of black subjects had a family income level below the poverty level[12] as compared with white subjects. This observation has been reported by other authors[1,11]. When the effect of clinical factors on knee scores is assessed, investigators must be careful not to draw false conclusions on the basis of confounding variables such as race and socioeconomic status.

There was a significant relationship between socioeconomic status and the knee scores. Impoverished subjects had significantly lower normalized total scores according to three of the four systems; the relationship between socioeconomic status and the knee score was marginally significant with the fourth system. Similarly, there was a significant relationship between socioeconomic status and the normalized scores for pain and for range of motion according to the knee score of The Hospital for Special Surgery; the normalized scores for function and for stability, deformity, and contracture according to the scoring system of The Knee Society; and the normalized score for muscle strength according to the system of Hungerford and Kenna and that of Hofmann et al.

A number of authors have addressed the complex relationship between socioeconomic status and health[1,25,27,29]. Adler et al. noted that "socioeconomic status ... is a strong and consistent predictor of morbidity and premature mortality." Syme and Berkman noted that individuals in lower socioeconomic groups tend to have higher rates of disability, morbidity, and mortality. The three most recognizable components of socioeconomic status are income level, education, and occupational status[1]. In the present study, we chose to use income level as our determinant of socioeconomic status because this information is quantifiable and was readily available for most (193) of our subjects. Adler et al. suggested that the poorer over-all health seen in lower socioeconomic groups is related to health-risk behaviors (smoking, use of alcohol, and so on), differential exposure to physical and social situations, stress, and lack of control over work circumstances. It is likely that many of these mechanisms affected the knee scores in our study; subjects who had a lower socioeconomic status (below the poverty level[12]) tended to report more pain, less function, a decreased range of motion and muscle strength, and poorer scores for stability, deformity, and contracture with at least one of the scoring systems. While the complex relationship between socioeconomic factors and knee scores remains somewhat obscure, it is clear that the effect of income level on total and component scores must be controlled for when the results of total knee replacements are analyzed.

We also observed a significant relationship between the knee scores and the number of major medical conditions. Subjects who had two major medical conditions or more had lower normalized total, function, and range-of-motion scores according to two of the four scoring systems. Charnley recognized the importance of factoring medical conditions into evaluations of the hip when he described category-C patients as those who have a condition that directly impairs walking and noted that different categories of patients should not be compared. Liang et al. noted the importance of medical conditions as a comorbidity and stated: "Concurrent active medical or operative problems may be associated with pain or with loss of function, potentially confounding the outcome of total hip arthroplasty." It is likely that Charnley's advice regarding total hip arthroplasty in category-C patients is applicable to total knee arthroplasty. The results of our investigation support this concept insofar as our subjects who had two major medical conditions or more reported poorer function and had significantly lower knee scores.

It is interesting to note that the average normalized total knee scores in our group of adults who were fifty years old or older was lower than might be anticipated on the basis of the reported scores in series of patients who have had a knee replacement. The normalized total scores for our subjects averaged 91, 95, 89, and 95 per cent according to the knee score of The Hospital for Special Surgery and the systems of Hungerford and Kenna, The Knee Society, and Hofmann et al., respectively.

With regard to age, gender distribution, and prevalence of major medical conditions, a review of the medical literature revealed only one study[23] in

which all of these data were reported for patients who had had a total hip replacement. Our study group was comparable with that series in terms of age and gender distribution and the prevalence of hypertension; however, the prevalence of coronary artery and chronic obstructive pulmonary disease was lower in our group. On the basis of this fact, we believe that our recruitment strategy was not more likely to attract less healthy people.

Every effort was made to avoid the introduction of selection bias into our study group. All of the volunteers were recruited without our providing knowledge of the purpose of the study or the type of examination to be performed. Despite these efforts, we acknowledge that subjects who volunteer for a free examination may have an underlying physical or psychological ailment that is not apparent at the time of recruitment. While our study group appeared to be representative of the population of Louisiana (in terms of race and family income), it is possible that socioeconomic differences and secondary gain may have played a role in our recruitment process.

We found numerous significant comparisons in this study as well as numerous comparisons that were not significant and that we did not report. It is possible that at least one of the claimed significant differences is spurious — that is, it resulted from chance alone. Such a possibility always exists, and its likelihood increases with the number of comparisons made. We made no attempt to correct for this phenomenon in such a complex study because our goal was not to make confirmatory findings but rather to report preliminary observations that raise the issue that knee-scoring systems may be biased by many demographic variables that are not relevant to their intended use. Thus, significance was used primarily as a means for us to extract potentially meaningful hypotheses that can be subjected to additional scrutiny if they are deemed to be of sufficient interest to the orthopaedic community.

