Bone 150 (2021) 116017

Contents lists available at ScienceDirect

Bone

journal homepage: www.elsevier.com/locate/bone

Full Length Article

Feasibility of a generalized convolutional neural network for automated identification of vertebral compression fractures: The Manitoba Bone Mineral Density Registry

Barret A. Monchka [a], Douglas Kimelman [b], Lisa M. Lix [a], William D. Leslie [b,c,*]

[a] Department of Community Health Sciences, University of Manitoba, Winnipeg, Manitoba, Canada
[b] Department of Radiology, University of Manitoba, Winnipeg, Manitoba, Canada
[c] Department of Internal Medicine, University of Manitoba, Winnipeg, Manitoba, Canada

ARTICLE INFO

Keywords:
Osteoporosis
Fracture
Dual-energy x-ray absorptiometry
Vertebral fracture assessment
Artificial intelligence
Machine learning

ABSTRACT

Background: Vertebral fracture assessment (VFA) images are acquired in dual-energy (DE) or single-energy (SE) scan modes. Automated identification of vertebral compression fractures, from VFA images acquired using GE Healthcare scanners in DE mode, has achieved high accuracy through the use of convolutional neural networks (CNNs). Due to differences between DE and SE images, it is uncertain whether CNNs trained on one scan mode will generalize to the other.

Purpose: To evaluate the ability of CNNs to generalize between GE DE and GE SE VFA scan modes.

Methods: 12,742 GE VFA images from the Manitoba Bone Mineral Density Program, obtained between 2010 and 2017, were exported in both DE and SE modes. VFAs were classified by imaging specialists as fracture present or absent using the modified algorithm-based qualitative (mABQ) method. VFA scans were randomly divided into independent training (60%), validation (10%), and test (30%) sets. Three CNN models were constructed by training separately on DE only, SE only, and a composite dataset comprised of both SE and DE VFAs. All three trained CNN models were separately evaluated against both SE and DE test datasets.

Results: Good performance was seen for CNNs trained and evaluated on the same scan mode. DE scans used for both training and evaluation (DE/DE) achieved 87.9% sensitivity, 87.4% specificity, and an area under the receiver operating characteristic curve (AUC) of 0.94. SE scans used for both training and evaluation (SE/SE) achieved 78.6% sensitivity, 90.6% specificity, AUC = 0.92. Conversely, CNNs performed poorly when evaluated on scan modes that differed from their training sets (AUC = 0.58). However, a composite CNN trained simultaneously on both SE and DE VFAs gave performance comparable to DE/DE (82.4% sensitivity, 94.3% specificity, AUC = 0.95); and provided improved performance over SE/SE (82.2% sensitivity, 92.3% specificity, AUC = 0.94). Positive predictive value was higher with the composite CNN compared with models trained solely on DE (74.5% vs. 58.7%) or SE VFAs (68.6% vs. 62.9%).

Conclusion: CNNs for vertebral fracture identification are highly sensitive to scan mode. Training CNNs on a composite dataset, comprised of both GE DE and GE SE VFAs, allows CNNs to generalize to both scan modes and may facilitate the development of manufacturer-independent machine learning models for vertebral fracture detection.

1. Introduction

Vertebral compression fractures are the most common osteoporosis-related fractures [1,2] and are associated with adverse health outcomes including chronic pain, disability, reduced quality of life, and mortality [3–6]. Substantial healthcare utilization has been attributed to vertebral fractures [7,8]; and as populations continue to age, vertebral fractures will place an increasing burden on healthcare systems. Prevalent vertebral fractures are well-established predictors of incident vertebral, hip, and other non-vertebral fractures [9–11]. Despite the importance of fracture detection, vertebral fracture diagnosis is prone to low interrater reliability and the majority of vertebral fractures are asymptomatic and

* Corresponding author at: Department of Medicine (C5121), 409 Tache Avenue, Winnipeg, Manitoba R2H 2A6, Canada.
E-mail address: bleslie@sbgh.mb.ca (W.D. Leslie).

