A Randomized Controlled Trial of Vertebroplasty for Osteoporotic Spine Fractures

D.F. Kallmes[1], **B.A. Comstock**[2], **P.J. Heagerty**[2], **J.A. Turner, Ph.D.**[2], **D.J. Wilson**[4], **T.H. Diamond**[5], **R. Edwards**[6], **L.A. Gray**[1], **L. Stout**[2], **S. Owen**[4], **W. Hollingworth**[3], **B. Ghdoke**[2], **D.J. Annesley-Williams**[7], **S.H. Ralston**[8], and **J.G. Jarvik**[2]

[1] Mayo Clinic, Rochester, MN/US

[2] University of Washington, Seattle, WA/US

[3] Department of Social Medicine, Bristol/UK

[4] Nuffield Orthopaedic Centre NHS Trust, Oxford/UK

[5] St George Hospital, University of New South Wales, Sydney/AU

[6] Gartnavel General Hospital, Glasgow/UK

[7] Nottingham University Hospital NHS Trust, Nottingham/UK

[8] Western General Hospital, University of Edinburgh, Edinburgh/UK

Abstract

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

Methods—In this multi-center trial, we randomly assigned patients with 1-3 painful, osteoporotic vertebral compression fractures to vertebroplasty or to a simulated vertebroplasty without cement. The primary outcomes were modified Roland-Morris Disability Questionnaire (RDQ) scores (range, 0–23) and patient ratings of average pain intensity in the preceding 24 hours (0–10 numerical rating scale) at one month. Patients were allowed to cross over after one month.

Results—All patients received their assigned interventions (68 vertebroplasty and 63 simulated vertebroplasty). The baseline characteristics were similar in the two groups. At one month, the vertebroplasty and control groups did not differ significantly on either the RDQ (treatment difference: 0.7; 95% CI: −1.3, 2.8; P = 0.49) or the pain rating (treatment difference: 0.7; 95% CI: −0.3, 1.7; P = 0.19). Both groups showed immediate improvement in disability and pain after the intervention. Although the groups did not differ significantly on any secondary outcome at one month, there was a trend toward a higher rate of clinically meaningful improvement in pain (30% decrease from baseline) in the vertebroplasty group (64% versus 48%, P = 0.06). At three months, there was a higher crossover rate in the control group (43% versus 12%, P<0.001)). There was one serious adverse event in each group.

Conclusions—Improvement in osteoporotic compression fracture pain and pain-related disability was similar in patients treated with vertebroplasty and patients treated with simulated vertebroplasty without cement.

Introduction

Spontaneous, painful vertebral fractures represent an important cause of morbidity and mortality among 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 management;[1] typically, such fracture duration ranges from several weeks to several months, or longer, for fractures that have not healed.

Numerous case series and several small, non-blinded, non-randomized controlled studies have suggested 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 (RCTs), the role of active treatment effects of PMMA versus nonspecific effects remains unknown.

We conducted an RCT, the INvestigational Vertebroplasty Safety and Efficacy Trial (INVEST), to evaluate the efficacy of PMMA infusion in vertebroplasty for patients with painful, osteoporotic compression fractures, as compared with a simulated vertebroplasty without PMMA. We hypothesized that patients assigned to vertebroplasty, as compared with patients assigned to the control procedure, would report less pain and back pain-related disability at one month (the primary endpoints).

Methods

Patients

We enrolled patients from five centers in the U.S., five centers in the United Kingdom, and one center in Australia. Sites were selected based on having an established vertebroplasty practice for osteoporotic fractures, enthusiasm of the local principal investigator, and the availability of a research coordinator. Study methods were described previously. [13] Because initial recruitment was slow, after enrolling the first 3 patients, we liberalized our inclusion criteria to: age 50 years or older; 1–3 painful, osteoporotic vertebral compression fractures between vertebral levels T4 and L5; inadequate pain relief with standard medical therapy; and current pain intensity rated at least 3 on a 0–10 scale. Fractures needed to be less than one year old, (as indicated by pain duration); we previously found that fracture duration, up to one year, was not associated with vertebroplasty response.[14] 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 past 60 days, no access to a telephone, inability to communicate in English, and dementia. We required subjects with fractures of uncertain age to have marrow edema on magnetic resonance imaging or increased vertebral body uptake on bone scan. The protocol was approved by the Institutional Review Boards at all sites and all patients gave 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 administered prior to randomization and at various times up to one year. The focus of this report is on outcomes at one month, the primary endpoint. We also describe outcomes at three, 14, and 90

days. The pre-specified primary outcome measures were the modified Roland-Morris Disability Questionnaire (RDQ) and a numerical rating scale (NRS; 0 = 'no pain' to 10 = 'pain as bad as could be') score of average back pain intensity in the preceding 24 hours. The RDQ is widely used to assess physical disability associated with back pain, and has been demonstrated to be valid, reliable, and responsive to change, [16–21] including in studies of vertebroplasty.[22] The modified RDQ [23] is scored on a 0–23 scale, with higher scores indicating greater physical disability. We present the (post-specified) proportion of patients who achieved a decrease of 30% on the RDQ and NRS measures of pain intensity, the minimal change on each scale considered to be clinically important. [24–26]

