Clin Orthop Relat Res (2015) 473:94–100
DOI 10.1007/s11999-014-3744-1

Clinical Orthopaedics and Related Research®
A Publication of The Association of Bone and Joint Surgeons®

SYMPOSIUM: 2014 KNEE SOCIETY PROCEEDINGS

What Can Be Learned From Minimum 20-year Followup Studies of Knee Arthroplasty?

John J. Callaghan MD, Christopher T. Martin MD, Yubo Gao PhD, Andrew J. Pugely MD, Steve S. Liu MD, Devon D. Goetz MD, Scott S. Kelley MD, Richard C. Johnston MD

Published online: 8 July 2014

Abstract

Background Long-term evaluation of knee arthroplasty should provide relevant information concerning the durability and performance of the implant and the procedure. Because most arthroplasties are performed in older patients, most long-term followup studies have been performed in elderly cohorts and have had low patient survivorship to final followup; the degree to which attrition from patient deaths over time in these studies might influence their results has been poorly characterized.

Questions/purposes The purpose of this study was to examine the results at 20-year followup of two prospectively followed knee arthroplasty cohorts to determine the following: (1) Are there relevant differences among the two implant cohorts in terms of revision for aseptic causes (osteolysis, or loosening)? (2) How does patient death over the long followup interval influence the comparison, and do the comparisons remain valid despite the high attrition rates?

Methods Two knee arthroplasty cohorts from a single orthopaedic practice were evaluated: a modular tibial tray (101 knees) and a rotating platform (119 knees) design. All patients were followed for a minimum of 20 years or until death (mean, 14.1 years; SD 5.0 years). Average age at surgery for both cohorts was > 70 years. The indications for the two cohorts were identical (functionally limiting knee pain) and was surgeon-specific (each surgeon performed all surgeries in that cohort). Revision rates through a competing risks analysis for implants and survivorship curves for patients were evaluated.

Results Both of these elderly cohorts showed excellent implant survivorship at 20 years followup with only small differences in revision rates (6% revision versus 0% revision for the modular tibial tray and rotating platform, respectively). However, attrition from patient deaths was

One of the authors certifies that he (JJC), or a member of his immediate family, has or may receive payments or benefits, during the study period, an amount more than USD 1,000,001 from DePuy (Warsaw, IN, USA) and less than USD 10,000 from Lippincott Williams & Wilkins (Baltimore, MD, USA). We (CTM, AJP) acknowledge use of the Bierbaum Research Fund (an institutional resident research fund).
All ICMJE Conflict of Interest Forms for authors and *Clinical Orthopaedics and Related Research*® editors and board members are on file with the publication and can be viewed on request.

Each author certifies that his or her institution approved or waived approval for the human protocol for this investigation and that all investigations were conducted in conformity with ethical principles of research.
This work was performed at the University of Iowa, Iowa City, IA, USA, and Des Moines Orthopaedic Surgeons, West Des Moines, IA, USA.

J. J. Callaghan (✉), S. S. Liu
Department of Orthopaedic Surgery, University of Iowa Hospitals and Clinics, 200 Hawkins Drive, 01029 JPP, Iowa City, IA 52242, USA
e-mail: John-Callaghan@uiowa.edu

J. J. Callaghan, R. C. Johnston
Department of Orthopaedic Surgery, VA Medical Center, Iowa City, IA, USA

C. T. Martin, A. J. Pugely
Department of Orthopaedic Surgery, University of Iowa Hospitals and Clinics, 200 Hawkins Drive, 01051 JPP, Iowa City, IA 52242, USA

Y. Gao, R. C. Johnston
Department of Orthopaedic Surgery, University of Iowa Hospitals and Clinics, 200 Hawkins Drive, 01066 JPP, Iowa City, IA 52242, USA

substantial and overall patient survivorship to 20-year followup was only 26%. Patient survivorship was significantly higher in patients < 65 years of age in both cohorts (54% versus 15%, p < 0.001 modular tray cohort, and 52% versus 26%, p = 0.002 rotating platform cohort). Furthermore, in the modular tray cohort, patients < 65 years had significantly higher revision rates (15% versus 3%, p = 0.0019).