https://doi.org/10.1016/j.bone.2021.116017
Received 18 January 2021; Received in revised form 3 May 2021; Accepted 16 May 2021
Available online 19 May 2021
8756-3282/© 2021 Elsevier Inc. All rights reserved.

undiagnosed [12,13]—resulting in failure to institute secondary prevention. Vertebral fracture assessment (VFA), a low-radiation procedure on DXA bone densitometers to acquire lateral images of the thoracolumbar spine, is widely recommended at the same time as bone mineral density (BMD) measurement to identify undiagnosed vertebral fractures [14–18].

Automated identification of vertebral compression fractures in VFA images is possible through the use of convolutional neural networks (CNNs)—specialized machine learning models for image analysis [19]. Integrating automated vertebral fracture detection techniques into clinical workflows has the potential to reduce vertebral fracture misdiagnoses and improve agreement among VFA reviewers, promoting timely patient care [19,20]. Artificial intelligence (AI) was previously evaluated as a clinical-aid and found to improve fracture detection in standard radiographs among emergency medicine physicians [21]. Despite the potential for AI-assisted fracture detection to improve diagnostic consistency leading to improved patient outcomes, limited research has been conducted to investigate the feasibility of automated vertebral fracture detection with VFA scans.

The two major manufacturers of VFA imaging equipment, GE Healthcare and Hologic, employ different VFA techniques: Hologic scanners acquire in single-energy (SE) mode and display SE images, while GE scanners acquire in dual-energy (DE) mode but display either DE or SE images, with DE images being the default. Considering the differences in imaging techniques between GE and Hologic devices, machine learning models that perform well on both DE and SE VFAs are needed for AI-assisted vertebral fracture identification to be widely incorporated into clinical workflows. However, due to differences between DE and SE VFA images, CNNs trained on one scan mode may not generalize well to the other. The purpose of this study was to evaluate the ability of CNNs to generalize between GE DE and GE SE VFA scan modes in order to facilitate the development of manufacturer-independent machine learning models for vertebral fracture detection.

2. Materials and methods

2.1. Study design, cohort, and data source

We conducted a retrospective cohort study using de-identified data from the Manitoba Bone Mineral Density Registry [22]. The Registry contains records of all bone densitometry services provided in Manitoba, Canada; and all VFAs performed since 2010 [22]. The study cohort has been previously described and was comprised of all Manitoba residents who underwent a VFA at the time of bone densitometry between February 2010 and December 2017 [19]. Indications for VFA were assessed by DXA technologists and were specified as having a T-score ≤ -1.5 (measured at the femoral neck, lumbar spine, or total hip); and one of (a) age ≥ 70 years, or (b) age 50–69 years and historical height loss $>$ 5 cm, or (c) age 50–69 years and measured height loss $>$ 2.5 cm, or (d) age 50–69 years and $\geq$7.5 mg daily glucocorticoid steroid use for at least three months in the past year [10,19]. VFAs were excluded if a vertebral fracture was considered uncertain, traumatic or pathological fractures were present, or VFA image quality or resolution was deemed insufficient [19]. The Health Research Ethics Board for the University of Manitoba approved this study.

2.2. Vertebral fracture assessment

Lateral VFA images of the thoracolumbar spine (T4–L4) were captured using GE first-generation (Prodigy) and second-generation (iDXA) scanners, and exported as bone maps (dual-energy) and low-energy maps (single-energy). Each VFA was originally interpreted by one of four expert physician readers and labeled as fracture present or absent (i.e., binary classification) using the modified algorithm-based qualitative (mABQ) method as previously described [10,19]. The VFA was typically reviewed by a second physician at the time that the final DXA report was available for approval. Review of ancillary imaging through the province-wide picture archiving and communication system (PACS) and consultation between physicians was freely performed to ensure diagnostic accuracy. VFA reviewers were diagnostic imaging specialists, each having between 10 and 31 years of DXA experience, and certified by the International Society for Clinical Densitometry [19]. We have previously shown that the vertebral fracture labeling in this dataset is strongly predictive of fracture outcomes (including non-vertebral and hip fracture) [10].