Pre-specified secondary outcomes included the Pain Frequency and Bothersomeness Scale, [23] the Study of Osteoporotic Fractures -Activities of Daily Living (SOF-ADL),[28] the EQ-5D[29] (a generic health status measure reflecting mobility, self-care, activity limitations, pain, and psychological distress), opioid medication use, and the Physical Component Summary (PCS) and Mental Component Summary (MCS) subscales of the self-administered SF-36 (version 2).[27] The PCS assesses limitations in self-care, physical, social and role activities, bodily pain, and poor perceived health, while MCS provides indication of psychological distress and social and role disability due to emotional problems. Patients were asked prior to 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, 50–800 procedures). Patients were brought to the fluoroscopy suite, administered conscious sedation, and prepared in sterile fashion. Using fluoroscopic guidance, the skin and subcutaneous tissues overlying the pedicle of the target vertebra or vertebrae were infiltrated with 1% lidocaine and the periosteum of the pedicle/s was infiltrated with 0.25% bupivicaine. Patients were then randomized to the full vertebroplasty procedure or to the control intervention.

For the vertebroplasty procedure, 11- 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 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. Both groups of patients were maintained at bedrest for 1–2 hours prior to discharge.

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

Randomization and Blinding

We used stratified, blocked randomization by clinical site to achieve roughly balanced groups. The block sizes varied between four and 12 patients and were concealed from the research assistants involved in recruitment. These assignments were generated by the data coordinating center (DCC) using a random number generator and then placed in numbered, opaque, sealed

envelopes, using a series of envelopes for each study site. We attempted to blind all patients and study personnel performing follow-up assessments to treatment assignments for the duration of the study. Only the study statisticians, who did not have any contact with study participants, saw unblinded data.

Sample Size

The study was conservatively powered initially to detect differences in both primary and secondary outcome measures (initial N = 250 subjects, two-sided alpha = 0.05, power > 80%, 2.5 point difference on Roland, 1.0 point difference on pain rating). After early difficulty in recruitment and planned interim analysis of the first 90 subjects, we revised our target sample size to 130 randomized subjects with approval from the study data and safety monitoring board (DSMB). The decision to modify the target enrollment was driven primarily by accrual rates and revised power calculations. With the reduced sample size, the study remains powered (>80%) to detect important differences in the primary outcome measures: a 3.0-point difference between groups on the RDQ (assumed SD = 6.7) and a 1.5-point difference on the pain rating (assumed SD = 2.7) at one month.[26]

Statistical Analysis

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 adjusting for baseline values of the outcome measure, recruitment site, and a treatment indicator as the predictor of interest. As a post-hoc analysis, we also used logistic regression models, adjusting for site and baseline values of the outcome measures, to compare the proportion of patients in each group who achieved at least a 30% improvement in RDQ and pain scores, as recommended by the Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials- II to assess the clinical importance of improvement.[31] Further, we performed two post hoc subgroup analyses to determine whether continuous and categorical (1–13, 14–26, and 27–52 weeks) measures of baseline pain duration (as an index of fracture age) interacted with treatment in predicting one-month pain intensity in the ANCOVA models. Formal evaluation of effect modification is based on a partial F-test of whether the 2 interaction terms equal zero. Inference was similar for each subgroup analysis and we report the results of the categorical subgroup analysis.

An independent DSMB reviewed the blinded study results every six months to evaluate safety and efficacy. The DSMB monitored events of death, paralysis, hospitalizations, new onset fractures, new radiculopathy or myelopathy, and infection. The DSMB used O'Brien-Fleming [32] stopping rules of $P < 0.001$ and $P < 0.019$ for two pre-specified interim analyses to evaluate the accumulating evidence for treatment efficacy, though the interim study results did not achieve either threshold. All statistical analyses were performed using R statistical software (version 2.7) [33] and primary pre-specified one-month outcomes were considered significant at $P < 0.043$. All reported P-values are two-sided and not adjusted for multiple testing.

The study was conceived by Drs. Kallmes and Jarvik, with subsequent design input from Drs. Heagerty, Hollingworth, and Turner and Mr. Comstock. The data were gathered by coordinators at each study site and sent to the DCC for quality control and analysis. Dr. Heagerty and Mr. Comstock performed the data analyses. Drs. Heagerty and Jarvik and Mr. Comstock vouch for the data and analysis. Dr. Kallmes wrote the initial draft with substantial contributions from the co-authors.