Conclusions These two cohorts demonstrate the durability of knee arthroplasty in older patients (the vast majority older than 65 years). Unfortunately, few patients lived to 20-year followup, thus introducing bias into the analysis. These data may be useful as a reference for the design of future prospective studies, and consideration should be given to enrolling younger patients to have robust numbers of living patients at long-term followup.

Level of Evidence Level III, therapeutic study. See Guidelines for Authors for a complete description of levels of evidence.

Introduction

Although Level I evidence is considered important for guiding clinical decision-making, this is impractical when it comes to evaluating the long-term durability and function of knee arthroplasty implants. To date, performing long-term longitudinal studies of specific devices has provided the best available evidence regarding the implant design characteristics most likely to provide lasting durability and satisfactory patient function.

However, because most arthroplasties are performed in older patients, most long-term followup studies have been performed in elderly cohorts and have had low patient survivorship to final followup. The majority of prior studies, including our own [1–4, 8, 14, 15], have used a Kaplan-Meier (KM) survivorship analysis to report revision rates [12]. A KM analysis reports the time to the event of interest, in this case revision of the implant, and assumes that the event happens independently from other potential competing events. However, death is a competing risk against revision in a long-term followup study. If a patient dies, they cannot possibly be revised. In a KM analysis, patients with a competing event are censured from the final result, introducing significant bias. This type of bias is particularly evident in elderly cohorts, which have high attrition from patient deaths, and prior authors have noted that this not only greatly diminishes the statistical power of the conclusions, but also tends to overestimate revision rates [7, 11].

D. D. Goetz
Des Moines Orthopaedic Surgeons, West Des Moines, IA, USA

S. S. Kelley
North Carolina Orthopaedic Clinic, Durham, NC, USA

Table 1. Patient demographics

Demographic	PFC cohort	LCS cohort
Mean age (years)	71 (range, 52–89)	70 (range, 37–81)
Sex (percent female)	59	56
Number of knees	101	118*
Number of patients	75	86
Diagnosis		
Primary osteoarthritis	85%	88%
Rheumatoid arthritis	13%	10%
Posttraumatic arthritis	1%	2%
Osteonecrosis	1%	0%

PFC = Press-Fit Condylar knee; LCS = Low Contact Stress.
* One lost to followup.

As a result, recent authors have advocated for the use of a cumulative incidence of competing risk analysis (CI), in which patients with a death are not censored from the results [7]. Compared with a KM analysis, which answers the question, “What is the risk of the event if no one ever dies?,” the CI analysis more directly answers the question, “What is the risk of the event?” [11].

In light of these potential biases, the purpose of the current study was to shed light on what can and cannot be learned from currently available long-term followup studies of knee arthroplasty designs. First, we provide an example of a CI analysis with minimum 20-year followup comparing two implant cohorts in terms of revision for aseptic causes (osteolysis, or loosening) to determine if relevant comparisons can be made across elderly cohorts of patients undergoing knee arthroplasty. Second, we more specifically investigate patient survivorship over the 20-year followup and attempt to determine how patient deaths influence the comparison of these cohorts. Data from the second aim may be useful in guiding the design of future prospective long-term followup studies.

Materials and Methods

This study received an exception from the institutional review board and was HIPAA-compliant. The methodology for each prior cohort’s review has previously been published [1, 4]. In brief, two prospective series of knee arthroplasty cohorts were performed in a single orthopaedic practice: a modular tibial tray (101 knees) and a rotating platform (119 knees). Demographics of the cohorts were similar (Table 1). All patients were followed longitudinally for over 20 years or until death. Followup evaluations were performed by a single surgeon (DDG) not involved in the initial surgical care of the patients. Radiographs were evaluated by two independent observers (GF, DH, MI, AM, MS) with agreement by consensus at each followup interval. One observer (JJC) reviewed all radiographs at each followup interval of both cohorts using the Knee Society radiographic assessment [6].