Gartland discussed the need for a control group in studies of outcomes of total joint arthroplasty. It has become common practice for scores of 80 to 89 (of a possible 100) points to be considered good and for those of 90 to 100 points to be considered excellent. On review of the orthopaedic literature, it becomes apparent that this practice developed as a matter of convention rather than as a result of the scientific method. If a total knee score of 93 points is to be considered excellent, it should be excellent in comparison with a data set of controls. In the absence of a data set, such scores reported in the literature should be considered of questionable value. In the present study, we present a data set that may serve as a useful control group when the results of total knee arthroplasty are reported.

The practice of comparing series of total knee arthroplasties to draw conclusions about operative techniques or implant designs on the basis of knee scores should be reconsidered. The results of the current investigation suggest that problems related to demographic differences represent confounding variables that must be accounted for if comparisons between study groups are to be meaningful.

NOTE: The authors thank Dr. Douglas Waldman, Dr. James Dunlap, Dr. Rolando Garcia, Dr. Randall Jennings, Dr. Andrew Palafox, Dr. Paul Pflueger, Dr. Bill Tejeiro, and Dr. Michael Wolfe for their assistance in completing this study; Marc Elliott, Ph.D., for performing all statistical analyses; and Michele LeBouef and Becky Trahan for technical assistance with the manuscript.

References: (1-29)

1. Adler, N. E.; Boyce, W. T.; Chesney, M. A.; Folkman, S.; and Syme, S. L.: Caring for the uninsured and underinsured: socioeconomic inequalities in health: no easy solution. J. Am. Med. Assn. 1993; 269:3140-3145.
2. Andersson, G.: Hip assessment: a comparison of nine different methods. J. Bone and Joint Surg. 1972; 54-B(4):621-625.
3. Apley, A. G.: Editorial. An assessment of assessment. J. Bone and Joint Surg. 1990; 72-B(6):957-958.
4. Brinker, M. R.; Lund, P. J.; Cox, D. D.; and Barrack, R. L.: Demographic biases found in scoring instruments of total hip arthroplasty. J. Arthroplasty 1996; 11:820-830.
5. Bryant, M. J.; Kernohan, W. G.; Nixon, J. R.; and Mollan, R. A. B.: A statistical analysis of hip scores. J. Bone and Joint Surg. 1993; 75-B(5):705-709.
6. Callaghan, J. J.; Dysart, S. H.; Savory, C. F.; and Hopkinson, W. J.: Assessing the results of hip replacement. A comparison of five different rating systems. J. Bone and Joint Surg. 1990; 72-B(6):1008-1009.
7. Callahan, C. M.; Drake, B. G.; Heck, D. A.; and Dittus, R. S.: Patient outcomes following tricompartmental total knee replacement. A meta-analysis. J. Am. Med. Assn. 1994; 271:1349-1357.
8. Charnley, J.: Numerical grading of clinical results. In Low Friction Arthroplasty of the Hip. Theory and Practice, pp. 20-24. New York, Springer, 1979.
9. Constant, C. R.: Age related recovery of shoulder function after injury. Master's Thesis, University College, Cork, Ireland, Aug. 1986.
10. Dorey, F., and Amstutz, H. C.: Discrepancies in the orthopaedic literature: why? A statistical explanation. In Instructional Course Lectures, The American Academy of Orthopaedic Surgeons. Vol. 42, pp. 555-564. Rosemont, Illinois, The American Academy of Orthopaedic Surgeons, 1993.
11. Egbert, L. D.; and Rothman, I. L.: Relation between the race and economic status of patients and who performs their surgery. New England J. Med. 1977; 297:90-91.
12. Federal Registry: Federal Poverty Income Guidelines for Fiscal Year 1993, pp. 5455-5457. Washington, D.C., Government Printing Office, Feb. 12, 1992.

