Clinical Study

Assessment of patient’s pain-related behavior at physical examination may allow diagnosis of recent osteoporotic vertebral fracture

Roberto Postacchini, MD a,b,*, Michela Paolino, MD c, Silvia Faraglia, MD b, Gianluca Cinotti, MD c, Franco Postacchini, MD c

a IUSM (Italian University of Sport and Movement), Rome, Italy. Department of Anatomy, Piazza Lauro De Bosis, 6 - 00135, Rome, Italy
b Section of Orthopedic Surgery, Israelitic Hospital, Piazza San Bartolomeo all’Isola 21 Rome 00186 Italy
c Department of Anatomy, Histology, Locomotor Apparatus and Legal Medicine, Section of Orthopedic Surgery and Traumatology, University Sapienza, Rome, Italy

Received 18 March 2013; revised 11 July 2013; accepted 13 July 2013

Abstract

BACKGROUND CONTEXT: Although innumerable studies have analyzed the multiple aspects of osteoporotic vertebral fractures, no study has focused on the clinical features related to spine pain in patients with recent osteoporotic vertebral compression fractures (VCFs).

PURPOSE: To determine whether the assessment of pain-related behavior (P-RB) of patients with osteoporotic VCFs of recent onset may allow the fracture to be strongly suspected, or even diagnosed, at physical examination.

STUDY DESIGN: Pain-related behavior of elderly patients attending an outpatient spine clinic was evaluated on the basis of six consecutive movements made on the examining table.

PATIENT SAMPLE: Fifty-six patients complaining only of lumbar or thoracic pain. The fractured patients (FPs), representing the fracture group (FG), were the 19 who had a recent VCF, whereas the control group (CG) consisted of the remaining 37 patients.

METHODS: Assessment of P-RB was based on six parameters: grimacing, sighing, clenching or blocking eyelids, gaping or strongly tightening the lips, need for help to take positions, and extreme difficulty to turn in the prone position. A score of 1 or a decimal was assigned to each parameter, the final score to each patient being 0 to 6. Three types of injury, acute (I), subacute (II), or chronic (III), were identified on the basis of the time elapsed from the probable occurrence of the fracture. The diagnosis of recent fracture was based on magnetic resonance images. Patients were videotaped during their movements. An examiner, unaware of the clinical history and diagnosis, gave a P-RB score to all patients and indicated whether they had to be placed in FG or CG, and also their presumable type of fracture. Subsequently, a DVD with the videotapes of all patients was given to three independent examiners, not specifically expert of spine conditions, who were asked to make the same evaluations as the first examiner.

RESULTS: The mean scores for P-RB given by the first examiner were 4.6 to FG and 0.7 to CG (p<.01). He identified as FPs 89% of those who were in FG. The type of fracture was indicated correctly in 88% of patients identified as FPs. The mean scores for the three types of fracture ranged from 5.4 (Type I) to 3.3 (Type III) (p<.001). The mean scores for P-RB given by the independent examiners to FG and CG were similar to those of the first examiner. The rates of correctness in identifying the type of fracture in patients indicated as FPs varied from 87% to 80%. The mean scores assigned to the patients included in the three types of fracture ranged from 5.4 to 2.8.

CONCLUSIONS: Pain-related behavior evaluation of patients with osteoporotic VCF during their movements on the examining table may allow to suspect, or even diagnose, the presence of a fracture, particularly in the initial 4 to 6 weeks after the occurrence. Even orthopedic surgeons not...
Introduction

The incidence of osteoporotic vertebral compression fractures (VCFs) has been rising in the last decades because of increasing age of the population. Although these fractures may be asymptomatic and heal undetected, many recent injuries cause severe pain and often become chronically painful and responsible for considerable disability [1]. Furthermore, the presence of multiple fractures in a patient, occurring in the course of months or years, may be associated with several complications, such as decrease in pulmonary function, hyperkyphosis, abdominal distension and early satiety leading to malnutrition, as well as increased mortality [2–5]. The recognition and adequate management of these injuries thus represent a problem of considerable clinical and social impact.

