ORIGINAL ARTICLE

A Randomized Trial of Vertebroplasty for Osteoporotic Spinal Fractures

David F. Kallmes, M.D., Bryan A. Comstock, M.S., Patrick J. Heagerty, Ph.D., Judith A. Turner, Ph.D., David J. Wilson, F.R.C.R., Terry H. Diamond, F.R.A.C.P., Richard Edwards, F.R.C.R., Leigh A. Gray, M.S., Lydia Stout, B.S., Sara Owen, M.Sc., William Hollingworth, Ph.D., Basavaraj Ghdoke, M.D., Deborah J. Annesley-Williams, F.R.C.R., Stuart H. Ralston, F.R.C.P., and Jeffrey G. Jarvik, M.D., M.P.H.

ABSTRACT

BACKGROUND

Vertebroplasty is commonly used to treat painful, osteoporotic vertebral compression fractures.

METHODS

In this multicenter trial, we randomly assigned 131 patients who had one to three painful osteoporotic vertebral compression fractures to undergo either vertebroplasty or a simulated procedure without cement (control group). The primary outcomes were scores on the modified Roland–Morris Disability Questionnaire (RDQ) (on a scale of 0 to 23, with higher scores indicating greater disability) and patients' ratings of average pain intensity during the preceding 24 hours at 1 month (on a scale of 0 to 10, with higher scores indicating more severe pain). Patients were allowed to cross over to the other study group after 1 month.

RESULTS

All patients underwent the assigned intervention (68 vertebroplasties and 63 simulated procedures). The baseline characteristics were similar in the two groups. At 1 month, there was no significant difference between the vertebroplasty group and the control group in either the RDQ score (difference, 0.7; 95% confidence interval [CI], –1.3 to 2.8; P=0.49) or the pain rating (difference, 0.7; 95% CI, –0.3 to 1.7; P=0.19). Both groups had immediate improvement in disability and pain scores after the intervention. Although the two groups did not differ significantly on any secondary outcome measure at 1 month, there was a trend toward a higher rate of clinically meaningful improvement in pain (a 30% decrease from baseline) in the vertebroplasty group (64% vs. 48%, P=0.06). At 3 months, there was a higher crossover rate in the control group than in the vertebroplasty group (51% vs. 13%, P<0.001). There was one serious adverse event in each group.

CONCLUSIONS

Improvements in pain and pain-related disability associated with osteoporotic compression fractures in patients treated with vertebroplasty were similar to the improvements in a control group. (ClinicalTrials.gov number, NCT00068822.)

From the Mayo Clinic, Rochester, MN (D.F.K., L.A.G.); the University of Washington, Seattle (B.A.C., P.J.H., J.A.T., L.S., B.G., J.G.J.); Nuffield Orthopaedic Centre NHS Trust, Oxford (D.J.W., S.O.), Gartnavel General Hospital, Glasgow (R.E.), the University of Bristol, Bristol (W.H.), Nottingham University Hospital NHS Trust, Nottingham (D.J.A.-W.), and Western General Hospital, University of Edinburgh, Edinburgh (S.H.R.) — all in the United Kingdom; and St. George Hospital, University of New South Wales, Sydney (T.H.D.). Address reprint requests to Dr. Kallmes at the Department of Radiology, Mayo Clinic, 200 First St. SW, Rochester, MN 55905, or at kallmes.david@mayo.edu.

This article (10.1056/NEJMoa0900563) was updated on March 8, 2012, at NEJM.org.

N Engl J Med 2009;361:569-79.

Spontaneous vertebral fractures are associated with pain, disability, and death in patients with osteoporosis. Percutaneous vertebroplasty, the injection of medical cement, or polymethylmethacrylate (PMMA), into the fractured vertebral body has gained widespread acceptance as an effective method of pain relief and has become routine therapy for osteoporotic vertebral fractures. Guidelines recommend vertebroplasty for fractures that have not responded to medical treatment.[1] Typically, the duration of such fractures ranges from several weeks to several months or longer for fractures that have not healed.

Numerous case series and several small, unblinded, nonrandomized, controlled studies have suggested the effectiveness of vertebroplasty in relieving pain from osteoporotic fractures.[2-12] The precise mechanism of action remains unknown. However, in the absence of blinded, randomized, controlled trials, the role of active treatment effects of PMMA versus nonspecific effects remains unknown.

In this randomized, controlled trial, called the Investigational Vertebroplasty Safety and Efficacy Trial (INVEST), we evaluated the efficacy of PMMA infusion in vertebroplasty for patients with painful osteoporotic compression fractures, as compared with a simulated procedure without PMMA. We hypothesized that patients who had undergone vertebroplasty would report less pain and back pain–related disability at 1 month (the primary outcomes) than those in the control group.

METHODS

PATIENTS

We enrolled patients at five centers in the United States, five centers in the United Kingdom, and one center in Australia. The sites were selected on the basis of having an established vertebroplasty practice for osteoporotic fractures, an enthusiastic local principal investigator, and an available research coordinator. The study methods have been described previously.[13] Because initial recruitment was slow, after the first three patients were enrolled, we liberalized the inclusion criteria to an age of 50 years or older, a diagnosis of one to three painful osteoporotic vertebral compression fractures between vertebral levels T4 and L5, inadequate pain relief with standard medical therapy, and a current rating for pain intensity of at least 3 on a scale from 0 to 10. Fractures needed to be less than 1 year old, as indicated by the duration of pain. We previously had found that a fracture duration of up to 1 year was associated with a good response to vertebroplasty.[14] For fractures of uncertain age, an additional requirement was marrow edema on magnetic resonance imaging or increased vertebral-body uptake on bone scanning.

Exclusion criteria were evidence or suspicion of neoplasm in the target vertebral body, substantial retropulsion of bony fragments, concomitant hip fracture, active infection, uncorrectable bleeding diatheses, surgery within the previous 60 days, lack of access to a telephone, inability to communicate in English, and dementia.