Results

Between June 2004 and August 2008, 131 patients were enrolled and randomized (Fig. 1). Sixty-eight patients were randomized to vertebroplasty and 63 to the control intervention; all received the allocated 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 prior to one month. One patient (1%) in the vertebroplasty group and two patients (3%) in the control group crossed over prior to one month.

The vertebroplasty and control groups did not differ significantly on either pre-specified primary outcomes of RDQ (treatment difference: 0.7; 95% CI, −1.3 to 2.8; P=0.49) or pain intensity (treatment difference: 0.7; 95% CI, −0.3 to 1.7; P=0.19) at one month (Table 2). Both groups showed substantial improvement in their back-related disability and pain immediately (three days) after the procedure, with comparable improvement between groups. The improvement in each group at three days was maintained at one month.

The treatment groups did not differ significantly on any of the secondary outcomes, including measures of pain and quality of life, at one month (Fig. 2). Further, the groups did not differ in the post-specified proportion of patients achieving clinically meaningful improvement in back pain-related physical disability at one month (40% of vertebroplasty patients vs. 41% of control intervention patients, P=0.99). There was a trend (P=0.06) toward a higher rate of clinically meaningful improvement in pain in the vertebroplasty group (64%) versus the control group (48%).

By three months, 8 (12%) patients in the vertebroplasty group and 27 (43%) patients in the control group crossed over to the other group (P<0.001). The vertebroplasty patients who crossed over reported higher disability and pain at three and 14 days, as compared with the other patients (Fig. 3). Control intervention patients who crossed over showed some early improvement after the control procedure that disappeared by the one-month assessment. However, even after they received the alternative intervention, neither the control nor the vertebroplasty patients who crossed over improved by three months to the extent of non-crossover subjects.

At 14 days, 63% of patients in the control group correctly guessed they had the control intervention, and 51% of patients in the vertebroplasty group correctly guessed they had the vertebroplasty. Both groups expressed similarly moderate confidence in their treatment guess on average (vertebroplasty mean=4.0, control mean=4.1; P=0.78). In the control group, 18 of 33 (55%) patients who adhered to treatment correctly guessed at 14 days that they had received the control intervention compared to 20 of 27 (74%) who eventually crossed over (P=0.12). Notably, among the eight vertebroplasty patients who crossed over to the control intervention, six (75%) guessed incorrectly at one month that they had received the control intervention.

In a post hoc subgroup analysis, the effect of treatment (vertebroplasty vs. control procedure) on one-month pain did not differ significantly across the three baseline pain duration categories (partial F-test P=0.58, 2 degrees of freedom). The treatment effect for patients with less than 13 weeks of pain (treatment difference: 0.8; 95% CI, −0.8 to 2.4; P=0.31) was similar to the results of the overall analysis. The treatment effect for patients with 14 – 26 weeks of pain was 1.3 (95% CI, −0.8 to 3.4; P=0.23) and that for patients with 27 – 52 weeks of pain was 0.0 (95% CI, −1.7 to 1.6; P=0.96).

Adverse events

One patient in the vertebroplasty group suffered injury to the thecal sac during the procedure, with resultant hospitalization. One patient in the control intervention group was hospitalized overnight after the procedure with tachycardia and rigors, of unclear etiology.

Discussion

Patients with osteoporotic vertebral fractures who were randomly assigned to either a full vertebroplasty or a control intervention consisting of a simulated vertebroplasty without infusion of PMMA did not differ significantly one month after the procedure on measures of back pain intensity, functional disability, and quality of life. In this study the confidence interval for the effect on RDQ is (−1.3, 2.8) which excludes a treatment benefit of 3 points or greater and therefore provides evidence against clinically meaningful treatment effects on functional disability. Similarly the confidence interval for the pain rating treatment effect is (−0.3, 1.7) which excludes effects of 2 points or greater. Both treatment groups showed immediate improvement in pain and disability after the procedure, and this improvement was sustained at one month. These results suggest that factors separate from instillation of PMMA may account for the observed clinical improvement following vertebroplasty. Such factors may include the impact of local anesthesia as well as nonspecific effects such as expectations of pain relief ("placebo effect"), natural history, and regression toward the mean.

The impact of the placebo effect on outcomes in this trial remain unclear. Previous studies have documented pain reduction in placebo trials, on the order of 6–7mm on a 100mm scale[34–36]. The treatment effect in our trial was substantially larger than these previous reports, although those reports included pharmacologic and psychological interventions in addition to physical interventions[34].

The vertebroplasty group showed a trend towards a higher proportion of patients with clinically meaningful improvement in pain at one month. Further, there was a higher crossover rate in the control group than in the vertebroplasty group after one 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 with our measure of pain intensity. However, we used a common, validated measure demonstrated to show 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 prevent this, some patients became unblinded and among the unblinded patients, those who still had pain and learned they were in the control group elected to cross over to vertebroplasty.