Innumerable studies have analyzed the etiology, diagnostic investigations, prognosis, prevention, and management of osteoporotic VCFs. However, an extensive review of the literature has failed to reveal any studies focused on the clinical features of recent osteoporotic VCFs. On the other hand, the knowledge of these features is of paramount importance to allow these injuries to be suspected and diagnosed early. It is well known, in fact, that osteoporotic VCFs are often diagnosed after several weeks from the occurrence, during which the vertebral body may undergo a collapse and/or wedge-shaped deformity.

The description of the clinical features of acute osteoporotic VCFs in journal articles or texts on spine care usually just reports that “percussion on a single level will be painful with nontender adjacent segments” [6], “pain is exacerbated by standing erect or with a change of position” [7], or “patients with acute pain may demonstrate a loss of bowel sounds, ileus, or bladder distension” [8]. However, our clinical experience agrees with those descriptions only partially. We had the impression that patients with a recent VCF very often behave in a particular way during physical examination, as to allow the examiner to differentiate them from patients with degenerative spine diseases and give rise immediately to the suspicion of fracture.

Keefe et al. [9] first introduced the pain assessment method, based on the observation of pain-related behavior (P-RB) in patients with low back pain, during physical examination. This method was then used by Walsh et al. [10] and Carragee et al. [11] to evaluate the severity of pain in asymptomatic subjects undergoing lumbar discography. The same method was used in this prospective, controlled study.

Our hypothesis was that P-RB of patients with an osteoporotic VCF might suggest the presence of a fracture, and also that orthopedic surgeons, not specifically familiar with spine diseases could make a correct diagnosis based on the behavior of those patients.

Materials and methods

In a 4-month period, all new patients aged 65 to 85 years who attended our outpatients’ spine clinic, because of pain in the lumbar or thoracic spine, were enrolled in the study, unless the pain was associated with radicular or cord symptoms in the lower limb(s) or related to vertebral infections or tumors.

Protocol

To assess P-RB during physical examination, the patients were asked to take six consecutive positions on the examining table: to sit on the edge of the bed, to lie supine, to turn on the flank, to take the prone position, to turn again on the flank, and to sit on the edge of the bed. Furthermore, they were asked to be taken with a videocamera during the movements. The patients who refused or did not give a written permission to be videotaped were not entered in the study. Six parameters were used to evaluate the patient’s behavior: grimacing, sighing, clenching or blocking eyelids, gaping or strongly tightening the lips, need for help by the examiner to take positions, and refusal, or extreme difficulty, to turn to the prone position. It was considered that the patient had demonstrated a P-RB when she/he had shown at least two of these types of behavior. The positivity of P-RB, reflecting the severity of pain, was graded 1 (very mild pain) to 6 (very severe pain), whereas 0 corresponded to a negative behavior.

The diagnosis of a recent vertebral fracture was made on the basis of magnetic resonance images. For a fracture to be considered of recent onset, the vertebral body had to be deformed or show a fracture, compared with the adjacent normal vertebrae and the signal, in a part or the entire vertebral body, had to be hypointense on the T1-weighted images, hyperintense on T2-weighted sequences, and hyperintense on fat-suppression sequences. The latter sequences were not available in two cases, but for both of them, there was no doubt that the vertebral body was fractured due to its deformation and the signal changes in T1 and T2 sequences. A fracture was considered as long standing when the vertebral body was deformed and the magnetic resonance imaging signals were normal. Furthermore, based on the time interval between the onset of pain and the diagnosis, three types of recent fractures were
arbitrarily identified on the basis of presumable occurrence of the fracture: Type I, acute (≤1 month), Type II, subacute (from 31 to 60 days), and Type III, chronic (3 or 4 months).