13. Galante, J.: Editorial. The need for a standardized system for evaluating results of total hip surgery. J. Bone and Joint Surg. April 1985; 67-A:511-512.
14. Galante, J.: Editorial. Evaluation of results of total hip replacement. J. Bone and Joint Surg. Feb. 1990; 72-A:159-160.
15. Gartland, J. J.: Orthopaedic clinical research. Deficiencies in experimental design and determinations of outcome. J. Bone and Joint Surg. Oct. 1988; 70-A:1357-1364.
16. Greenfield, S.: Editorial. The state of outcome research: are we on target? New England J. Med. 1989; 320:1142-1143.
17. Gross, M.: A critique of the methodologies used in clinical studies of hip-joint arthroplasty published in the English-language orthopaedic literature. J. Bone and Joint Surg. Oct. 1988; 70-A:1364-1371.
18. Hofmann, A. A.; Murdock, L. E.; Wyatt, R. W. B.; and Alpert, J. P.: Total knee arthroplasty. Two- to four-year experience using an asymmetric tibial tray and a deep trochlear-grooved femoral component. Clin. Orthop. 1991; 269:78-88.
19. Hungerford, D. S.; and Kenna, R. V.: Preliminary experience with a total knee prosthesis with porous coating used without cement. Clin. Orthop. 1983; 176:95-107.
20. Ilstrup, D. M.; Nolan, D. R.; Beckenbaugh, R. D.; and Coventry, M. B.: Factors influencing the results in 2,012 total hip arthroplasties. Clin. Orthop. 1973; 95:250-262.
21. Insall, J. N.; Dorr, L. D.; Scott, R. D.; and Scott, W. N.: Rationale of the Knee Society clinical rating system. Clin. Orthop. 1989; 248:13-14.
22. Insall, J. N.; Ranawat, C. S.; Aglietti, P.; and Shine, J.: A comparison of four models of total knee-replacement prostheses. J. Bone and Joint Surg. Sept. 1976; 58-A:754-765.
23. Koide, M.; Pilon, R. N.; Vandam, L. D.; and Lowell, J. D.: Anesthetic experience with total hip replacement. Clin. Orthop. 1974; 99:78-85.
24. Liang, M. H.; Katz, J. N.; Phillips, C.; Sledge, C.; Cats-Baril, W.; and The American Academy of Orthopaedic Surgeons Task Force on Outcome Studies: The total hip arthroplasty outcome evaluation form of The American Academy of Orthopaedic Surgeons. J. Bone and Joint Surg. June 1991; 73-A:639-646.
25. McLeod, J. D.; and Kessler, R. C.: Socioeconomic status differences in vulnerability to undesirable life events. J. Health and Soc. Behav. 1990; 31:162-172.
26. Metropolitan Life Insurance Company: Metropolitan height and weight tables. Statist. Bull. Metropol. Insur. Co., 64: 2-9, 1983.
27. Otten, M. W., Jr.; Teutsch, S. M.; Williamson, D. F.; and Marks, J. S.: The effect of known risk factors on the excess mortality of black adults in the United States. J. Am. Med. Assn.

TABLE I
THE TOTAL, COMPONENT, AND NORMALIZED SCORES FOR THE TWO HUNDRED SUBJECTS

Knee-Scoring System	Total Score* *(Points)*	Normalized Total Score *(Per cent)*	Pain Score* *(Points)*	Normalized Pain Score *(Per cent)*
The Hospital for Special Surgery[22]	90.8 (22 to 100)	91	27.1 (0 to 30)	90
Hungerford and Kenna	95.3 (10 to 100)	95	47.0 (0 to 50)	94
The Knee Society[21]	177.6 (-15.5 to 200)	89	45.2 (0 to 50)	90
Hofmann et al.	94.9 (26.5 to 100)	95	37.0 (0 to 40)	93

*The score is given as the average with the range in parentheses.
†This component is not part of the system of Hungerford and Kenna or that of Hoffman et al.
‡This component is not part of the system of The Knee Society.