The first cohort, consisting of 101 knees in 75 patients operated on between 1988 and 1991, received a modular posterior cruciate-retaining Press-Fit Condylar (PFC) prosthesis (Johnson and Johnson Professional, Inc, Raynham, MA, USA) [1]. The mean age at the time of surgery was 71 years (range, 52–89 years). Diagnosis was primary osteoarthritis in 86 (85%) knees (Table 1). No patients were lost to followup at 20 years.

The second cohort, consisting of 119 knees in 86 patients operated on between 1985 and 1988, received the cemented Low Contact Stress (LCS) rotating platform tibial and femoral implants mated with a cemented Townley all-polyethylene dome patellar component (DePuy, Warsaw, IN, USA) [4]. The mean age at the time of the original surgery was 70 years (range, 37–81 years). Diagnosis was osteoarthritis in 105 (88%) knees (Table 1). One patient (one knee) was lost to followup at 20 years.

The indications for surgery using the two implants studied (LCS and PFC) were identical (functionally limiting knee pain). They were performed in the same practice in close to a sequential time interval (1985–1988 and 1988–1991, respectively).

For statistical analysis, we compared implant survivorship across cohorts according to CI methods [7, 11]. Because a patient death precludes revision surgery, patient death and implant revision are competing risks in long-term followup studies. Thus, a competing risk analysis allows for an assessment of implant revision rates while taking into account the competing risk of patient death during the study period. The primary endpoint was revision for aseptic implant failure and radiographic loosening. In contrast to implant survivorship, patient survivorship is a binary outcome for which there is no competing risk. Thus, to determine the relationship of patient age at the time of surgery to likelihood of patient survivorship to final followup, we performed survivorship analysis of the patients themselves out to minimum 20-year followup using

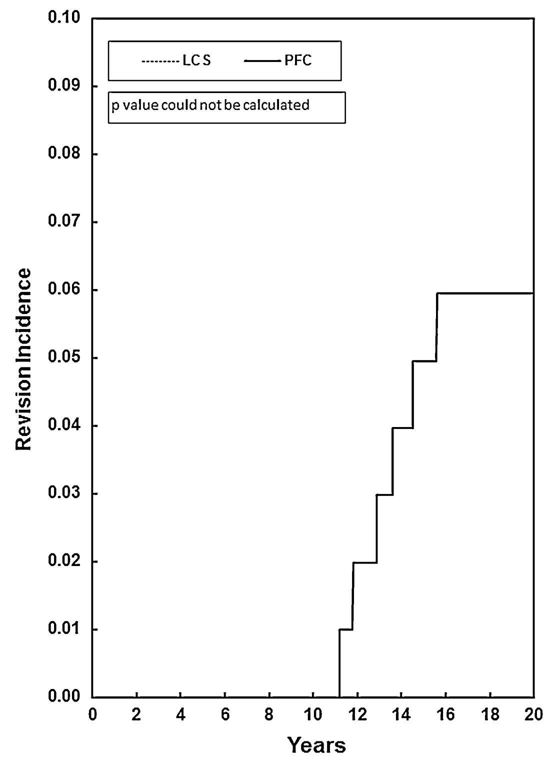

Fig. 1 Competing risk analysis of implant failure (for osteolysis, or implant loosening) as the endpoint for the two cohorts was evaluated. The incidence of revision was higher in the modular tray cohort (PFC) as compared with the rotating platform cohort (LCS), but no statistical comparison could be made.

Table 2. Comparison of incidence of mortality and revision at 20 years followup

Patient cohort	Number of knees	Incidence of revision	Incidence of mortality
PFC knee			
All patients	101	6% (6 knees)	75% (76 knees)
Younger than 65 years	26	15% (4 knees)	46% (12 knees)
Older than 65 years	75	3% (2 knees)	85% (64 knees)
LCS knee			
All patients	118	0% (0 knees)	74% (87 knees)
Younger than 65 years	33	0% (0 knees)	48% (16 knees)
Older than 65 years	85	0% (0 knees)	84% (71 knees)
All implant types			
All patients	219	3% (6 knees)	74% (163 knees)
Younger than 65 years	59	7% (4 knees)	47% (28 knees)
Older than 65 years	160	1% (2 knees)	84% (135 knees)

PFC = Press-Fit Condylar knee; LCS = Low Contact Stress.