**First stage**

All patients were interviewed by a single young orthopedic surgeon (PM), who recorded the duration and site of the reported pain. Then she asked a senior orthopedic surgeon (PF), who visited patients in an adjacent room, to attend. He asked the patients the permission to be videotaped while doing the six movements described previously. After videotaping, he left the visiting room and gave a score to each P-RB (1 or a decimal for each parameter), indicated who was a fractured patient (FP) and who was a control, and to which of the three types of recent fracture FP could be assigned. The sheet, containing the score and the type of fracture, was put in an envelope that was closed and left in the patient’s clinical chart after it was marked with an “F,” for those patients whom the senior orthopedic surgeon considered to have a recent fracture. Then the young surgeon continued on his own the physical examination and evaluated the investigations that the patients had already undergone. Additional investigations were prescribed or immediately carried out in hospital when necessary and possible. The final diagnosis was written in the clinical chart for those patients who were found to have a recent osteoporotic VCF based on the imaging studies. These patients, as well as those who were asked to return after additional investigations, had one of the three treatment modalities: analgesic medications, light corset, and/or physiotherapy (most patients with chronic fracture), the prescription of a three-point corset or a vertebral body augmentation procedure (most patients with acute or subacute fracture). The choice between the latter two modalities was made on the basis of the interval of time elapsed from the fracture, the severity of spine pain, and the patient’s decision to undergo a conservative or surgical treatment.

The videotape of each patient was recorded in a DVD, which was placed in an envelope, containing the clinical chart, number. The envelopes were numbered progressively, adding an “F” to the number when an osteoporotic VCF was definitely diagnosed and then were closed.

In the study period, 69 patients met the criteria to be enrolled in the study. Thirty of them, none with a recent fracture, refused to be videotaped. Thus, the study cohort consisted of 56 patients, 19 of whom had a final diagnosis of recent fracture. Then the envelopes containing the evaluations of the senior orthopedic surgeon were opened and the scores for P-RB were recorded. The diagnosis that he made, as shown by “F” on the envelope and the type of fracture that he indicated were compared with the final diagnosis reported on, and the type of fracture emerging from, the clinical chart.

Three orthopedic surgeons (PR, FS, CG) not involved in the patients’ evaluation or treatment and not specifically expert in spine conditions were asked to assess all DVDs. To this purpose, three copies of DVDs of the 52 patients were done. Each DVD of a series was put in a new envelope with the same progressive number marked on the previous envelope. A series of DVDs was given to each examiner, who thus had no clinical information on the patients. The examiners were asked to give a score of 0 to 1 for each P-RB to each patient in the study cohort, to detect which patients were FPs and which controls, and to indicate whether the fracture was Type I, II, or III for those patients who they identified as FPs. The assessments were written on a sheet which was put again into the individual patient’s envelope together with DVD. The examiners were asked to evaluate three times the DVDs in three different days and report their assessments in three different sheets without knowing about the previous assessments. They were also invited not to communicate their evaluations to each other.

**Statistical analysis**

Mean, standard deviation, and 95% confidence interval were computed for all set measurements. The Lilliefors (Kolmogorov-Smirnov) test on normality, performed to check all assessed variables, showed a non-normal distribution (p<0.05). The Mann-Whitney test was used to assess the differences between the mean of the scores assigned to both FG and CG. The differences between the mean scores, given to FPs by the senior orthopedic surgeon and the independent examiners, were assessed with the Wilcoxon signed-rank test. The same was done with the CG scores. The Chi-square test was used to assess differences between percentages. A reliability analysis, using the Kappa statistic, was performed to determine the consistency within and among raters. The ability of our six-movements test.