This study has several limitations. First, because of reluctance of both physicians and patients to accept a control intervention arm for a longer time period, we allowed crossover at one month. This complicates interpretations of group differences in outcome after one month. However, there is evidence that nearly all the benefits of vertebral augmentation accrue within the first month.[37] Additionally, given that the half-life of bupivicaine is only three hours, any benefit from bupivicaine would have disappeared by one month. Second, we did not compare groups on other medical treatments received that might have affected patient outcomes. Third, persistence of pain after vertebroplasty or fracture healing may indicate etiologies for the pain other than fracture, a possibility that our baseline imaging excludes to a certain extent, but not entirely. Fourth, consistent with our previous findings that fracture age was not associated with response to vertebroplasty[14], in this study we did not find a differential treatment effect according to baseline duration of pain, yet it remains possible that vertebroplasty may be effective only for fractures of a certain age or healing stage. Lastly, 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 prior to cement infusion. [38]

In conclusion, at least up to one month, clinical improvement in patients suffering from painful, osteoporotic vertebral fractures is similar between patients treated with vertebroplasty and those treated with a simulated vertebroplasty, without infusion of PMMA. These data suggest that further studies should be undertaken to determine whether long-term outcome is similar between groups, especially because our crossover study design limits the ability of this study to shed light on long-term efficacy of vertebroplasty.

Acknowledgments

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(7):827–31. [PubMed: 12847189]

2. Alvarez L, Alcaraz M, Perez-Higueras A, et al. Percutaneous vertebroplasty: functional improvement in patients with osteoporotic compression fractures. Spine 2006;31(10):1113–8. [PubMed: 16648745]

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(3):113–7. [PubMed: 16460295]

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(4):257–65. [PubMed: 12681451]

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(7):1623–8. [PubMed: 16091504]

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(2):360–7. [PubMed: 15662480]

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(4):683–9. [PubMed: 17416821]

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(9 Pt 1):883–6. [PubMed: 12354821]

9. Peh WC, Gilula LA, Peck DD. Percutaneous vertebroplasty for severe osteoporotic vertebral body compression fractures. Radiology 2002;223(1):121–6. [PubMed: 11930056]

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(8):1313–20. [PubMed: 16923978]

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 osteoporotic vertebral compression fractures. The VERTOS study. AJNR Am J Neuroradiol 2007;28 (3):555–60. [PubMed: 17353335]

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 (2 Pt 1):139–48. [PubMed: 11830619]

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. [PubMed: 18096050]

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(10):1860–3. [PubMed: 11733316]

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(1):73–84. [PubMed: 8551813]

16. Roland M, Morris R. A study of the natural history of back pain. Part 1: Development of a reliable and sensitive measure of disability in low-back pain. Spine 1983;8(2):141–4. [PubMed: 6222486]

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. [PubMed: 2950601]

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. [PubMed: 1408311]

19. Underwood MR, Barnett AG, Vickers MR. Evaluation of two time-specific back pain outcome measures. Spine 1999;24:1104–12. [PubMed: 10361660]

20. Beurskens AJHM, de Vet HCW, Koke AJA. Responsiveness of functional status in low back pain: a comparison of different instruments. Pain 1996;65:71–6. [PubMed: 8826492]

21. Roland M, Fairbank J. The Roland-Morris Disability Questionnaire and the Oswestry Disability Questionnaire. Spine 2000;25:3115–24. [PubMed: 11124727]

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(10):2652–7. [PubMed: 16286418]

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(17):1899–909. [PubMed: 8560339]

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. [PubMed: 18055266]

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. [PubMed: 11690728]

26. Ostelo RWJG, Deyo RA, Stratford P, et al. Interpreting change scores for pain and functional status in low back pain. Spine 2008;33:90–4. [PubMed: 18165753]

27. Ware, JE.; Kosinski, M.; Dewey, JE. How to score version two of the SF-36 health survey. Lincoln, RI: QualityMetric Incorporated; 2000.

28. Ettinger B, Black DM, Nevitt MC, et al. Contribution of vertebral deformities to chronic back pain and disability. The Study of Osteoporotic Fractures Research Group. J Bone Miner Res 1992;7(4):449–56. [PubMed: 1535172]

29. Brooks R. EuroQol: the current state of play. Health Policy 1996;37:53–72. [PubMed: 10158943]

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(10):1897–904. [PubMed: 9403451]

31. Dworkin RH, Turk DC, Farrar JT, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 2005;113(1–2):9–19. [PubMed: 15621359]

32. O'Brien PC, Fleming TR. A multiple testing procedure for clinical trials. Biometrics 1979;35(3):549–56. [PubMed: 497341]

33. Team RDC. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2008.

34. Hrobjartsson A, Gotzsche PC. Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. N Engl J Med 2001;344(21):1594–602. [PubMed: 11372012]

35. Hrobjartsson A, Gotzsche PC. Is the placebo powerless? Update of a systematic review with 52 new randomized trials comparing placebo with no treatment. J Intern Med 2004;256(2):91–100. [PubMed: 15257721]