KM methods [12]. Curves were truncated at 20 years in each analysis for similar comparison across cohorts. Statistical analysis was performed using SPSS 13.0 software (SPSS Inc, Chicago, IL, USA). Cumulative risk analysis requires at least one event in each compared cohort to calculate a p value.

Results

Relevance of Revision

Overall durability was excellent in both cohorts with only six knees (6%) and zero knees (0%) revised for loosening in the PFC and LCS cohorts, respectively (Table 2; Fig. 1). Because there were no revisions in the elderly cohort of LCS knees, we cannot estimate CIs or provide a p value for the comparison between the two cohorts. To determine if patient age played a role in the incidence of revision, we substratified the PFC group according to patient age. After stratifying, the incidence of revision was much higher in patients aged < 65 years (15%; 95% CI, 5%–32%) as compared with patients > 65 years (3%; 95% CI, 0.5%–8%) ($p = 0.0188$) (Fig. 2A). Again, a similar comparison could not be made for the LCS cohort because the overall incidence of revision was 0% (Fig. 2B).

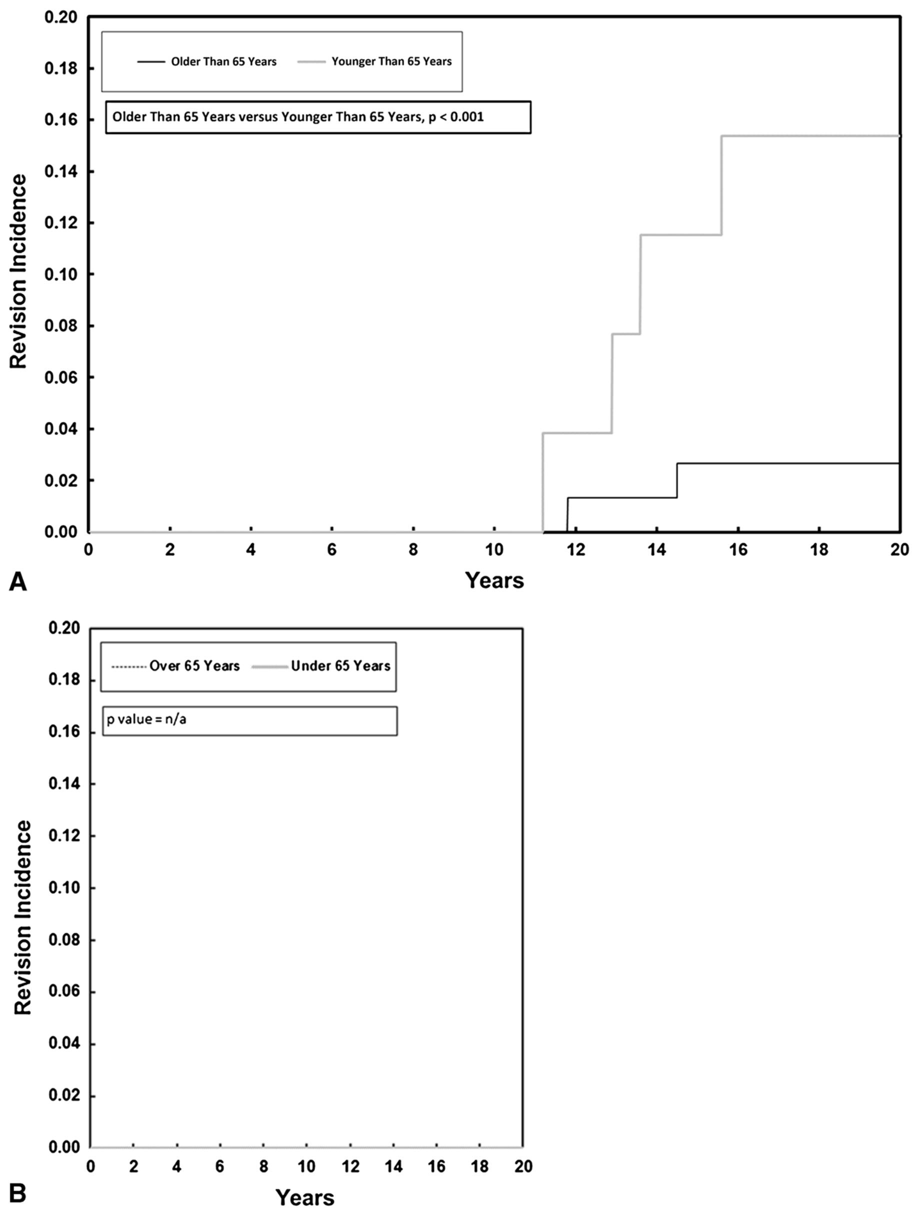

Fig. 2A–B Competing risk analysis of implant survival over time was analyzed by patients > 65 years and < 65 years for the modular tray cohort (PFC) (**A**) and the rotating platform cohort (LCS) (**B**).