<table>
<thead>
<tr>
<th>Age (y)</th>
<th>Sex</th>
<th>Pain duration (d)</th>
<th>Site of fracture</th>
</tr>
</thead>
<tbody>
<tr>
<td>65–74</td>
<td>9</td>
<td>F 16 Type I Type II Type III</td>
<td>T8 1 L1 4</td>
</tr>
<tr>
<td>75–82</td>
<td>10</td>
<td>M 3 1–15 3 31–45 4 61–75 3</td>
<td>T10 2 L2 3</td>
</tr>
<tr>
<td>16–30</td>
<td>8</td>
<td>46–60 1 76–120 0</td>
<td>T11 3 L3 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>T12 4 L4 1</td>
</tr>
</tbody>
</table>

VCF, vertebral compression fracture; F, female; M, male.

Table 1

Demographic data, duration of pain, and level of fracture in the 19 patients with VCF.
to discriminate FPs from control patients was assessed using the receiver operating characteristics (ROC) curve analysis, the area under ROC curve (AUC) being calculated by the trapezoid method. A level of significance at $p = .05$ was set for all tests. SPSS for Windows (version 17.0, SPSS, Chicago, IL, USA) was used for the statistical analysis.

### Results

In FG, the vast majority of patients were females, the mean age being 72 years (range, 65 to 82 years). The fracture was Type I in 11 patients, Type II in 5, and Type III in 3. The demographic data, duration of pain, and site of fracture are reported in detail in Table 1. Control group consisted of 26 women and 11 men with a mean age of 73 years (range, 65 to 84 years). There was no significant difference between the two groups in regard to age and sex.

The mean score of the senior orthopedic surgeon for P-RB to the patients in FG was 4.6 points (Table 2), whereas that assigned to those in CG was 0.70 points ($p < .001$). He identified as FPs 17 (89%) of the 19 patients in FG (Table 2). The mean scores, given to the patients who had originally been assigned to the three types of fractures, are reported in Table 2. The highest score for Type I was 6, assigned to four patients (Fig. 1), and the highest for Type II was 4.5 (Fig. 2). None of the FPs scored less than 3 points, whereas it was the case with all those in CG, the majority of whom had a score of 0 points (Fig. 3). The differences in the scores between Type I and Type II or III were significant ($p < .001$ for both), whereas the difference between Type II and Type III was not significant ($p > .05$). The type of fracture was indicated correctly in 15 (88%) of the 17 patients identified as FPs (Table 2). The two errors concerned a Type I fracture, mistaken as Type II, and a Type III, assigned to Type II.

The mean scores of the independent examiners to FG and CG are reported in Table 3; the differences between the two groups being significant for all of the three examiners ($p < .001$), with no significant difference compared with those of the senior orthopedic surgeon ($p = .3$). All examiners identified as FPs the majority of the patients of FG (Table 3), with no significant difference compared with the rate of correctness of the senior orthopedic surgeon ($p = .4$). The mean scores for P-RB of FPs ranged from 4.0 to 4.3 points (Table 3). The rates of correctness in identifying the type of fracture of those indicated as FPs varied from 87% to 80% (Table 3), the differences with the rate of the senior orthopedic surgeon being not significant ($p = .9$). The mean scores for P-RB given by each examiner to the patients originally assigned to Type I, Type II, or Type III fracture are reported in Table 3. The differences

<table>
<thead>
<tr>
<th>Type</th>
<th>Correctness in identifying FPs</th>
<th>Mean scores for P-RB of FPs</th>
<th>Correctness in identifying FT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type I</td>
<td>11 (of 11)</td>
<td>Type I 5.4</td>
<td>Type I 10</td>
</tr>
<tr>
<td>Type II</td>
<td>4 (of 5)</td>
<td>Type II 3.9</td>
<td>Type II 3</td>
</tr>
<tr>
<td>Type III</td>
<td>2 (of 3)</td>
<td>Type III 3.3</td>
<td>Type III 2</td>
</tr>
<tr>
<td>Total</td>
<td>89%</td>
<td>Mean 4.6</td>
<td>Total 88%</td>
</tr>
</tbody>
</table>

FPs, patients of the fracture group identified as fractured; P-RB, pain-related behavior; FT, fracture type.