The protocol was approved by the institutional review board at each study center. All patients provided written informed consent.

MEASURES

At baseline, patients completed the self-report version of the Charlson comorbidity index[15] and provided demographic and clinical information. Evaluation measures were performed before randomization and at various times up to 1 year. The focus of this report is the primary outcomes at 1 month. We also describe outcomes at 3, 14, and 90 days. The prespecified primary outcome measures were scores on the modified Roland–Morris Disability Questionnaire (RDQ) and patients' ratings of average back-pain intensity during the preceding 24 hours (on a scale of 0 to 10, with higher scores indicating more severe pain). The RDQ is widely used to assess physical disability associated with back pain and has been shown to be valid, reliable, and responsive to change in several studies,[16-21] including a study of vertebroplasty.[22] The modified RDQ[23] is scored on a scale of 0 to 23, with higher scores indicating greater physical disability. We present the (post-specified) proportion of patients who had a decrease of 30% or more on the RDQ and measures of pain intensity, which was the minimal change on each scale that was considered to be clinically important.[24-26]

Prespecified secondary outcomes included scores on the Pain Frequency Index and the Pain Bothersomeness Index,[23] the Study of Osteoporotic Fractures–Activities of Daily Living (SOF–ADL) scale,[27] and the European Quality of Life–5 Dimensions (EQ–5D) scale[28] (a generic health-status measure, reflecting mobility, self-care, activity limitations, pain, and psychological distress); the use of opioid medications; and scores on the Physical

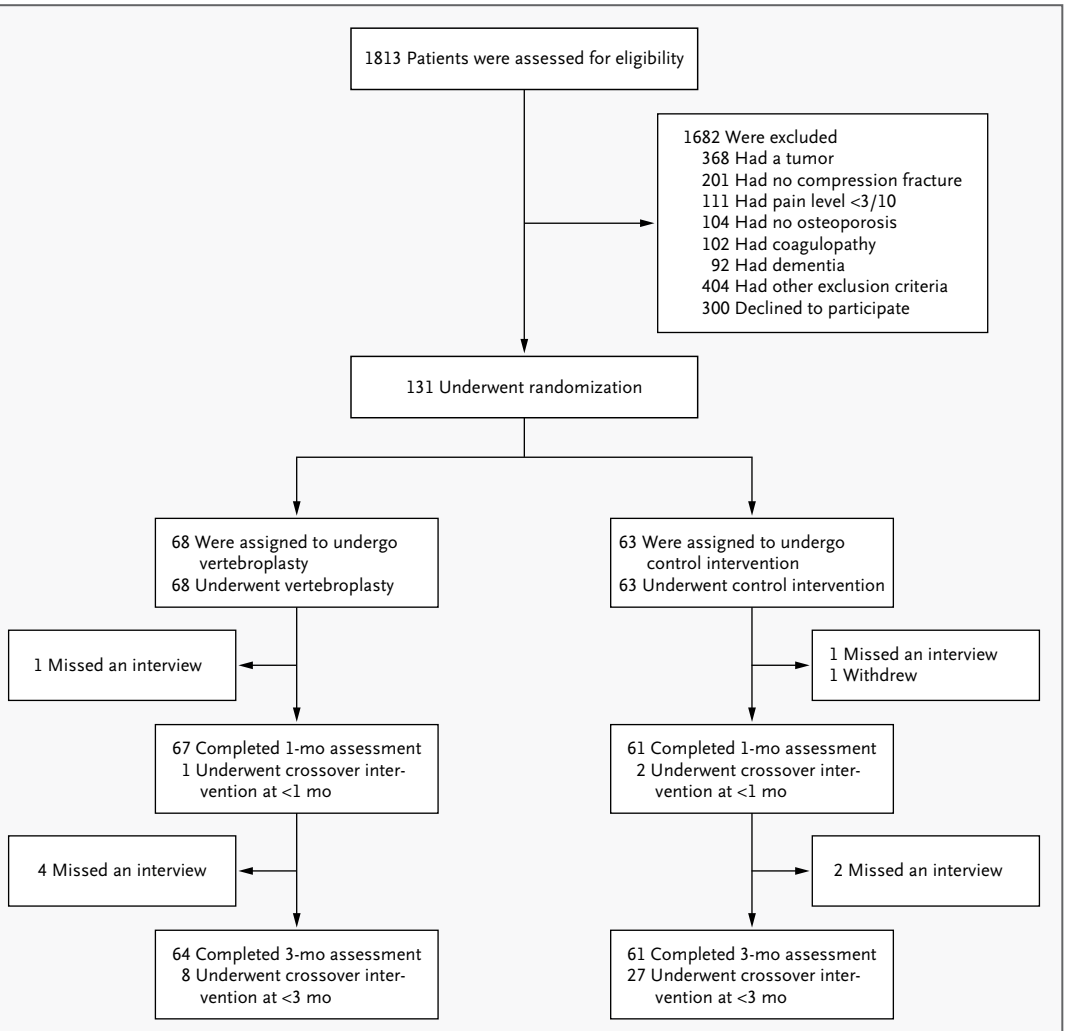

Figure 1. Enrollment and Outcomes.

The single patient in the vertebroplasty group who missed the 1-month assessment completed the 3-month assessment. The two patients in the control group who missed the 1-month assessment also missed the 3-month assessment.

Component Summary (PCS) and Mental Component Summary (MCS) subscales of the self-administered Medical Outcomes Study 36-Item Short-Form General Health Survey (SF-36), version 2.[29] The PCS assesses limitations in self-care and physical, social, and role activities; bodily pain; and perceived health. The MCS provides an indication of psychological distress and social and role disability because of emotional problems. Patients were asked before discharge on the day of the procedure and at each follow-up assessment to guess which procedure they had undergone and to rate their confidence in their guess on a scale from 0 (no confidence) to 10 (complete confidence).