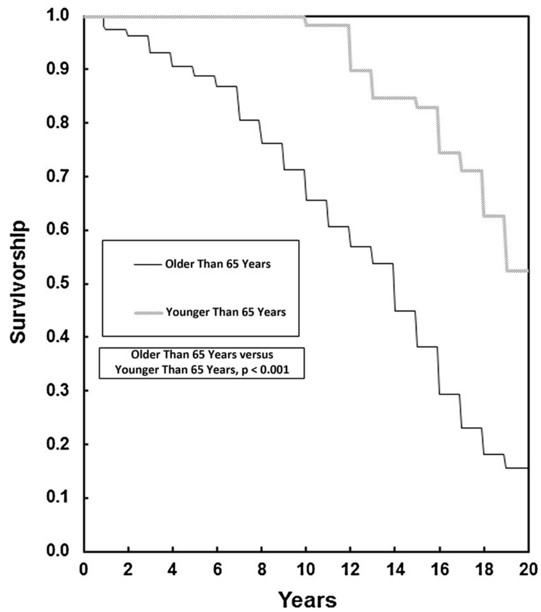

Fig. 3 Patient survivorship over the 20-year followup interval combined all patients from both cohorts.

Patient Survivorship

Most of the patients in both cohorts had died over the 20-year span of this study. Average patient age at surgery was relatively old (average, 70 years) and combined 20-year patient survival across the two cohorts was only 26%. However, survivorship was much higher in the younger patients. Twenty-year patient survivorship for patients > 65 years of age was 16% (95% CI, 10%–22%) and for patients < 65 years of age was 53% (95% CI, 40%–65%) ($p < 0.0001$) (Fig. 3).

For the PFC knee cohort, overall patient survivorship at minimum 20-year followup was 25% (76 deaths) for all patients. However, for patients < 65 years of age, survivorship was 54% (95% CI, 35%–73%) (12 deaths), and for patients > 65 years of age, survivorship was 15% (95% CI, 8%–24%) (64 deaths) ($p < 0.0001$) (Fig. 4A).

For the LCS knee cohort, patient survivorship at minimum 20-year followup was 26% (87 deaths) for all patients. However, for patients < 65 years of age, survivorship was 52% (95% CI, 35%–68%) (16 deaths), and for patients > 65 years, the survivorship was 16% (95% CI, 9%–25%) (71 deaths) ($p < 0.001$) (Fig. 4B).