Fig. 1. A 76-year-old woman who had a T8 fracture 14 days before the examination. (A) Images showing the patient while trying to take a supine position, (B) lying on the left flank, (C) standing up to sit and (D) almost seated.
between the mean scores of Type I and Type III were significant for all three examiners ($p < 0.01$). Instead, the differences between Type I and Type II and those between Type II and Type III were not significant ($p = 0.1$). However, even the patients of the latter type, scored higher than those in CG ($p < 0.01$). Four of the total eight errors of the examiners were Type II fractures mistaken as Type III or vice versa.

The intra-rater and inter-raters reliabilities in scoring P-RB of patients assigned to three types of fracture were 0.94 and 0.88, respectively.

With a cutoff for diagnosis of 2/6 P-RBs, AUC of the ROC curve was 0.994 (SE, 0.007) for the senior orthopedic surgeon (Fig. 4); the six-movements test that was used showed a sensitivity and specificity of detecting FPs of 95% and 97%, respectively, the positive predictive value (PPV), negative predictive value (NPV), and accuracy being 95%, 97%, and 96%, respectively. Also, AUC of the ROC curve was 0.994 (SE, 0.007) for the first independent examiner (Fig. 5). The AUC of the ROC curve was 0.993 (SE, 0.008) for either the second and third independent examiner (Fig. 6). For the first two independent examiners, the test showed a sensitivity of 95%, a specificity of 97%, and a PPV, NPV, and accuracy of 95%, 97%, and 96%, respectively. For the third examiner, the sensitivity and specificity were 91% and 95%, respectively, and PPV, NPV, and accuracy were 91%, 95%, and 96%, respectively.

### Discussion

Little is still known about the clinical features of osteoporotic VCFs not only by general practitioners but probably also by orthopedic surgeons or neurosurgeons. This consideration stems from the fact that these fractures are often missed or diagnosed after weeks or months from the occurrence. Our study showed that most elderly patients with an osteoporotic VCF of recent onset have a peculiar behavior that should lead to strongly suspect, or even diagnose, the presence of a fracture at physical examination.

The assessment of P-RB is a consolidated method for evaluating a subject’s severity of pain. It has been widely used in experimental animals to determine the effect of anatomic lesions, substances, or medications [12–15], whereas in human subjects, the studies mostly focused on pain-related fear [16–19]. Walsh et al. [10], to evaluate P-RB of subjects undergoing discography, videotaped them during disc injection. The same procedure was followed by Carragee et al. [11] to assess the pain response to lumbar discography in patients without low back symptoms. In the present study, we applied a similar protocol and videotaped either patients with an osteoporotic VCF or control patients to assess their behavior while doing six movements. No methods of patient’s subjective pain evaluation was used because our aim was to determine whether osteoporotic
VCFs could be suspected by an examiner at physical examination of patients.

Most patients in our cohort with an acute (Type I) fracture showed a behavior suggesting a very severe to excruciating pain compared with no or very mild pain in the majority of control patients. Most patients with a subacute (Type II) or chronic (Type III) injury exhibited a behavior indicating a progressive decrease in pain over time. These impressions were reflected by the significantly higher mean scores assigned both by the initial examiner and the independent examiners to the patients with Type I fracture, compared with those with Type III. However, a few patients with a Type II fracture assessed in the first 2 weeks after injury had only slightly lower scores than those with Type I fracture. Furthermore, even the patients with Type III fracture reached a higher mean score than the controls. These results indicate that careful observation of any elderly patient with a recent osteoporotic VCF may allow the injury to be suspected. This is true particularly for patients with acute fracture and those with subacute injury assessed in the initial weeks of the second month after injury.