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(3):443–52. [PubMed: 12406519]

37. Wardlaw D, Cummings SR, Van Meirhaeghe J, et al. Efficacy and safety of balloon kyphoplasty compared with non-surgical care for vertebral compression fracture (FREE): a randomised controlled trial. Lancet. 2009

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(8):1085–100. [PubMed: 17277923]
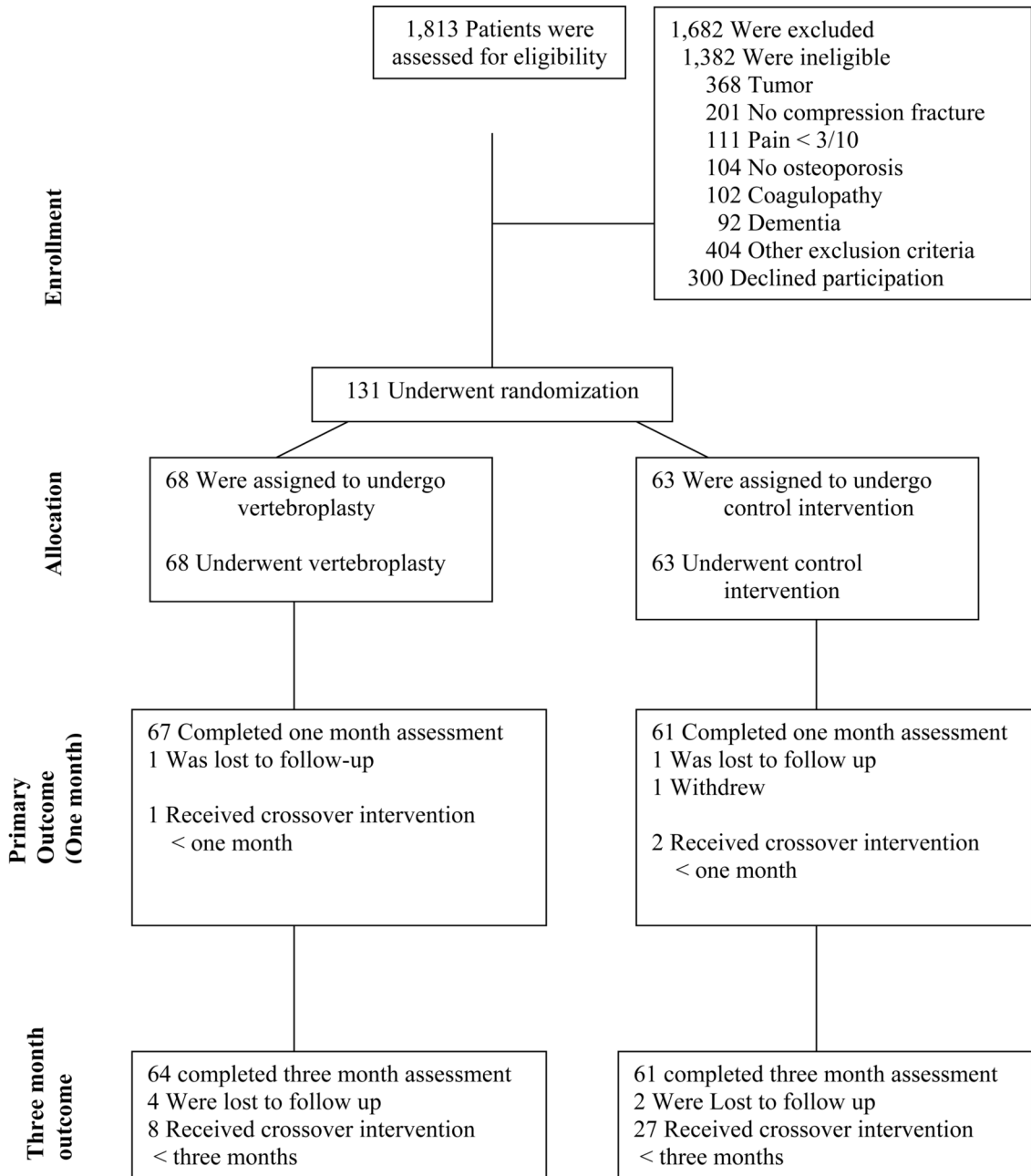
```
┌─────────────────────────┐   ┌──────────────────────────────────┐
│  1,813 Patients were    │   │ 1,682 Were excluded              │
│  assessed for eligibility│  │   1,382 Were ineligible          │
└─────────────────────────┘   │      368 Tumor                   │
                              │      201 No compression fracture  │
                              │      111 Pain < 3/10             │
                              │      104 No osteoporosis          │
                              │      102 Coagulopathy             │
                              │       92 Dementia                │
                              │      404 Other exclusion criteria │
                              │   300 Declined participation     │
                              └──────────────────────────────────┘
```