Discussion

Long-term followup studies of implant designs for knee arthroplasty represent the best available evidence for investigating implant durability and patient postsurgical outcomes. Multiple studies have reported excellent midterm survivorship of implants in patients > 65 years of age [1, 4, 9, 10, 13, 16–20]. However, there are few reports of cohorts studied for a minimum of 20 years or until patient death, and the available studies have been criticized for having high rates of patient attrition from patient deaths. The present authors have performed a number of these long-term followup studies and hoped to provide some insight into the benefit of continuing to devote resources into this time-consuming endeavor. With this background, the authors evaluated two cohorts of patients undergoing knee arthroplasty, which were each longitudinally followed for a minimum of 20 years to answer the following questions: (1) Given the bias present in a study with high patient attrition, can relevant comparisons of implant durability be made across the two cohorts? (2) How does patient age affect the rates of attrition over the long followup interval? Overall, we found that both implant types were durable in the elderly cohorts studied, but that the analysis was limited by high rates of patient attrition, and no firm statistical comparisons could be made across the two cohorts. Revision rates were higher in the younger cohort of PFC knees, and patient survivorship was much higher for patients < 65 years in both cohorts. Thus, the enrollment of younger patients would likely have allowed for a more reliable comparison, and the data presented here may provide some insight into how to design future long-term followup studies.

Our study has several limitations. First, our study is a nonrandomized, retrospective review of prospectively collected data. This study design introduces the possibility of selection bias, because the patients were not randomized across implant designs. Furthermore, there is the possibility of assessor bias because the assessors were not blinded to the implant type. However, all patients were evaluated both clinically (DDG) and radiographically (JJC) by surgeons who were not involved in the initial care of the patients including the surgical procedure, which we feel helps to minimize this risk. Second, it is possible that we were underpowered to detect differences in patient or implant survivorship, introducing the possibility of a Type II error. Third, the patients operated on 20 to 30 years ago probably are of different demographics and may have different life expectancy than those being operated on today. Finally, revision is not an ideal measure of implant performance, because patients may be dissatisfied or have a poor functional outcome without requesting or undergoing revision surgery.

In answering our first question of durability, in this older cohort of patients, overall implant survivorship was excellent across both implant types with regard to revision for aseptic causes with only 6% and 0% of the PFC and LCS cohorts undergoing revision, respectively. We used a CI analysis, in which patients who died were not censored

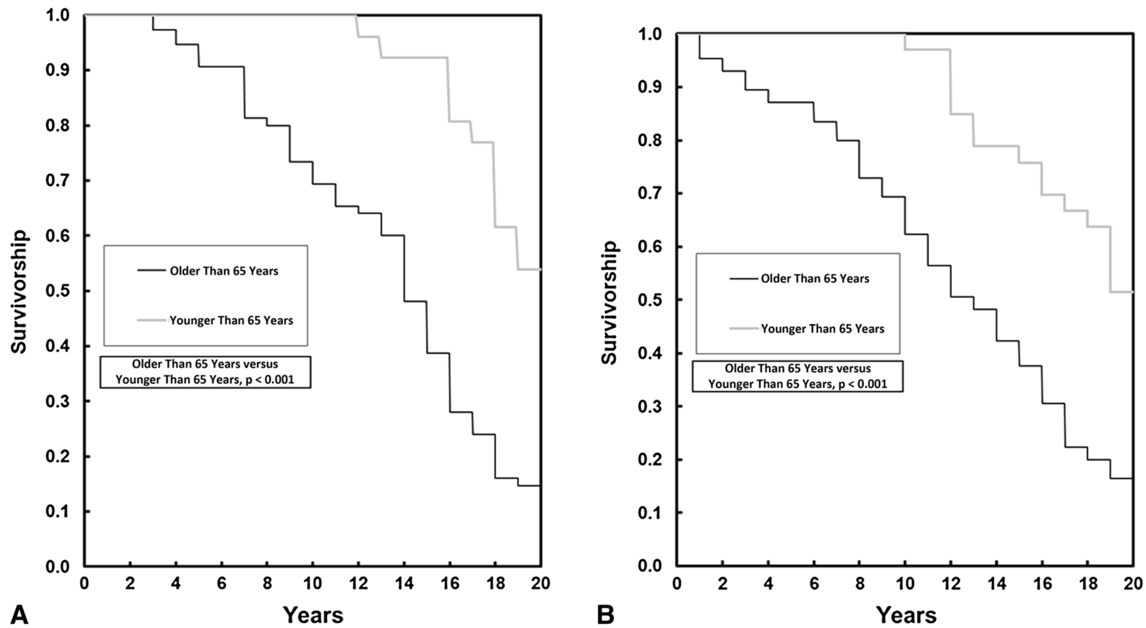

Fig. 4A–B Patient survivorship over the 20-year followup interval was separated by implant type for the modular tray cohort (PFC) (**A**) and the rotating platform cohort (LCS) (**B**).