Table 3
Assessments of the three independent examiners

<table>
<thead>
<tr>
<th>Correctness in identifying FPs</th>
<th>Mean scores for P-RB</th>
<th>Correctness in identifying FT</th>
<th>Mean scores for P-RB of FPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ex. 1 16 (84%) 4.1 0.7</td>
<td>14 (of 16) 87% 5.2 4.3 2.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex. 2 16 (84%) 4.0 0.9</td>
<td>13 (of 16) 81% 5.3 3.6 3.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex. 3 15 (79%) 4.3 1.0</td>
<td>12 (of 15) 80% 5.4 4.1 3.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

FPs, patients of the fracture group identified as fractured; P-RB, pain-related behavior; Ex., examiner; FT, fracture type.

Fig. 3. A 72-year-old woman of the control group, with chronic low back pain, (A) while taking a supine position, (B) turning on the left flank, (C) taking the prone position, and (D) beginning to stand up to sit on the edge of the bed.

Fig. 4. Receiver operating characteristic curve obtained with fractured patients and control group patients of the senior orthopedic surgeon’s series. The area under the curve was 0.994.
The first examiner of both the fractured and control patients was a senior orthopedic surgeon who was blind of the final diagnosis of each patient but more expert in the behavior of patients with an osteoporotic VCF than the three independent examiners. Nevertheless, the evaluations of the latter coincided almost consistently with those of the first examiner. This finding not only corroborates the results of this study but also indicates that any orthopedic surgeon can be able to identify patients with a recent osteoporotic fracture by careful observation of their behavior during physical examination. Our six-movements test of assessment of P-RB of patients with a recent VCF thus appears a valid diagnostic instrument because it was found to have an extremely high sensitivity, specificity, and accuracy when applied either by the first, expert, examiner and by the independent, less expert, examiners.

Neither the first examiner nor the independent examiners identified all patients with a fracture, but the percentages of correctness ranged from 89% to 79%. The reason for the mistakes was probably related to the fact that patients in the last half of the second month and in the third month after fracture may have behaviors that can simulate those of non-FPs. This interpretation is consistent with the finding that most mistaken patients had a late Type II or a Type III fracture. Similarly, the type of fracture was not indicated correctly for all patients who were identified as FPs. However, the first examiner did not indicate correctly the type of fracture only in two cases, whereas the total number of mistakes made by the independent examiners was eight. These errors may depend on the similarity of behaviors of patients with Type I fracture and the behaviors of those evaluated in the first weeks of the second month. On the other hand, patients with late Type II fracture may have similar behaviors as those with Type III injury. Furthermore, the reason why the independent observers made a higher number of mistakes than the senior orthopedic surgeon in the identification both of FPs and of the type of fracture is probably related to their less experience with osteoporotic VCFs.

All patients in CG had a chronic low back pain, except for a few who had an acute spine pain. The latter had higher scores than the others, but their maximum score did not reach that of any of the FPs. These findings are conceivable.
considering that, in the absence of any injury, elderly patients rarely have a severe acute pain as younger patients may complain of. Furthermore, not only patients with spine infections but also those with radicular symptoms, whose scores might be high, were excluded from the study. This was done because the typical patient with an osteoporotic vertebral fracture complains only of back pain.

All our patients lived in a European country and were white. Nevertheless, we believe that patients with a recent osteoporotic VCF of any industrialized country and any culture would have a very similar behavior to that of our FPs. This conviction stems from data indicating that the prevalence of low back pain is similar in diverse industrialized countries [20] and the degree of decrease in quality of life after osteoporotic VCF is not dissimilar in patients living in five different western countries [21]. In addition, we found that our FG included individuals of even very different social and economic status, as well as personal culture.

A limitation of this study is that FG included fractures of T8–L4 and none of the upper thoracic spine, which are rare anyway. A further limitation is the relatively low number of patients in FG. Nonetheless, the study appears to be relevant because it shows that early identification of an osteoporotic VCF can often be done by mere observation of patient’s behavior, thus allowing correct investigations and adequate treatments to be carried out early.

References