Enrollment

```
              ┌──────────────────────────────┐
              │ 131 Underwent randomization  │
              └──────────────────────────────┘
```

Allocation

68 Were assigned to undergo vertebroplasty	63 Were assigned to undergo control intervention
68 Underwent vertebroplasty	63 Underwent control intervention

Primary Outcome (One month)

67 Completed one month assessment	61 Completed one month assessment
1 Was lost to follow-up	1 Was lost to follow up
	1 Withdrew
1 Received crossover intervention < one month	2 Received crossover intervention < one month

Three month outcome

64 completed three month assessment	61 completed three month assessment
4 Were lost to follow up	2 Were Lost to follow up
8 Received crossover intervention < three months	27 Received crossover intervention < three months

Figure 1.
Enrollment, Randomization, and Follow-up of the Study Participants.
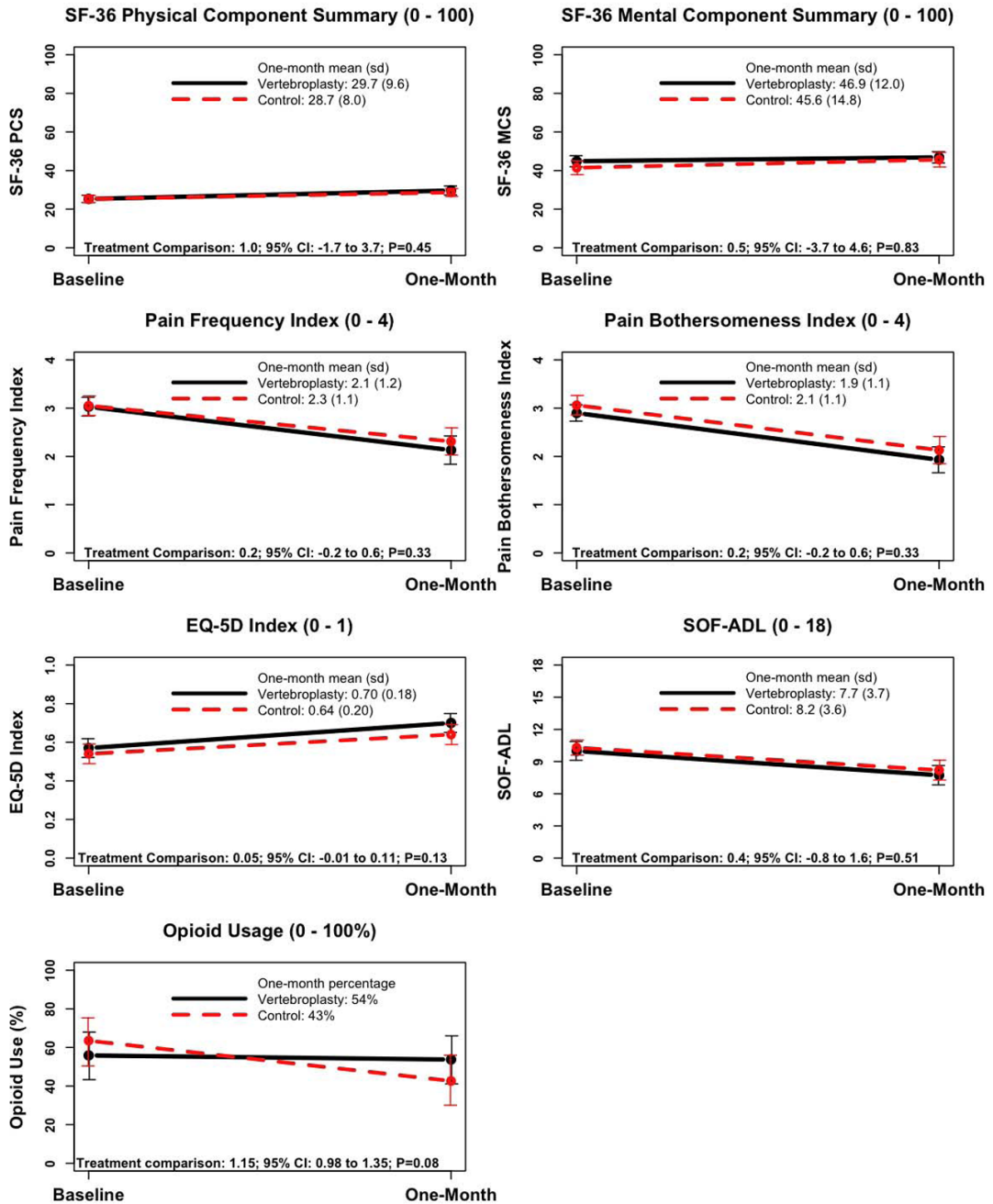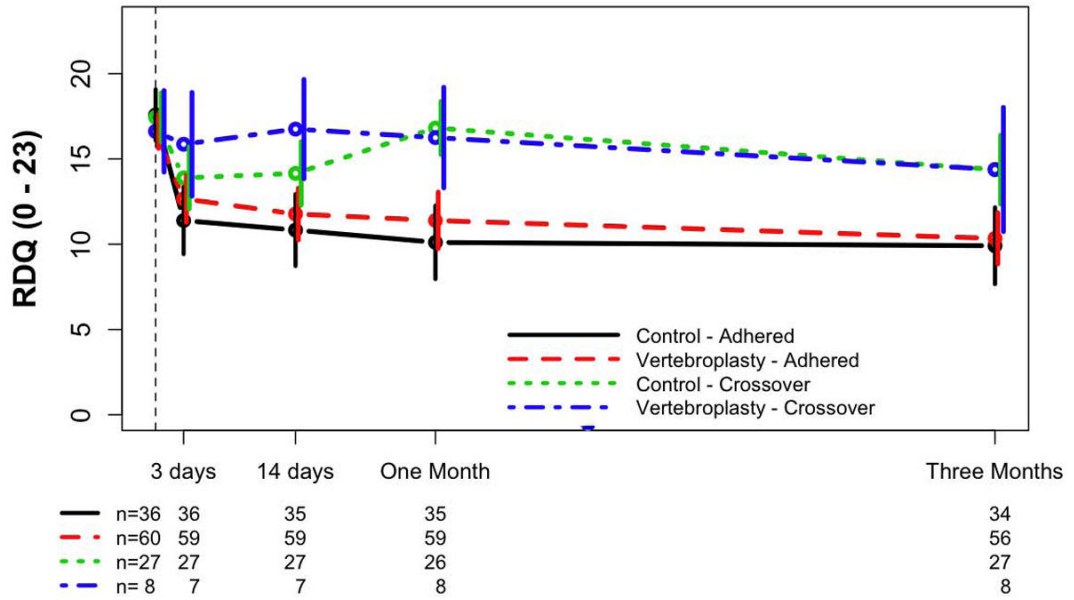
Figure 2.
Treatment Comparisons on Secondary Outcome Measures (Intent-to-Treat Analyses) at One Month.
† For continuous outcome measures, treatments were compared from ANCOVA models that adjusted for randomized treatment assignment, the baseline value of the outcome measure, and recruitment site. Treatment effects are coded such that negative treatment effects favor the control procedure and positive treatment effects favor vertebroplasty. Treatment comparisons and confidence intervals for one-month opioid use are reported as an odds ratio from a logistic regression model, adjusting for baseline opioid use and recruitment site. Odds ratios less than 1.0 favor vertebroplasty and greater than 1.0 favor the control procedure.