from the result and which has greater statistical validity than a KM analysis in the presence of competing risks [7]. These data support the claim that knee arthroplasty is a durable operation, especially in the elderly, because most patients died with their original implant. However, no firm statistical comparisons could be made across the two cohorts, and attrition from patient death was high in both cohorts. Implant revision was clearly lower in the older cohort of patients older than 65 years of age with the PFC knee as compared with patients $<$ 65 years (15% versus 3%, $p = 0.0188$). Thus, as the average age of a patient undergoing knee arthroplasty continues to decrease [5], the results from these mostly elderly cohorts may not be relevant to modern patient populations.

In regard to the influence of patient age on attrition rates in long-term followup studies, the vast majority of each cohort was dead at final followup with only 25% and 26% of knees surviving to 20 years (modular tray and rotating platform, respectively). However, survivorship was much higher in the younger patients. Twenty-year patient survivorship for patients $>$ 65 years of age was 16% (95% CI, 10%–22%) and for patients $<$ 65 years of age was 53% (95% CI, 40%–65%) ($p < 0.0001$). This finding raises two important points. First, it emphasizes the importance of accounting for patient death in long-term followup studies. Most prior authors, including our group, have used KM analyses to report implant survivorship. However, the high rate of patient death we identified here clearly violates a key assumption of the KM analysis, ie, the assumption that the event of interest occurs independently from other confounding events. A patient death would preclude them from having revision surgery. Thus, a KM analysis is the wrong tool for the job. Recognizing this limitation, we chose to use a CI analysis for the comparison of implant revision rates across ages and implant designs. In a CI analysis, the patients who died are not censored, thus more directly answering the question, "What is the risk of the event?" [11]. Second, patient survival more than doubled for those $<$ 65 years of age ($p < 0.0001$) compared with those $>$ 65 years of age. Our results indicated that the enrollment of elderly patients in these prior studies is not sufficient for an accurate long-term assessment of implant durability both from the standpoint of determining 20-year durability of the implant as well as determining the ability of the implant to provide reasonable functional activity over the entire interval of followup. Thus, future long-term followup studies would likely benefit from enrolling younger patients, because this group is clearly much more likely to survive to final followup.

In summary, our results support the claim that knee arthroplasty is a durable operation in older patients. However, patient survivorship by the end of study period was very low, which raises two important points. First, because patient death is a competing risk against revision, the widespread use of a KM analysis, both by ourselves and others, is inappropriate as a tool for reporting revision rates. We recommend that investigators use patient survivorship curves that consider carefully all competing risks when planning and reporting on long-term implant results. Second, both the incidence of revisions as well as the survivorship of the patients was much different in the younger cohort of patients < 65 years of age. In view of the low likelihood that older patients will require revision surgery at any time in their remaining years, we suggest that clinicians focus their efforts at ensuring regular followup among their younger patients. For future investigators interested in long-term followup studies, the patient survivorship curves we provided may be useful for determining the necessary composition of patients in terms of patient age and numbers of patients needed to have adequate numbers for statistically valid comparisons. This may require multicenter studies of young patients to enroll the robust numbers needed to perform the most clinically relevant long-term followup studies.

Acknowledgments We thank the following for their review of radiographs at each of the previous followup intervals: Gary Fetzer MD, David Hennessy MD, Michael Iossi MD, Andrew Malin MD, and Matthew Squire MD.