STUDY TREATMENT

All vertebroplasty practitioners in the trial were highly experienced, having performed a mean of approximately 250 procedures (range, 50 to 800). Patients were brought to the fluoroscopy suite, where conscious sedation was induced and sterile preparation for surgery was performed. Using fluoroscopic guidance, the practitioner infiltrated the skin and subcutaneous tissues overlying the pedicle of the target vertebra or vertebrae with 1% lidocaine and infiltrated the periosteum of the pedicles with 0.25% bupivacaine. Patients were then randomly assigned to undergo either the full vertebroplasty procedure or the control intervention.

For the vertebroplasty procedure, 11-gauge or 13-gauge needles were passed into the central aspect of the target vertebra or vertebrae. Barium-opacified PMMA was prepared on the bench and infused under constant lateral fluoroscopy into the vertebral body. Infusion was stopped when the

Table 1. Baseline Characteristics of the Patients.*

Characteristic	Vertebroplasty Group (N=68)	Control Group (N=63)
Study center — no. (%)		
United States		
Mayo Clinic	14 (21)	16 (25)
Other than Mayo Clinic	15 (22)	12 (19)
United Kingdom	26 (38)	26 (41)
Australia	13 (19)	9 (14)
Age — yr	73.4±9.4	74.3±9.6
White race — no. (%)†	67 (99)	60 (95)
Female sex — no. (%)	53 (78)	46 (73)
Education — no. (%)		
Less than high school	36 (53)	26 (41)
High school	13 (19)	23 (37)
Some college	10 (15)	7 (11)
College graduate	9 (13)	7 (11)
Married or living with partner — no. (%)	42 (62)	27 (43)
Employment status — no. (%)		
Employed full- or part-time	7 (10)	5 (8)
Retired	42 (62)	39 (62)
Disabled	10 (15)	10 (16)
Other	9 (13)	9 (14)
Current smoker — no. (%)	12 (18)	9 (14)
Comorbidity index‡	1.9±2.1	2.0±1.9
Receiving worker's compensation — no. (%)	9 (13)	7 (11)
Pain duration		
Mean — wk	16	20
Interquartile range — wk	10–36	8–38
Period — no. (%)		
1–13 wk	30 (44)	24 (38)
14–26 wk	14 (21)	15 (24)
27–39 wk	8 (12)	9 (14)
40–52 wk	16 (24)	15 (24)
No. of spinal levels treated — no. (%)		
1	48 (71)	41 (65)
2	13 (19)	14 (22)
3	7 (10)	8 (13)
Self-reported use of opioid analgesic — no. (%)	38 (56)	40 (63)
RDQ score §	16.6±3.8	17.5±4.1
Average pain intensity during past 24 hr¶	6.9±2.0	7.2±1.8

Table 1. (Continued.)

Characteristic	Vertebroplasty Group (N=68)	Control Group (N=63)
SF-36 score‖		
Physical component	25.3±7.8	25.3±7.3
Mental component	44.8±11.8	41.5±14.1
Pain Frequency Index score**	3.0±0.8	3.1±0.8
Pain Bothersomeness Index score**	2.9±0.7	3.1±0.8
EQ–5D score††	0.57±0.18	0.54±0.23
SOF–ADL score‡‡	10.0±3.6	10.3±2.8

* Plus–minus values are means ±SD.
† Race was self-reported.
‡ Scores on the comorbidity index range from 0 to 28, with higher scores indicating greater severity.[15]
§ Scores on the Roland–Morris Disability Questionnaire (RDQ) range from 0 to 23, with higher scores indicating more severe disability.
¶ The pain-intensity rating ranges from 0 (no pain) to 10 (worst pain).
‖ Scores on the Medical Outcomes Study 36-Item Short-Form General Health Survey (SF-36), version 2, range from 0 to 100, with lower scores indicating a worse outcome.
** Scores on the Pain Frequency Index and Pain Bothersomeness Index range from 0 to 4, with higher scores indicating more severe pain.
†† Scores on the European Quality of Life–5 Dimensions (EQ–5D) scale range from –0.1 to 1.0, with higher scores indicating a better quality of life.
‡‡ Scores on the Study of Osteoporotic Fractures–Activities of Daily Living (SOF–ADL) scale range from 0 to 18, with higher scores indicating more back-related disability.

PMMA reached to the posterior aspect of the vertebral body or entered an extraosseous space, such as the intervertebral disk or an epidural or paravertebral vein.[30] During the control intervention, verbal and physical cues, such as pressure on the patient's back, were given, and the methacrylate monomer was opened to simulate the odor associated with mixing of PMMA, but the needle was not placed and PMMA was not infused. After the procedure, both groups of patients were monitored in the supine position for 1 to 2 hours before discharge.

Patients were told at the time of consent that they would be allowed to cross over to the other procedure 1 month or later after the intervention if adequate pain relief was not achieved. Specific numerical thresholds of outcome measures were not used for allowance of crossover. Patients were seen in the clinic for the 1-month follow-up visit by a vertebroplasty practitioner to discuss whether to cross over to receive the alternative therapy.

No commercial entity paid for any materials used in the study. Research funds paid for all costs related to the control interventions. Costs of the vertebroplasty procedure were billed to insurance.

RANDOMIZATION AND BLINDING

We used stratified, blocked randomization according to study center to achieve roughly balanced groups. The block sizes ranged from 4 to 12 patients, and assignments were concealed from the research assistants involved in recruitment. These assignments were generated by the data coordinating center with the use of a random-number generator and were then placed in numbered, opaque, sealed envelopes, with a series of envelopes for each study center. The protocol specified that study-group assignments should be concealed from all patients and study personnel who performed follow-up assessments for the duration of the study. Only the study statisticians, who did not have any contact with the patients, saw unblinded data.