2.3. Image pre-processing

Exported VFAs were converted to Portable Network Graphics (PNG) format, and resized to 600×360 pixels. To account for the variability in VFA image proportions and to maintain a consistent width, images were left and right padded with black borders. Contrast was enhanced by narrowing the pixel intensity range; and images were normalized by rescaling pixel values to [0, 1]. Dual-energy and single-energy images were pre-processed in a consistent manner.

VFA scans were randomly divided into independent training (60%), validation (10%), and test (30%) sets. Oversampling was used to balance training classes: VFAs containing fractures were copied five times, for a total of six instances each; then, random sampling without replacement was used to select additional fracture images to be copied one more time—until class sizes were equal. To form the composite dataset, the dual-energy and single-energy image sets were separately balanced prior to combining.

2.4. Image analysis

Using supervised machine learning, three CNN models were constructed by training separately on DE only (GE Prodigy and iDXA combined), SE only, and a composite dataset comprised of both DE and SE VFAs. All three trained CNNs were separately evaluated against both DE and SE scans (six combinations). For each training/evaluation pair, the CNN architecture that performed best against the validation set was chosen. Architectures considered include Inception-ResNet-v2 [23], DenseNet [24], and ensembles formed by taking the maximum and mean predicted probabilities of the two models. Analysis was performed using Python (v3.5.2), and the Keras (v2.2.0) and TensorFlow (v1.9.0) machine learning libraries. Final training was performed on the entire development set (training + validation), before evaluating performance against the test set.

Hyperparameters were chosen based on predictive performance against the validation set. Model weights were initialized using the Glorot uniform scheme. Optimization was performed using Adam stochastic gradient descent with a binary cross-entropy loss function, and a cosine-based annealing learning rate scheduler (initial learning rate = 10^{-4}). Training was conducted over 10 epochs using a batch size of 12, with random input shuffling and data augmentation. Random transformations applied during training, to minimize overfitting, include rotation [-30°, $+30^\circ$], shear [-20°, $+20^\circ$], width and height shift [−20%, +20%], brightness [−10%, +10%], and zoom [−10%, +10%]. VFA images were classified as fracture or non-fracture using sigmoid activation with a cut-off of 0.5.

2.5. Statistical analysis

We measured CNN performance using accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), balanced accuracy (arithmetic mean of sensitivity and specificity), F1-score (harmonic mean of sensitivity and PPV), and area under the receiver operating characteristic curve (AUC). Bootstrapping with 1000 samples was used to construct 95% confidence intervals (CIs).

3. Results

3.1. Baseline characteristics

A total of 12,742 VFA scans were included in this study, 1356 (10.6%) of which were same patient repeats. Among the 13,395 VFAs performed between 2010 and 2017, 653 (4.9%) scans were excluded: vertebral fracture considered uncertain by an expert reader (N = 393), traumatic or pathological fracture present (N = 5), VFA image quality or resolution deemed insufficient (N = 255) [19]. The development set, which was used for model training and hyperparameter optimization, was comprised of 8920 scans (70%); while the test set accounted for 3822 (30%) of the included VFAs.

The baseline characteristics of patients with a VFA were distributed similarly between the development and test sets (Table 1). Out of 12,742 VFAs included in this study, 16.6% (N = 2116) were diagnosed with a vertebral fracture by the imaging specialist. The mean age (standard deviation) was 75.8 years (± 6.9); and 94.3% of patients were female (N = 12,013). GE second-generation (iDXA) imaging equipment was used to capture 68.1% of scans; while the remaining VFAs (31.9%) were performed using GE first-generation (Prodigy) scanners.