References

1. Callaghan JJ, Beckert MW, Hennessy DW, Goetz DD, Kelley SS. Durability of a cruciate-retaining TKA with modular tibial trays at 20 years. *Clin Orthop Relat Res.* 2013;471:109–117.
2. Callaghan JJ, O'Rourke MR, Iossi MF, Liu SS, Goetz DD, Vittetoe DA, Sullivan PM, Jonston RC. Cemented rotating-platform total knee replacement. a concise follow-up, at a minimum of fifteen years, of a previous report. *J Bone Joint Surg Am.* 2005;87:1995–1998.
3. Callaghan JJ, Squire MW, Goetz DD, Sullivan PM, Johnston RC. Cemented rotating-platform total knee replacement. A nine to twelve-year follow-up study. *J Bone Joint Surg Am.* 2000;82: 705–711.
4. Callaghan JJ, Wells CW, Liu SS, Goetz DD, Johnston RC. Cemented rotating-platform total knee replacement: a concise follow-up, at a minimum of twenty years, of a previous report. *J Bone Joint Surg Am.* 2010;92:1635–1639.
5. Cram P, Lu X, Kaboli PJ, Vaughan-Sarrazin MS, Cai X, Wolf BR, Li Y. Clinical characteristics and outcomes of Medicare patients undergoing total hip arthroplasty, 1991–2008. *JAMA.* 2011;305:1560–1567.
6. Ewald FC. The Knee Society total knee arthroplasty roentgenographic evaluation and scoring system. *Clin Orthop Relat Res.* 1989;248:9–12.
7. Fennema P, Lubsen J. Survival analysis in total joint replacement: an alternative method of accounting for the presence of competing risk. *J Bone Joint Surg Br.* 2010;92:701–706.
8. Fetzer GB, Callaghan JJ, Templeton JE, Goetz DD, Sullivan PM, Kelley SS. Posterior cruciate-retaining modular total knee arthroplasty: a 9- to 12-year follow-up investigation. *J Arthroplasty.* 2002;17:961–966.
9. Gill GS, Joshi AB. Long-term results of kinematic condylar knee replacement. An analysis of 404 knees. *J Bone Joint Surg Br.* 2001;83:355–358.
10. Gill GS, Joshi AB, Mills DM. Total condylar knee arthroplasty. 16- to 21-year results. *Clin Orthop Relat Res.* 1999;367:210–215.
11. Grunkemeier GL, Wu Y. Interpretation of nonfatal events after cardiac surgery: actual versus actuarial reporting. *J Thorac Cardiovasc Surg.* 2001;122:216–219.
12. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc.* 1958;53:457–481.
13. Ma HM, Lu YC, Ho FY, Huang CH. Long-term results of total condylar knee arthroplasty. *J Arthroplasty.* 2005;20:580–584.
14. Malin AS, Callaghan JJ, Bozic KJ, Liu SS, Goetz DD, Sullivan N, Kelley SS. Routine surveillance of modular PFC TKA shows incerasing failures after 10 years. *Clin Orthop Relat Res.* 2010;468:2469–2476.
15. O'Rourke MR, Callaghan JJ, Goetz DD, Sullivan PM, Johnston RC. Osteolysis associated with a cemented modular posterior-cruciate-substituting total knee design: five to eight-year follow-up. *J Bone Joint Surg Am.* 2002;84:1362–1371.
16. Pavone V, Boettner F, Fickert S, Sculco TP. Total condylar knee arthroplasty: a long-term followup. *Clin Orthop Relat Res.* 2001,388:18–25.
17. Rodricks DJ, Patil S, Pulido P, Colwell CW Jr. Press-fit condylar design total knee arthroplasty. Fourteen to seventeen-year follow-up. *J Bone Joint Surg Am.* 2007;89:89–95.
18. Rodriguez JA, Bhende H, Ranawat CS. Total condylar knee replacement: a 20-year followup study. *Clin Orthop Relat Res.* 2001;388:10–17.
19. Sextro GS, Berry DJ, Rand JA. Total knee arthroplasty using cruciate-retaining kinematic condylar prosthesis. *Clin Orthop Relat Res.* 2001;388:33–40.
20. Weir DJ, Moran CG, Pinder IM. Kinematic condylar total knee arthroplasty. 14-year survivorship analysis of 208 consecutive cases. *J Bone Joint Surg Br.* 1996;78:907–911.