STATISTICAL ANALYSIS

The study initially had a power of more than 80% to detect differences in both primary and secondary outcomes in 250 patients, with a two-sided alpha of 0.05, on the basis of a 2.5-point difference on the RDQ and a 1.0-point difference on the pain rating. After early difficulty in recruitment and a planned interim analysis of the first 90 patients, we reduced the target sample size to 130 patients, with approval from the independent data and safety monitoring board. The decision to modify the target enrollment was driven primarily by accrual rates and revised power calculations. With the reduced sample size, the study had a power of more

than 80% to detect important differences in the primary outcome measures — a 3.0-point difference between groups on the RDQ (with an assumed SD of 6.7) and a 1.5-point difference on the pain rating (with an assumed SD of 2.7) — at 1 month.[26]

For our primary analyses, we used an intention-to-treat strategy, with patients analyzed in their assigned group. Treatment effects and confidence intervals were calculated from analysis of covariance (ANCOVA) models with adjustment for baseline values of the outcome measure, recruitment site, and an indicator of study group as the predictor of interest. In a post hoc analysis, we used logistic-regression models with adjustment for site and baseline values of the outcome measures to compare the proportion of patients in each group who had at least a 30% improvement in the RDQ score and pain rating, as recommended by the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials II to assess the clinical importance of improvement.[31] Furthermore, we performed two post hoc subgroup analyses to determine whether continuous and categorical measures of the duration of baseline-intensity pain (as an index of fracture age) interacted with treatment in predicting the pain intensity at 1 month in the ANCOVA models; these measures compared the periods from 1 to 13 weeks, from 14 to 26 weeks, and from 27 to 52 weeks. Formal evaluation of effect modification was based on a partial F-test of whether the two interaction terms equaled zero. The results were similar for the two analyses, and we report the results for the categorical measures.

The data and safety monitoring board reviewed the blinded study results every 6 months to evaluate safety and efficacy and monitored any deaths, events involving paralysis, hospitalizations, new-onset fractures, new radiculopathy or myelopathy, and infection. The board used O'Brien–Fleming[32] stopping rules of $P<0.001$ and $P<0.019$ for two prespecified interim analyses in order to evaluate the accumulating evidence of treatment efficacy; the interim study results did not reach either threshold. All statistical analyses were performed with the use of R statistical software, version 2.7.[33] A P value of less than 0.043 for between-group differences in the primary outcomes was considered to indicate statistical significance. All reported P values are two-sided and have not been adjusted for multiple testing.

RESULTS

PATIENTS

From June 2004 through August 2008, a total of 131 patients were enrolled and underwent randomization (Fig. 1). Of these patients, 68 were assigned to undergo vertebroplasty and 63 to undergo the control intervention; all underwent the assigned intervention. The baseline characteristics of the groups were similar (Table 1). One patient (1%) in the vertebroplasty group and two patients (3%) in the control group were lost to follow-up before 1 month. One patient (1%) in the vertebroplasty group and two patients (3%) in the control group crossed over to the other group before 1 month.

The two study groups did not differ significantly with respect to either of the two prespecified primary outcomes at 1 month. The mean (±SD) RDQ score in the vertebroplasty group was 12.0±6.3, as compared with 13.0±6.4 in the control group (adjusted treatment effect, 0.7; 95% confidence interval [CI], −1.3 to 2.8; $P=0.49$). The mean pain-intensity rating was 3.9±2.9 in the vertebroplasty group and 4.6±3.0 in the control group (adjusted treatment effect, 0.7; 95% CI, −0.3 to 1.7; $P=0.19$) (Table 2). The two study groups had substantial improvement in back-related disability and pain immediately (3 days) after the procedure, with similar improvement in the two groups. The improvement in each group at 3 days was maintained at 1 month.

The study groups did not differ significantly on any of the secondary outcomes, including measures of pain and quality of life, at 1 month (Fig. 2). Furthermore, the two groups did not differ in the post-specified proportion of patients who had clinically meaningful improvement in physical disability related to back pain at 1 month (40% of patients in the vertebroplasty group and 41% of patients in the control group, $P=0.99$). There was a trend toward a higher rate of clinically meaningful improvement in pain in the vertebroplasty group than in the control group (64% vs. 48%, $P=0.06$).

At 3 months, 9 patients (13%) in the vertebroplasty group and 32 patients (51%) in the control group had crossed over to the other group and had undergone the alternative procedure ($P<0.001$). The patients in the vertebroplasty group who crossed over reported higher levels of disability and pain at 3 days and 14 days, as compared with patients who did not cross over (Fig. 3). Patients in the

control group who crossed over had some early improvement after the control procedure, but this improvement had disappeared by the 1-month assessment. However, even after they underwent the alternative intervention, patients who were originally assigned to either the vertebroplasty group or the control group did not have the same level of improvement at 3 months as did patients who did not cross over.

At 14 days, 63% of patients in the control group correctly guessed that they had undergone the control intervention, and 51% of patients in the vertebroplasty group correctly guessed that they had undergone vertebroplasty. Patients in both the vertebroplasty group and the control group expressed a moderate degree of confidence, on a scale of 0 (not certain) to 10 (extremely certain), in their treatment guess (mean scores, 4.0 and 4.1, respectively; P=0.78). In the control group, 18 of 33 patients (55%) who did not cross over to vertebroplasty correctly guessed at 14 days that they had undergone the control intervention, as compared with 20 of 27 patients (74%) who eventually crossed over (P=0.12). Notably, among the eight patients in the vertebroplasty group who crossed over to the control group, six (75%) guessed incorrectly at 1 month that they had received the control intervention.

In a post hoc subgroup analysis, the effect of treatment (vertebroplasty vs. control procedure) on pain at 1 month did not differ significantly across the three baseline pain-duration categories (P=0.58). The treatment effect for patients with less than 13 weeks of pain (difference in score, 0.8; 95% CI, −0.8 to 2.4; P=0.31) was similar to the results for the overall analysis. The treatment effect for patients with 14 to 26 weeks of pain was 1.3 (95% CI, −0.8 to 3.4; P=0.23), and the effect for patients with 27 to 52 weeks of pain was 0.0 (95% CI, −1.7 to 1.6; P=0.96).