3.2. Predictive performance

Good performance was seen for CNNs trained and evaluated on the same scan mode (Table 2). SE scans used for both training and evaluation (SE/SE) achieved a balanced accuracy (arithmetic mean of sensitivity and specificity) of 84.6% (95% CI: 83.1–86.2%), an F1-score (harmonic mean of sensitivity and PPV) of 69.9% (95% CI: 67.4–72.5%), and AUC = 0.92 (95% CI: 0.91–0.93%). Compared with SE/SE, higher predictive performance was observed when DE scans were used for both training and evaluation (DE/DE): balanced accuracy = 87.7% (95% CI: 86.2–89.0%), F1-score = 78.6% (95% CI: 76.1–80.9%), AUC = 0.94 (95% CI: 0.93–0.95%). Conversely, CNNs performed poorly when evaluated on scan modes that differed from their training sets: The DE-trained model evaluated against SE VFAs obtained balanced accuracy = 55.2%, AUC = 0.58; while the SE-trained model evaluated against DE VFAs attained balanced accuracy = 54.4%, AUC = 0.58.

A composite model, simultaneously trained on all SE and DE scans (N = 17,840), gave performance comparable to DE/DE and provided improved performance over SE/SE. When evaluated against DE scans, the composite CNN achieved balanced accuracy = 88.3% (95% CI: 86.8–89.9%), F1-score = 78.6% (95% CI: 76.3–80.9%), AUC = 0.95 (95% CI: 0.94–0.96); while evaluating the composite model against SE produced balanced accuracy = 87.3% (95% CI: 85.7–89.9%), F1-score = 74.8% (95% CI: 72.3–77.2%), AUC = 0.94 (95% CI: 0.93–0.95). The composite model produced positive predictive values noticeably better than models trained and evaluated on the same scan mode: 74.5% (95% CI: 71.3–77.8%) compared with 58.7% (95% CI: 55.6–61.6%) from the DE model, and 68.6% (95% CI: 65.1–71.7%) compared with 62.9% (95% CI: 59.6–66.1%) from the SE model. NPV was comparable between the composite model and models trained and evaluated on the same scan mode. Class activation heatmaps (Fig. 1) demonstrate compatibility between localized fracture predictions made by the composite CNN and the human reference.

For different training/evaluation pairs, certain CNN architectures performed better than the others during validation and final evaluation. Inception-ResNet-v2 performed best for the composite model and SE/SE. Taking the maximum predicted probability of a DenseNet-Inception-ResNet-v2 ensemble produced the best results for DE/DE and the SE-trained model evaluated against DE, while a mean ensemble performed best for the CNN trained on DE and evaluated against SE images.

Table 1
Characteristics of patients undergoing a vertebral fracture assessment by image set assignment.

Characteristic	Image set		Overall
	Development (training + validation)	Test	
No. of VFAs	8920 (70.0%)	3822 (30.0%)	12,742
Age (years)	75.9 ± 6.9	75.7 ± 7.2	75.8 ± 6.9
Sex			
Female	8422 (94.4%)	3591 (94.0%)	12,013 (94.3%)
Male	498 (5.6%)	231 (6.0%)	729 (5.7%)
BMI (kg/m^2)	26.1 ± 5.0	26.1 ± 5.1	26.1 ± 5.0
Lumbar spine T-score	−1.9 ± 1.2	−1.9 ± 1.2	−1.9 ± 1.2
DXA scanner type			
GE Prodigy (1st generation)	2807 (31.5%)	1255 (32.8%)	4062 (31.9%)
GE iDXA (2nd generation)	6113 (68.5%)	2567 (67.2%)	8680 (68.1%)
Vertebral fracture diagnosis	1470 (16.5%)	646 (16.9%)	2116 (16.6%)

Data are presented as mean ± standard deviation, or N (%). VFA = vertebral fracture assessment; BMI = body mass index; DXA = dual-energy x-ray absorptiometry.