Normalized Function Score *(Per cent)*	Range-of-Motion Score* *(Points)*	Normalized Range-of-Motion Score *(Per cent)*	Muscle Strength Score*‡ *(Points)*	Normalized Muscle Strength Score *(Per cent)*
90	15.6 (9 to 18)	87	9.7 (5 to 10)	97
—	19.7 (10 to 20)	98	9.3 (0 to 10)	93
87	23.9 (12.5 to 25)	96	—	—
—	24.8 (17.5 to 25)	99	9.3 (0 to 10)	93

Stability, Deformity, Contracture Score* *(Points)*	Normalized Stability, Deformity, Contracture Score *(Per cent)*
18.8 (5 to 20)	94
19.4 (0 to 20)	97
22.1 (-20 to 25)	88
23.8 (9 to 25)	95

TABLE III

DATA FOR THE TWO HUNDRED SUBJECTS (FOUR HUNDRED KNEES) ACCORDING TO A MODIFICATION OF THE SCORING SYSTEM OF THE KNEE SOCIETY[21]

Component	No.*
Pain	400
None	276 (69%)
Mild or occasional	64 (16%)
Stairs only	7 (2%)
Walking and stairs	16 (4%)
Moderate	
Occasional	29 (7%)
Continual	5 (1%)
Severe	3 (<1%)
Range of motion	400
≥125°	232 (58%)
120 to 124°	101 (25%)
110 to 119°	36 (9%)
90 to 109°	18 (5%)
<90°	13 (3%)
Stability (maximum movement in any position)	400
Anteroposterior	
<5 mm	395 (99%)
5 to 10 mm	5 (1%)
>10 mm	0
Mediolateral	
<5°	378 (95%)
6 to 9°	5 (1%)
10 to 14°	17 (4%)
15°	0
Contractures and alignment	400
Flexion contracture	
None	365 (91%)
5 to 10°	28 (7%)
11 to 15°	4 (1%)
16 to 20°	1 (<1%)
>20°	2 (<1%)
Extension lag	
None	390 (98%)
<10°	9 (2%)
10 to 20°	0
>20°	1 (<1%)
Alignment (valgus)	
5 to 10°	326 (82%)
<5 or >10°	74 (19%)
Function	200
Walking	
Unlimited	154 (77%)
>10 blocks	3 (2%)
5 to 10 blocks	11 (6%)
<5 blocks	25 (13%)
Household	6 (3%)
Unable	1 (<1%)
Stairs	
Normal up and down	134 (67%)
Normal up, down with rail	11 (6%)
Up and down with rail	49 (25%)
Up with rail, down unable	2 (1%)
Unable	4 (2%)

Walking aids	200
None	188 (94%)
One cane	7 (4%)
Two canes	0
Crutches or walker	5 (3%)

*The numbers for pain, range of motion, stability, and contraction and alignment are given for knees, and the numbers for function and walking aids are given for subjects.

TABLE IV
RELATIONSHIPS BETWEEN NORMALIZED TOTAL AND COMPONENT SCORES AND PATIENT VARIABLES*

Knee-Scoring System	Age	Gender	Relative Body Weight	Race	Family Income below Poverty Level	Two Major Medical Conditions or More
The Hospital for Special Surgery[22]						
Normalized score						
Total	<0.0001	NS	NS	NS	0.022	0.012
Pain	NS	NS	NS	NS	0.039	NS
Function	<0.0001	NS	NS	NS	NS	0.0045
Range of motion	<0.0001	NS	NS	NS	0.021	0.0037
Muscle strength	<0.0001	NS	NS	NS	NS	NS
Stability, deformity, contracture	<0.0001	NS	NS	NS	NS	NS
Hungerford and Kenna						
Normalized score						
Total	<0.0001	NS	NS	NS	0.035	NS
Pain	0.04	NS	NS	NS	NS	NS
Range of motion	<0.0001	NS	NS	NS	NS	NS
Muscle strength	<0.0001	NS	NS	NS	0.031	NS
Stability, deformity, contracture	<0.0001	NS	NS	NS	NS	NS
The Knee Society[21]						
Normalized score						
Total	<0.0001	NS	NS	NS	0.023	0.017
Pain	NS	NS	NS	NS	NS	NS
Function	<0.0001	NS	NS	NS	0.031	0.0036
Range of motion	0.0042	NS	NS	NS	NS	0.036
Stability, deformity, contracture	<0.0001	NS	NS	NS	0.038	NS
Hofmann et al.						
Normalized score						
Total	<0.0001	NS	NS	NS	NS	NS
Pain	NS	NS	NS	NS	NS	NS
Range of motion	0.0005	NS	NS	NS	NS	NS
Muscle strength	<0.0001	NS	NS	NS	0.04	NS
Stability, deformity, contracture	<0.0001	NS	NS	NS	NS	NS

*NS = not significant. The component score for function is not part of the system of Hungerford and Kenna or that of Hoffman et al., and the component score for muscle strength is not part of the scoring system of The Knee Society.