Table 2. Primary Outcomes (Intention-to-Treat Analyses).*

Measure	Vertebroplasty Group	Control Group	Treatment Effect (95% CI)†	P Value†
RDQ‡				
At baseline	16.6±3.8	17.5±4.1		
At 3 days	13.0±5.2	12.5±5.5	−0.9 (−2.7 to 0.8)	0.30
At 14 days	12.4±5.8	12.3±5.9	−0.6 (−2.4 to 1.2)	0.35
At 1 mo	12.0±6.3	13.0±6.4	0.7 (−1.3 to 2.8)	0.49
Pain intensity§				
At baseline	6.9±2.0	7.2±1.8		
At 3 days	4.2±2.8	3.9±2.9	−0.4 (−1.5 to 0.5)	0.37
At 14 days	4.3±2.9	4.5±2.8	0.1 (−0.8 to 1.1)	0.77
At 1 mo	3.9±2.9	4.6±3.0	0.7 (−0.3 to 1.7)	0.19

* Plus–minus values are means ±SD.
† Between-group comparisons, confidence intervals, and P values were calculated with the use of analysis-of-covariance models with adjustment for study-group assignment, baseline value of the outcome measure, and study center. Negative treatment effects favor the control procedure, and positive treatment effects favor vertebroplasty.
‡ Scores on the Roland–Morris Disability Questionnaire (RDQ) range from 0 to 23, with higher scores indicating more severe disability.
§ Scores on the pain-intensity scale range from 0 (no pain) to 10 (worst pain).

ADVERSE EVENTS

One patient in the vertebroplasty group had an injury to the thecal sac during the procedure, with resultant hospitalization. One patient in the control group was hospitalized overnight after the procedure with tachycardia and rigors of unknown cause.

DISCUSSION

Patients with osteoporotic vertebral fractures who were randomly assigned to undergo either a full vertebroplasty or a control intervention consisting of a simulated vertebroplasty without infusion of PMMA did not differ significantly at 1 month after the procedure on measures of back-pain intensity, functional disability, and quality of life. In this study, the confidence interval for the comparison of the RDQ score (−1.3 to 2.8) excluded a treatment benefit of 3 points or more and therefore provided evidence against clinically meaningful treatment effects with respect to functional disability. Similarly, the confidence interval for the comparison of pain ratings (−0.3 to 1.7) excluded a benefit of 2 points or more. Patients in the two study groups showed immediate improvement in pain and disability after the procedure, and this improvement was sustained at 1 month. These results suggest that factors aside from the instillation of PMMA may have accounted for the observed clinical improvement after vertebroplasty. Such factors may include the effect of local anesthesia, as well as nonspecific effects, such as expectations of pain relief (the so-called placebo effect), the natural history of the fracture, and regression toward the mean.

The possible role of the placebo effect on outcomes in this trial remains unclear. Previous studies have documented pain reduction in placebo groups, on the order of 6 to 7 mm on a 100-mm