4. Discussion

We assessed the ability of CNNs to generalize between GE DE and GE SE VFAs and found CNNs are highly sensitive to scan mode and have poor predictive performance when evaluated on scan modes that differ from those in their training set. However, we found that simultaneously training on both GE dual- and GE single-energy images allows CNNs to generalize to both scan modes and improves the identification of vertebral fractures in GE SE VFAs. Incorporating automated vertebral fracture detection into clinical workflows could reduce misdiagnoses, improve fracture management and prevention, and could ultimately lead to improved patient outcomes. For automated fracture detection to be broadly available to assist clinical decision making, machine learning models must have good performance against both SE and DE VFAs, and this study is a first step towards the development of manufacturer-independent models that perform well on both scan modes.

We observed slightly better performance for detecting vertebral fractures in DE scans compared with SE, whether contrasting DE/DE with SE/SE, or comparing evaluations of the composite model against DE and SE. This may relate to the more prominent soft tissue effects on SE scans. Compared with SE, DE scans provide greater contrast between the vertebral bodies and surrounding image features. These differences may contribute to the poor performance observed when evaluating CNNs on scan modes that differ from their training set. Simultaneously training on both scan modes may help CNNs learn to focus on the vertebral bodies and ignore image features irrelevant to vertebral fracture identification. Our results suggest that incorporating DE scans in the training process allows CNNs to partly overcome the challenges inherent in SE image analysis.

Limited research has been conducted to automate vertebral fracture detection in VFA images using machine learning methods. Derkatch et al. used CNNs to identify vertebral fractures in the same cohort used by this study, but limited their analysis to DE VFAs [19]. Our DE-trained model had comparable predictive performance with Derkatch et al.; while our composite CNN produced similar AUC but with slightly lower sensitivity and higher specificity, whether evaluated against SE or DE VFAs. To the best of our knowledge, no previous study has automated the detection of vertebral compression fractures in SE VFAs, nor have any previous studies constructed a machine learning model for vertebral fracture detection that generalizes to both dual- and single-energy VFAs.

We acknowledge several strengths and weaknesses. The machine learning models developed in this study allow for fully automated vertebral fracture identification in VFA images, without requiring any prior feature engineering or image segmentation. This study was conducted using a relatively large number of VFA scans collected over a 7

Table 2
Convolutional neural network performance for vertebral fracture identification.

Evaluation dataset	Training dataset	Accuracy (%)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	AUC
Single-energy	Single-energy	88.6 (87.5–89.5)	78.6 (75.6–81.6)	90.6 (89.5–91.5)	62.9 (59.6–66.1)	95.4 (94.7–96.1)	0.92 (0.91–0.93)
	Composite	90.6 (89.7–91.5)	82.2 (79.3–85.4)	92.3 (91.4–93.3)	68.6 (65.1–71.7)	96.2 (95.6–96.9)	0.94 (0.93–0.95)
	Dual-energy	60.4 (58.9–61.9)	47.4 (43.4–51.4)	63.1 (61.5–64.7)	20.7 (18.6–22.8)	85.5 (84.0–86.9)	0.58 (0.55–0.60)
Dual-energy	Single-energy	70.4 (69.0–71.9)	30.3 (26.8–33.7)	78.6 (77.2–79.9)	22.3 (19.6–25.0)	84.7 (83.5–86.1)	0.58 (0.56–0.61)
	Composite	92.3 (91.4–93.1)	82.4 (79.3–85.3)	94.3 (93.4–95.1)	74.5 (71.3–77.8)	96.3 (95.7–97.0)	0.95 (0.94–0.96)
	Dual-energy	87.5 (86.5–88.6)	87.9 (85.2–90.4)	87.4 (86.3–88.5)	58.7 (55.6–61.6)	97.3 (96.6–97.9)	0.94 (0.93–0.95)

PPV = positive predictive value; NPV = negative predictive value; AUC = area under the receiver operating characteristic curve; Composite = dataset comprised of both single-energy and dual-energy vertebral fracture assessment images. 95% confidence intervals are shown in parentheses.