Roland Disability Questionnaire (0 - 23)

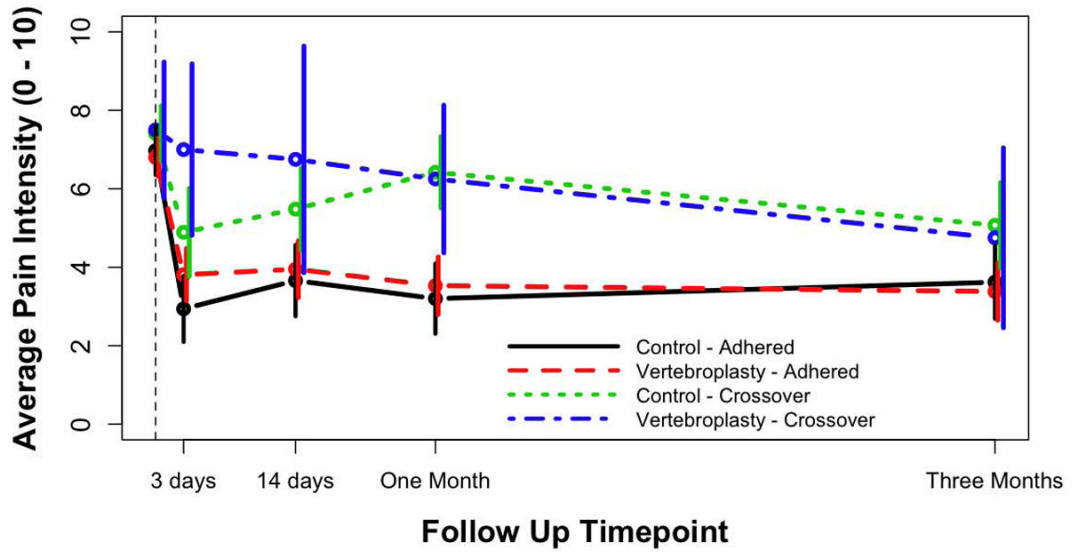## Average Pain Intensity (0 - 10)

Figure 3.
Mean (95% confidence intervals) Roland-Morris Disability Questionnaire (RDQ) and pain numerical rating scores (average pain intensity, prior 24 hours) by treatment adherence over time. Treatment "crossover" was defined as having had the non-randomized procedure prior to the three-month study assessment.