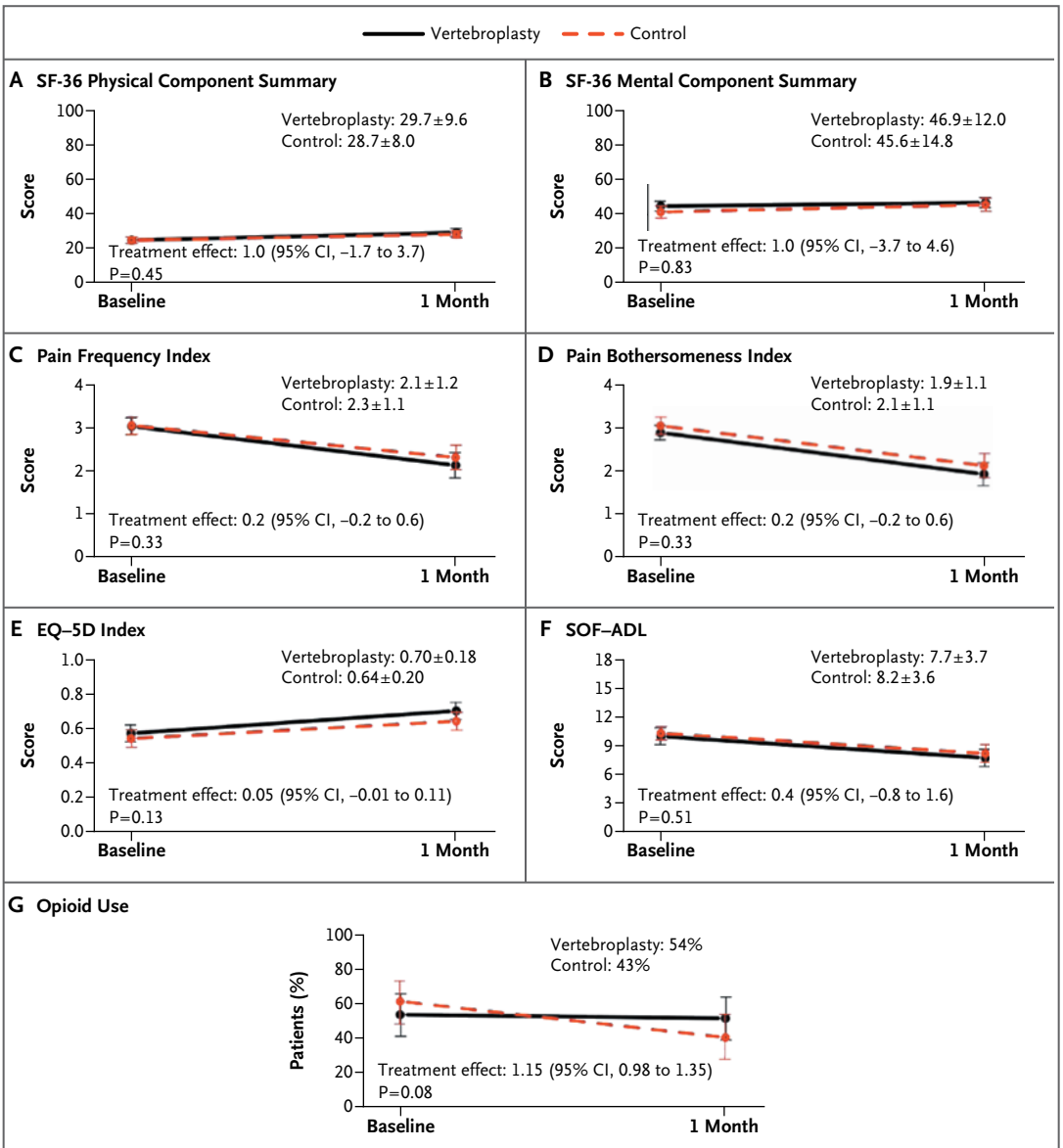

Figure 2. Secondary Outcome Measures at 1 Month (Intention-to-Treat Analyses).

Prespecified secondary outcomes included mean (±SD) scores on the Medical Outcomes Study 36-Item Short-Form General Health Survey (SF-36), version 2, including both the Physical Component Summary (Panel A) and the Mental Component Summary (Panel B), the Pain Frequency Index (Panel C), the Pain Bothersomeness Index (Panel D), the European Quality of Life–5 Dimensions (EQ–5D) scale (Panel E), and the Study of Osteoporotic Fractures–Activities of Daily Living (SOF–ADL) scale (Panel F), as well as the proportion of patients who were taking any opioid analgesics (Panel G). Scores on the SF-36 range from 0 to 100, with lower scores indicating a worse outcome. Scores on the Pain Frequency Index and the Pain Bothersomeness Index range from 0 to 4, with higher scores indicating more severe pain. Scores on the EQ-5D scale range from −0.1 to 1.0, with higher scores indicating a better quality of life. Scores on the SOF–ADL scale range from 0 to 18, with higher scores indicating more back-related disability. For continuous outcome measures, treatments were compared with the use of analysis-of-covariance models with adjustment for study-group assignment, baseline value of the outcome measure, and study center, with all positive numbers favoring vertebroplasty. The treatment effect for opioid use is reported as an odds ratio from a logistic-regression model with adjustment for baseline opioid use and study center. The I bars denote 95% confidence intervals.

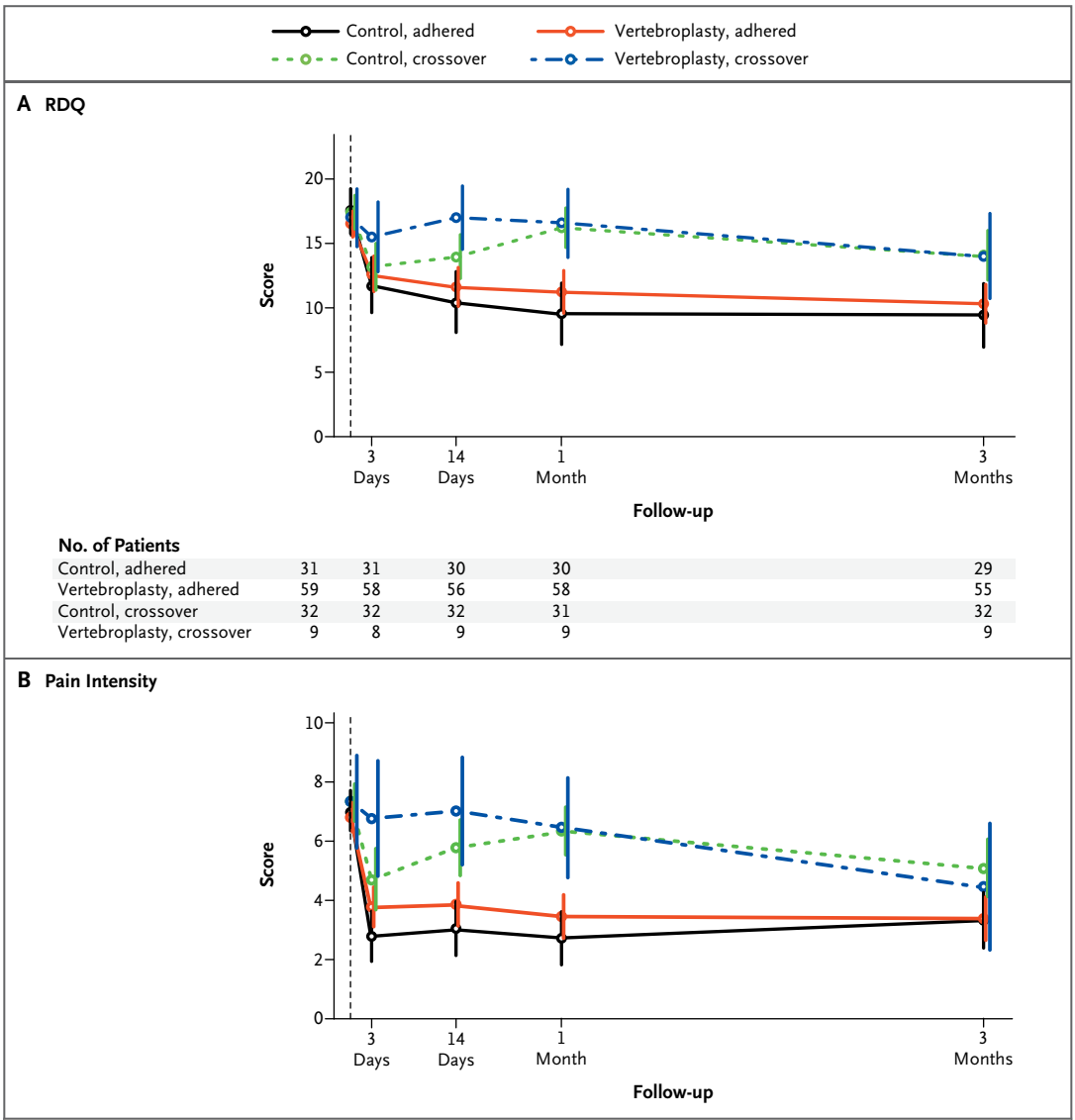

Figure 3. Scores on Measures of Disability and Pain over a 3-Month Period.