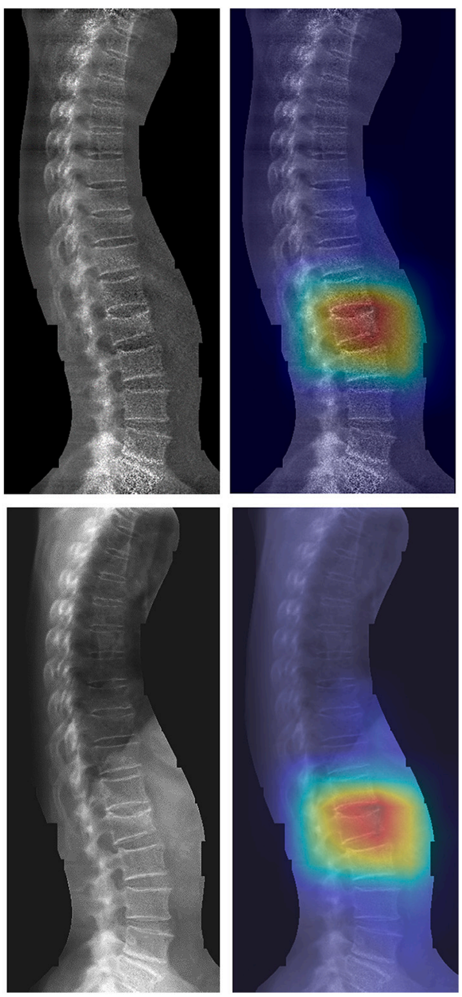

Fig. 1. Example vertebral fracture assessment exported as both dual-energy (top-left) and single-energy (bottom-left) showing a moderate vertebral compression fracture at L2. Class activation heatmaps (right) indicate image regions relevant to the composite convolutional neural network, trained simultaneously on dual- and single-energy images, for vertebral fracture identification.

year period, and labeled by diagnostic imaging specialists with a minimum of 10 years of DXA experience. Although the study cohort was disproportionally female, our data had excellent population coverage—including all Manitoba residents referred for VFA following bone densitometry. Predictive performance evaluation was limited to an internal test set, and external generalizability was not assessed against an independent dataset. This study was limited to SE images from a single scanner manufacturer (GE Healthcare). Further work is needed to directly assess fracture detection against Hologic SE VFAs, and further assess the feasibility of manufacturer-independent prediction models. Although it was not feasible in this study to assess the accuracy of the predicted fracture location, comparing the agreement between class activation heatmaps and human-annotated fracture location would provide additional insight into model performance and would be a valuable future contribution to this research area.

5. Conclusion

CNNs for vertebral fracture identification are highly sensitive to scan mode and do not generalize when trained on only one scan mode. Training CNNs on a composite dataset, comprised of both GE DE and GE SE VFAs, allows CNNs to generalize to both scan modes, improves fracture identification in GE SE VFA scans, and supports the feasibility of manufacturer-independent models for automated vertebral fracture detection.

CRediT authorship contribution statement

Barret A. Monchka: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization. **Douglas Kimelman:** Resources, Writing – review & editing, Visualization. **Lisa M. Lix:** Writing – review & editing, Visualization. **William D. Leslie:** Conceptualization, Formal analysis, Resources, Writing – review & editing, Supervision, Data curation, Project administration, Funding acquisition.

Acknowledgements

We acknowledge funding from the Visual and Automated Disease Analytics (VADA) Program to the first author; and the Rady Innovation Fund, Rady Faculty of Health Sciences, University of Manitoba. The authors acknowledge the Manitoba Centre for Health Policy for use of data contained in the Population Health Research Data Repository (HIPC 2016/2017-29). Results and conclusions are those of the authors and no official endorsement by the Manitoba Centre for Health Policy, Manitoba Health, Seniors and Active Living, or other data providers is intended or should be inferred. The authors also acknowledge valuable non-author contributions from Sheldon Derkatch, Chris Kirby, Steven Reda, Mohammad Jafari Jozani, and Michael Davidson.