Table 1

Baseline Characteristics of Study Participants.

Characteristic	Vertebroplasty (N=68)	Control (N=63)
Study Recruitment Site - no. (%)		
Mayo Clinic	14 (21%)	16 (26%)
United Kingdom	26 (38%)	26 (41%)
Australia	13 (19%)	9 (14%)
USA (Non-Mayo Clinic)	15 (22%)	12 (19%)
Age - mean yr (sd)	73.4 (9.4)	74.3 (9.6)
White - no. (%)	67 (99%)	60 (95%)
Female - 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 other - no. (%)	42 (62%)	27 (43%)
Employment - no. (%)		
Full/Part Time	7 (10%)	5 (8%)
Retired	42 (62%)	39 (62%)
Disabled	10 (15%)	10 (16%)
Other	9 (13%)	9 (14%)
Smoker - no. (%)	12 (18%)	9 (14%)
Comorbidity index[15] (0–28) - mean (sd)	1.9 (2.1)	2.0 (1.9)
Workers' compensation - no. (%)	9 (13%)	7 (11%)
Pain duration (weeks) - no. (%)		
1 – 13	30 (38%)	24 (44%)
14 – 26	14 (24%)	15 (21%)
27 – 39	8 (14%)	9 (12%)
40 – 52	16 (24%)	15 (23%)
median weeks (IQR)	16 (10 – 36)	20 (8 – 38)
Number of levels treated – no. (%)		
1	48 (71%)	41 (65%)
2	13 (19%)	14 (22%)
3	7 (10%)	8 (13%)
Self-Reported Opioid Analgesic Use – no. (%)	38 (56%)	40 (63%)
RDQ[*] – mean (sd)	16.6 (3.8)	17.5 (4.1)
Average Pain Intensity last 24 hours[†] - mean (sd)	6.9 (2.0)	7.2 (1.8)
SF-36 (v2) - mean (sd)		
Physical Component Summary[§]	25.3 (7.8)	25.3 (7.3)
Mental Component Summary[§]	44.8 (11.8)	41.5 (14.1)

Characteristic	Vertebroplasty (N=68)	Control (N=63)
Pain Frequency Index¶ - mean (sd)	3.0 (0.8)	3.1 (0.8)
Pain Bothersome Index¶ - mean (sd)	2.9 (0.7)	3.1 (0.8)
EQ-5D** - mean (sd)	0.57 (0.18)	0.54 (0.23)
SOF-ADL†† - mean (sd)	10.0 (3.6)	10.3 (2.8)

[*] RDQ ranges from 0 to 23 with higher scores indicating worse physical disability.

[†] Average Pain Intensity ranges from 0 = 'no pain' to 10 = 'pain as bad as could be'.

[§] Norm-based SF-36 (v2) scores are presented. In a healthy population, SF-36 scores have a mean of 50 and standard deviation of 10 with lower scores indicating worse outcome.

[¶] Pain Frequency and Bothersome indices range from 0 to 4 (4 = worst pain).

[**] EQ-5D ranges from −0.1 to 1.0 with higher scores indicating better health.

[††] SOF-ADL scores range from 0 to 18 with higher scores indicating more back-related disability.

Table 2

Treatment Comparisons on Primary Outcomes (Intent-to-Treat analyses).

Primary Outcome Measures	Vertebroplasty M (SD)	Control M (SD)	Treatment Comparison (95% CI)[§]	P-value[§]
RDQ[*]				
Baseline	16.6 (3.8)	17.5 (4.1)		
3 Days	13.0 (5.2)	12.5 (5.5)	−0.9 (−2.7, 0.8)	0.30
14 Days	12.4 (5.8)	12.3 (5.9)	−0.6 (−2.4, 1.2)	0.35
1 Month	12.0 (6.3)	13.0 (6.4)	0.7 (−1.3, 2.8)	0.49
Average Pain Intensity[†]				
Baseline	6.9 (2.0)	7.2 (1.8)		
3 Days	4.2 (2.8)	3.9 (2.9)	−0.4 (−1.5, 0.5)	0.37
14 Days	4.3 (2.9)	4.5 (2.8)	0.1 (−0.8, 1.1)	0.77
1 Month	3.9 (2.9)	4.6 (3.0)	0.7 (−0.3, 1.7)	0.19

[*] RDQ ranges from 0 to 23 with higher scores indicating worse physical disability.

[†] Average Pain Intensity ranges from 0 = 'no pain' to 10 = 'pain as bad as could be'.

[§] Treatment comparisons, confidence intervals and P-values are calculated from ANCOVA models that adjust for randomized treatment assignment, the baseline value of the outcome measure, and recruitment site. Negative treatment effects favor the control procedure and positive treatment effects favor vertebroplasty.