Scores on the Roland–Morris Disability Questionnaire (RDQ) range from 0 to 23, with higher scores indicating more severe disability (Panel A). Scores on the pain-intensity rating range from 0 (no pain) to 10 (worst pain) (Panel B). Patients were classified as having adhered to their random study-group assignment if they did not subsequently undergo the alternative procedure before the 3-month follow-up. Patients who underwent the alternative procedure during this period were said to have crossed over to the other study group. At 3 months, 8 patients (12%) in the vertebroplasty group and 27 patients (43%) in the control group had crossed over to the other group and had undergone the alternative procedure ($P<0.001$). The black vertical lines indicate the time when baseline measures were taken, and the colored vertical lines represent 95% confidence intervals.

scale.[34-36] The treatment effect in our trial was substantially larger than those in previous studies, even though the previous studies included both pharmacologic and psychological interventions in addition to physical interventions.[34]

There was a trend toward a higher proportion of patients in the vertebroplasty group with clinically meaningful improvement in pain at 1 month. Furthermore, there was a higher crossover rate in the control group than in the vertebroplasty group after 1 month. The reasons for the higher crossover rate are unknown. It is possible that more

patients in the control group than in the vertebroplasty group had unsatisfactory pain outcomes but that we were unable to detect this difference with our measure of pain intensity. However, we used a common, validated measure that has been shown to indicate responsiveness to clinical improvement. It is possible that vertebroplasty was more effective than the control intervention for a subgroup of patients; further research is needed to explore this possibility. Finally, it is possible that despite efforts to conceal study-group assignments, some patients became aware of their assigned intervention, and those who still had pain and learned that they were in the control group may have elected to cross over to the vertebroplasty group.

Our study had several limitations. First, we allowed crossover at 1 month because both physicians and patients were reluctant to accept a longer period. This factor complicated the interpretation of between-group differences in outcomes after 1 month. However, there is evidence that nearly all the benefits of vertebral augmentation occur within the first month.[37] In addition, since the half-life of bupivacaine is only 3 hours, any benefit from this drug would have disappeared at 1 month. Second, we did not compare the study groups with respect to other medical treatments that they received that might have affected their outcomes. Third, the persistence of pain after vertebroplasty or fracture healing may indicate causes of the pain other than fracture, a possibility that our baseline imaging excluded to a certain extent but not entirely. Fourth, even though there was no differential treatment effect according to the baseline duration of pain, a result that is consistent with our previous finding that the fracture age is not associated with the response to vertebroplasty,[14] it remains possible that vertebroplasty is effective only for fractures of a certain age or healing stage. Finally, we limited our study to vertebroplasty and did not evaluate the efficacy of kyphoplasty, which is similar to vertebroplasty except that intraosseous balloons are inflated before cement infusion.[38]

In conclusion, at 1 month, clinical improvement in patients with painful osteoporotic vertebral fractures was similar among those treated with vertebroplasty and those treated with a simulated procedure. These data suggest that further studies should be undertaken to determine whether the long-term outcome is similar in the two groups, especially because our crossover study design limited our ability to shed light on the long-term efficacy of vertebroplasty.

Supported by a grant (R01-AR49373) from the National Institute of Arthritis and Musculoskeletal and Skin Diseases.

Dr. Kallmes reports receiving consulting fees from Zelos Therapeutics and grant support from ArthroCare, Stryker, Cardinal, and Cook and serving as an unpaid consultant to Bone Support; Dr. Heagerty, receiving grant support from GlaxoSmithKline; Dr. Annesley-Williams, receiving lecture fees from Stryker; Dr. Ralston, receiving consulting fees from Procter & Gamble, Novartis, and Merck, lecture fees from Novartis and Procter & Gamble, and grant support from Novartis and Wyeth; and Dr. Jarvik, receiving consulting fees from HealthHelp and having an equity interest in PhysioSonics and Nevro. No other potential conflict of interest relevant to this article was reported.

We thank Arash Ehteshami Rad at the Mayo Clinic; Tom Marshall and Clare Darrah at the Norfolk and Norwich University Hospital, Norwich, United Kingdom; Avery Evans and Selene Boutin at the University of Virginia; Juan Tejada and Vijay Bharati at Indiana University; Andy DeNardo and Judy Jackson at Indianapolis Methodist Hospital; Jonas Goldstein at Asheville Radiology Associates, Asheville, NC; and Sophie Clare at the University of Washington.