References

[1] B.L. Riggs, L.J. Melton, The worldwide problem of osteoporosis: insights afforded by epidemiology, Bone. 17 (1995) S505–S511, https://doi.org/10.1016/8756-3282(95)00258-4.
[2] R. Burge, B. Dawson-Hughes, D.H. Solomon, J.B. Wong, A. King, A. Tosteson, Incidence and economic burden of osteoporosis-related fractures in the United States, 2005–2025, J. Bone Miner. Res. 22 (2007) 465–475, https://doi.org/10.1359/jbmr.061113.
[3] U.A. Al-Sari, J. Tobias, E. Clark, Health-related quality of life in older people with osteoporotic vertebral fractures: a systematic review and meta-analysis, Osteoporos. Int. 27 (2016) 2891–2900, https://doi.org/10.1007/s00198-016-3648-x.
[4] M.C. Nevitt, B. Ettinger, D.M. Black, K. Stone, S.A. Jamal, K. Ensrud, M. Segal, H. K. Genant, S.R. Cummings, The association of radiographically detected vertebral

fractures with back pain and function: a prospective study, Ann. Intern. Med. 128 (1998) 793–800, https://doi.org/10.7326/0003-4819-128-10-199805150-00001.
[5] H.A. Fink, K.E. Ensrud, D.B. Nelson, R.P. Kerani, P.J. Schreiner, Y. Zhao, S. R. Cummings, M.C. Nevitt, Disability after clinical fracture in postmenopausal women with low bone density: the fracture intervention trial (FIT), Osteoporos. Int. 14 (2003) 69–76, https://doi.org/10.1007/s00198-002-1314-y.
[6] C. Cooper, E.J. Atkinson, S.J. Jacobsen, W.M. O'Fallon, L.J. Melton, Population-based study of survival after osteoporotic fractures, Am. J. Epidemiol. 137 (1993) 1001–1005, https://doi.org/10.1093/oxfordjournals.aje.a116756.
[7] A.A. Baaj, K. Downes, A.R. Vaccaro, J.S. Uribe, F.L. Vale, Trends in the treatment of lumbar spine fractures in the United States: a socioeconomics perspective - clinical article, J. Neurosurg. Spine 15 (2011) 367–370, https://doi.org/10.3171/2011.5.SPINE10934.
[8] W.D. Leslie, L.M. Lix, G.S. Finlayson, C.J. Metge, S.N. Morin, S.R. Majumdar, Direct healthcare costs for 5 years post-fracture in Canada: a long-term population-based assessment, Osteoporos. Int. 24 (2013) 1697–1705, https://doi.org/10.1007/s00198-012-2232-2.
[9] D.M. Black, N.K. Arden, L. Palermo, J. Pearson, S.R. Cummings, Prevalent vertebral deformities predict hip fractures and new vertebral deformities but not wrist fractures, J. Bone Miner. Res. 14 (1999) 821–828, https://doi.org/10.1359/jbmr.1999.14.5.821.
[10] J.T. Schousboe, L.M. Lix, S.N. Morin, S. Derkatch, M. Bryanton, M. Alhrbi, W. D. Leslie, Prevalent vertebral fracture on bone density lateral spine (VFA) images in routine clinical practice predict incident fractures, Bone. 121 (2019) 72–79, https://doi.org/10.1016/j.bone.2019.01.009.
[11] L.J. Melton, E.J. Atkinson, C. Cooper, W.M. O'Fallon, B.L. Riggs, Vertebral fractures predict subsequent fractures, Osteoporos. Int. 10 (1999) 214–221, https://doi.org/10.1007/s001980050218.
[12] H.A. Fink, D.L. Milavetz, L. Palermo, M.C. Nevitt, J.A. Cauley, H.K. Genant, D. M. Black, K.E. Ensrud, What proportion of incident radiographic vertebral deformities is clinically diagnosed and vice versa? J. Bone Miner. Res. 20 (2005) 1216–1222, https://doi.org/10.1359/JBMR.050314.
[13] P.D. Delmas, L. De Van Langerijt, N.B. Watts, R. Eastell, H. Genant, A. