ABSTRACT

Background: We analyzed the International Classification of Diseases, 10th edition (ICD-10) diagnostic codes, procedure codes, and radiographic image codes for vertebral fracture (VF) used in the database of Health Insurance Review and Assessment Service (HIRA) of Korea to establish a validated operational definition for identifying patients with osteoporotic VF in claims data.

Methods: We developed three operational definitions for detecting VFs using 9 diagnostic codes, 5 procedure codes and 4 imaging codes. Medical records and radiographs of 2,819 patients, who had primary and subordinated codes of VF between January 2016 and December 2016 at two institutions, were reviewed to detect true vertebral fractures. We evaluated the sensitivity and positive predictive value (PPV) of the operational definition in detecting true osteoporotic VF and obtained the receiver operating characteristic (ROC) curve.

Results: Among the 2,819 patients who had primary or secondary diagnosis codes for VF, 995 patients satisfied at least one of the criteria for the operational definition of osteoporotic VF. Of these patients, 594 were judged as having true fractures based on medical records and radiographic examinations. The sensitivity and PPV were 62.5 (95% confidence interval [CI], 59.4–65.6) and 59.7(95% CI, 56.6–62.8) respectively. In the receiver operating characteristic analysis, area under the curve (AUC) was 0.706 (95% CI, 0.688–0.724).

Conclusion: Our findings demonstrate the validity of our operational definitions to identify VF more accurately using claims data. This algorithm to identify VF is likely to be useful in future studies for diagnosing osteoporotic VF.

Keywords: Claim Database; Vertebral Fracture; Operational Definition; Osteoporosis
INTRODUCTION

Vertebral fracture (VF) is the most common osteoporotic fracture, that occurs in 30–50% of people aged over 50 years. Patients with VF experience marked impairment in mobility and increased mortality; this impairs the patients’ quality-of-life and poses a serious socioeconomic burden on the healthcare system, especially among older adults. VF can present with non-specific symptoms or be asymptomatic, thus making clinical detection difficult. The diagnosis of VFs requires spine imaging by lateral radiography. However, it may also be diagnosed by chance on an X-ray requested for other reasons.

Recently, large-scale medical databases have been used for big data research worldwide. Medical databases are powerful tools that support clinical and epidemiologic studies of disease burden and treatment outcomes. Specifically, a claims database can provide large-scale nationwide information for analysis while also preventing selection bias. When properly used, this information allows researchers to simulate predictive scenarios through the use of artificial intelligence such as machine learning and deep learning. However, claims databases were established for financial reimbursement and not medical research; therefore, analysis using claims data may be limited due to potential lack of clinical information and possibility of coding errors.

The lack of details on individual cases, such as VF injury mechanism (i.e., trauma type, osteoporosis), radiographs, and bone mineral density results, poses significant limitations on conclusions that may be drawn through analysis using claims database. Among previous studies, osteoporotic VF has had varied definitions when using large-scale medical claim data; there is no standard operational definition for osteoporotic VFs. Therefore, researchers should develop algorithms to identify osteoporotic VF prior to the study to reliably validate their operational definitions. For this reason, establishing and validating the operational definition of osteoporotic VF is necessary, especially when conducting research using national health claim data.

Here, we analyzed the International Classification of Diseases, 10th edition (ICD-10) diagnostic codes, procedure codes, and radiographic image codes for VF used in the database of Health Insurance Review and Assessment Service (HIRA) of South Korea to establish a validated operational definition for identifying patients with osteoporotic VF in claims data.

METHODS

Development of operational definition to identify osteoporotic VFs

Claims data submitted by our hospital to the national health insurance service were used in this study. The dataset contains demographic data, including age and sex; physician information; and hospital-related comorbidities, as described by the diagnostic codes.

Three of the authors (SMP, a spine surgery specialist with 7 years of experience; YKL, a hip surgery and osteoporosis specialist with 13 years of experience; and TYK, a hip surgery and osteoporosis specialist with 16 years of experience) collaboratively developed three criteria using 9 diagnostic codes (M48.4, M48.5, M49.5, M80.8, S22.0, S22.1, S32.0, S32.7, T08.0), 5 procedure codes (N0471, N0472, N0473, N0474, N0630) and 4 imaging codes (G430, G440, G450, G460) of the ICD-10 code (Table 1). Meeting any one of the three criteria was indicative of a VF.
If the patient had multiple claims for VF, only the first claim was included. We excluded patients with conditions that were considered high-impact traumas (multiple fractures).

Criterion 1 was a combination of 9 diagnostic codes and 5 procedure codes. The date of procedure coding was defined as the index date. A VF diagnostic code should have been coded within 1 month before the procedure codes.

Criterion 2 indicated an admission with 9 diagnostic codes as primary diagnosis. Date of the admission was defined as the index date.

Criterion 3 was a combination of 9 diagnostic codes as primary diagnosis and 4 imaging procedure codes irrespective of admission of the patient. The coding date of imaging was defined as the index date. The diagnosis of VF should have been coded as primary diagnosis within 1 month before or after the coding of imaging procedure.

**Evaluation of the criterion-related validity**

Two evaluators (NH, an endocrinology specialist with 6 years of experience and SL, an endocrinology specialist with 3 years of experience) independently reviewed the medical records and radiographs of 2,819 patients who had at least one of the diagnostic codes for VF as primary or secondary diagnosis at the outpatient clinic or at admission between 1 January 2016 and 31 December 2016 in our tertiary institutions. Patients were diagnosed with incident osteoporotic VF when 1) radiographs showed a decrease in vertebral height > 25% according to semiquantitative Genant method, 2) the patient had developed acute back pain within the most recent 3 months, and 3) patients had no history of major trauma or falls greater than 2-meter based on medical records. Any disparity in the diagnosis was resolved by adjudication of third reviewer (YR, an endocrinology specialist with 20 years of experience).
Statistical analysis
We evaluated the sensitivity and positive predictive value (PPV) of the operational definition in detecting true osteoporotic VF and obtained the receiver operating characteristic (ROC) curve. All analyses were performed using STATA v16.0 (Stata Corp., College Station, TX, USA).

Ethics statement
The Institutional Review Board (IRB) of our hospital approved this study (IRB number: 4-2021-1531). The requirement for informed consent was waived by the review board due to retrospective nature of the study design.

RESULTS
Demographic data
Among the 2,819 patients who had primary or secondary diagnosis codes for VF, 995 patients satisfied at least one of the criteria for the operational definition of osteoporotic VF. There was a female preponderance (651/955, 65.43%), and the mean patient age was 65.27 ± 16.71 years (Table 2). Of these patients, 594 were judged as having true fractures based on medical records and radiographic examinations; furthermore, individuals who satisfied operational definition for osteoporotic VF had higher prevalence of prior VF compared to those without.

Validity of the 3 operational definitions
Criteria 1 and 2 identified 5.05% and 1.01% of the patients with VF, respectively, while criterion 3 identified 53.20% of patients with VF (316/594) (Fig. 1).

The sensitivity and PPV of the operational definition of incident VF were 62.5 (95% confidence interval [CI], 59.4–65.6) and 59.7 (95% CI, 56.6–62.8) respectively (Table 3). Area under the ROC curve for operational definition was 0.706 (95% CI, 0.688–0.724).

DISCUSSION
This study demonstrated the diagnostic validity of three operational definitions of VFs that were developed to identify osteoporosis-related VFs from the HIRA claims database. The VF

**Table 2. Patient demographics and clinical characteristics of the study population**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Total (n = 2,819)</th>
<th>Incident VF by operational definition (n = 995)</th>
<th>Non-VF by operational definition (n = 1,824)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, yr</td>
<td>65.40 ± 16.76</td>
<td>65.27 ± 16.71</td>
<td>65.48 ± 16.80</td>
<td>0.747</td>
</tr>
<tr>
<td>Sex, female</td>
<td>1871 (66.37)</td>
<td>651 (65.43)</td>
<td>1,220 (66.89)</td>
<td>0.433</td>
</tr>
<tr>
<td>Prior diagnosis codes of vertebral fracture</td>
<td>1,121 (39.8)</td>
<td>440 (44.22)</td>
<td>681 (37.34)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>CKD</td>
<td>172 (6.10)</td>
<td>60 (6.03)</td>
<td>112 (6.14)</td>
<td>0.907</td>
</tr>
<tr>
<td>COPD</td>
<td>67 (2.38)</td>
<td>25 (2.51)</td>
<td>42 (2.30)</td>
<td>0.727</td>
</tr>
<tr>
<td>Heart failure</td>
<td>109 (3.87)</td>
<td>32 (3.22)</td>
<td>77 (4.22)</td>
<td>0.186</td>
</tr>
<tr>
<td>RA</td>
<td>58 (2.06)</td>
<td>16 (1.61)</td>
<td>42 (2.30)</td>
<td>0.214</td>
</tr>
<tr>
<td>Diabetes</td>
<td>457 (16.21)</td>
<td>135 (13.57)</td>
<td>322 (17.65)</td>
<td>0.005</td>
</tr>
<tr>
<td>Hypertension</td>
<td>985 (34.94)</td>
<td>293 (29.45)</td>
<td>692 (37.94)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Atrial fibrillation</td>
<td>144 (6.85)</td>
<td>38 (6.70)</td>
<td>76 (6.92)</td>
<td>0.870</td>
</tr>
<tr>
<td>Any malignancy</td>
<td>290 (10.29)</td>
<td>88 (8.84)</td>
<td>202 (11.07)</td>
<td>0.063</td>
</tr>
<tr>
<td>True incident VF cases</td>
<td>950 (33.68)</td>
<td>594 (59.70)</td>
<td>356 (19.49)</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

Numeric parameters are expressed as mean ± standard deviation in parentheses. Categorical parameters are expressed as counts and percentages in parentheses. CKD = chronic kidney disease, COPD = chronic obstructive pulmonary disease, RA = rheumatoid arthritis, VF = vertebral fracture.
Operational definitions were based on a combination of diagnosis, procedure and imaging codes from the medical records. And we demonstrated these operational definitions based on medical records and radiographs. Our operational definition identified patients with osteoporosis-related VF with a high accuracy by combining diagnosis, procedure, and imaging codes.

In our country, the diagnosis code of claim data is assigned according to the importance of treatment or examination during the treatment period. A primary diagnosis is the condition that consumes the most medical resources, especially for newly diagnosed diseases. Secondary, or subordinated, diagnostic codes imply the use of fewer medical resources than those needed for primary diagnoses. For this reason, the order of diagnostic codes is often used as a criterion for developing operational definitions using claims data, and classifying the codes as primary or secondary. Additionally, imaging and procedure codes were used along with diagnostic codes to increase the accuracy of the operational definitions.

Here, we analyzed the operational definition developed using these codes.

To date, various operational definitions have been proposed for the identification of patients with osteoporotic fractures including VF, hip fracture and distal radius fracture. Unlike other osteoporotic fractures, identification of patients with osteoporosis-related VFs in claims data is challenging because the claims database does not include data on injury mechanism or bone mineral density. Moreover, the patient’s condition often improves with simple conservative treatments, such as prescription pain relievers, precluding the need for specialty treatments such as vertebroplasty or fusion surgery. Moreover, multiple diagnostic codes assigned during follow-up evaluations after the VF are problematic in accurate diagnostic identification though medical records.

Table 3. Diagnostic performance of operational definitions for VF

<table>
<thead>
<tr>
<th>No. of VF case</th>
<th>No. of TP</th>
<th>No. of FN</th>
<th>No. of FP</th>
<th>No. of TN</th>
<th>Sensitivity, % (95% CI)</th>
<th>PPV, % (95% CI)</th>
<th>AUC (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>950</td>
<td>594</td>
<td>356</td>
<td>401</td>
<td>1,471</td>
<td>62.5 (59.4–65.6)</td>
<td>59.7 (56.6–62.8)</td>
<td>0.706 (0.688–0.724)</td>
</tr>
</tbody>
</table>

VF = vertebral fracture, TP = true positive, FN = false negative, FP = false positive, TN = true negative, PPV = positive predictive value, AUC = area under curve.
Most osteoporotic VFs are treated conservatively, and there is a wide variation in the therapeutic options used and treatment lengths followed by different physicians. Therefore, previous claims database studies, which mainly used diagnosis codes and/or procedure codes, reported inconsistent incidence rates of osteoporotic VFs.\textsuperscript{18,28,35} Thus, diagnostic, imaging, and procedural codes should be combined to enable accurate identification of cases of osteoporotic VF. Furthermore, prior medical history of VF leads to further inconsistency between the true VF rate and the rate of VFs detected by currently used operational definitions; in the currently available literature, most (99.9%) patients who did not have incident VF but had primary or subordinated codes of VF had a previous history of VF (Table 1). In a previous study, the PPV of incident VF using only the diagnostic code was relatively low by 46%; however, it slightly increased to 61% in this study owing to the inclusion of imaging codes.\textsuperscript{39}

There were some limitations to this study. First, we sourced our data from a tertiary referral center, and a substantial number of the selected patients had comorbidities. If patient populations from other institutions were studied, the results could be inconsistent. Therefore, a multi-institutional study that combines a larger population data set from various databases should be conducted in the future to corroborate our results. Second, our operational definitions cannot be generalized to other countries that use different coding systems. Hence, although further validation is needed to translate our approach to other coding systems and study the populations of different countries, our results may serve as a foundation for such studies. Third, the PPV of our operational definition for VF was lower than those reported for hip fracture (59.1–77.4\%)\textsuperscript{41} and wrist fracture (95.6–98.2\%).\textsuperscript{40} Most patients with VF are treated conservatively and visit the outpatient clinic several times during the treatment course; this unique clinical feature may account for this difference.

Our findings demonstrate the validity of our operational definitions to identify VFs more accurately using claims data. This algorithm to identify VF is likely to be useful in future studies for diagnosing osteoporotic VF. Although the use of three operational definitions could misclassify incident VF, these results may be useful for future observational studies. Additional validation studies that can reduce the effect of confounding factors with improved algorithms and larger datasets should be conducted to corroborate and expand on our findings.

REFERENCES

2. Cummings SR, Melton LJ. Epidemiology and outcomes of osteoporotic fractures. Lancet 2002;359(9319):1761-7. \textsuperscript{PubMed | CrossRef}
   PUBLMED | CROSSREF

   PUBLMED | CROSSREF

   PUBLMED | CROSSREF

   PUBLMED | CROSSREF

   PUBLMED | CROSSREF

   PUBLMED | CROSSREF

   PUBLMED | CROSSREF

   PUBLMED | CROSSREF

   PUBLMED | CROSSREF

   PUBLMED | CROSSREF

   PUBLMED | CROSSREF

   PUBLMED | CROSSREF

   PUBLMED | CROSSREF

   PUBLMED | CROSSREF

   PUBLMED | CROSSREF

   PUBLMED | CROSSREF

   PUBLMED | CROSSREF

   PUBLMED | CROSSREF