REFERENCES

1. McGraw JK, Cardella J, Barr JD, et al. Society of Interventional Radiology quality improvement guidelines for percutaneous vertebroplasty. J Vasc Interv Radiol 2003;14:827-31.
2. Alvarez L, Alcaraz M, Perez-Higueras A, et al. Percutaneous vertebroplasty: functional improvement in patients with osteoporotic compression fractures. Spine 2006;31:1113-8.
3. Diamond TH, Bryant C, Browne L, Clark WA. Clinical outcomes after acute osteoporotic vertebral fractures: a 2-year non-randomised trial comparing percutaneous vertebroplasty with conservative therapy. Med J Aust 2006;184:113-7.
4. Diamond TH, Champion B, Clark WA. Management of acute osteoporotic vertebral fractures: a nonrandomized trial comparing percutaneous vertebroplasty with conservative therapy. Am J Med 2003; 114:257-65.
5. Do HM, Kim BS, Marcellus ML, Curtis L, Marks MP. Prospective analysis of clinical outcomes after percutaneous vertebroplasty for painful osteoporotic vertebral body fractures. AJNR Am J Neuroradiol 2005;26:1623-8.
6. Kobayashi K, Shimoyama K, Nakamura K, Murata K. Percutaneous vertebroplasty immediately relieves pain of osteoporotic vertebral compression fractures and prevents prolonged immobilization of patients. Eur Radiol 2005;15:360-7.
7. Layton KF, Thielen KR, Koch CA, et al. Vertebroplasty, first 1000 levels of a single center: evaluation of the outcomes and complications. AJNR Am J Neuroradiol 2007;28:683-9.
8. McGraw JK, Lippert JA, Minkus KD, Rami PM, Davis TM, Budzik RF. Prospective evaluation of pain relief in 100 patients undergoing percutaneous vertebroplasty: results and follow-up. J Vasc Interv Radiol 2002;13:Suppl 1:883-6.
9. Peh WC, Gilula LA, Peck DD. Percutaneous vertebroplasty for severe osteoporotic vertebral body compression fractures. Radiology 2002;223:121-6.
10. Voormolen MH, Lohle PN, Lampmann LE, et al. Prospective clinical follow-up after percutaneous vertebroplasty in patients with painful osteoporotic vertebral compression fractures. J Vasc Interv Radiol 2006;17:1313-20.
11. Voormolen MH, Mali WP, Lohle PN, et al. Percutaneous vertebroplasty compared with optimal pain medication treatment: short-term clinical outcome of patients with subacute or chronic painful osteo-

porotic vertebral compression fractures. AJNR Am J Neuroradiol 2007;28:555-60.
12. Zoarski GH, Snow P, Olan WJ, et al. Percutaneous vertebroplasty for osteoporotic compression fractures: quantitative prospective evaluation of long-term outcomes. J Vasc Interv Radiol 2002;13: Suppl 1:139-48.
13. Gray LA, Jarvik JG, Heagerty PJ, et al. INvestigational Vertebroplasty Efficacy and Safety Trial (INVEST): a randomized controlled trial of percutaneous vertebroplasty. BMC Musculoskelet Disord 2007;8: 126. (Accessed July 13, 2009, at http://www.biomedcentral.com/1471-2474/8/126.)
14. Kaufmann TJ, Jensen ME, Schweickert PA, Marx WF, Kallmes DF. Age of fracture and clinical outcomes of percutaneous vertebroplasty. AJNR Am J Neuroradiol 2001;22:1860-3.
15. Katz JN, Chang LC, Sangha O, Fossel AH, Bates DW. Can comorbidity be measured by questionnaire rather than medical record review? Med Care 1996;34:73-84.
16. Roland M, Morris R. A study of the natural history of back pain. Part I: development of a reliable and sensitive measure of disability in low-back pain. Spine 1983;8:141-4.
17. Deyo RA. Comparative validity of the Sickness Impact Profile and shorter scales for functional assessment in low-back pain. Spine 1986;11:951-4.
18. Jensen MP, Strom SE, Turner JA, Romano JM. Validity of the Sickness Impact Profile Roland scale as a measure of dysfunction in chronic pain patients. Pain 1992;50:157-62.
19. Underwood MR, Barnett AG, Vickers MR. Evaluation of two time-specific back pain outcome measures. Spine 1999;24: 1104-12.
20. Beurskens AJ, de Vet HC, Köke AJ. Responsiveness of functional status in low back pain: a comparison of different instruments. Pain 1996;65:71-6.
21. Roland M, Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. Spine 2000;25:3115-24.
22. Trout AT, Kallmes DF, Gray LA, et al. Evaluation of vertebroplasty with a validated outcome measure: the Roland-Morris Disability Questionnaire. AJNR Am J Neuroradiol 2005;26:2652-7.
23. Patrick DL, Deyo RA, Atlas SJ, Singer DE, Chapin A, Keller RB. Assessing health-related quality of life in patients with sciatica. Spine 1995;20:1899-909.
24. Dworkin RH, Turk DC, Wyrwich KW, et al. Interpreting the clinical importance of treatment outcomes in chronic pain clinical trials: IMMPACT recommendations. J Pain 2008;9:105-21.
25. Farrar JT, Young JP, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. Pain 2001;94:149-58.
26. Ostelo RW, Deyo RA, Stratford P, et al. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. Spine 2008; 33:90-4.
27. Ettinger B, Black DM, Nevitt MC, et al. Contribution of vertebral deformities to chronic back pain and disability. J Bone Miner Res 1992;7:449-56.
28. Brooks R. EuroQol: the current state of play. Health Policy 1996;37:53-72.
29. Ware JE, Kosinski M, Dewey JE. How to score version two of the SF-36 health survey. Lincoln, RI: QualityMetric, 2002.
30. Jensen ME, Evans AJ, Mathis JM, Kallmes DF, Cloft HJ, Dion JE. Percutaneous polymethylmethacrylate vertebroplasty in the treatment of osteoporotic vertebral body compression fractures: technical aspects. AJNR Am J Neuroradiol 1997;18: 1897-904.
31. Dworkin RH, Turk DC, Farrar JT, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 2005;113:9-19.
32. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. Biometrics 1979;35:549-56.
33. The R Development Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2008.
34. Hróbjartsson A, Gøtzsche PC. Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. N Engl J Med 2001;344:1594-602.
35. *Idem.* Is the placebo powerless? Update of a systematic review with 52 new randomized trials comparing placebo with no treatment. J Intern Med 2004;256: 91-100.
36. Vase L, Riley JL 3rd, Price DD. A comparison of placebo effects in clinical analgesic trials versus studies of placebo analgesia. Pain 2002;99:443-52.
37. Wardlaw D, Cummings SR, Van Meirhaeghe J, et al. Efficacy and safety of balloon kyphoplasty compared with nonsurgical care for vertebral compression fracture (FREE): a randomised controlled trial. Lancet 2009;373:1016-24.
38. Taylor RS, Fritzell P, Taylor RJ. Balloon kyphoplasty in the management of vertebral compression fractures: an updated systematic review and meta-analysis. Eur Spine J 2007;16:1085-100.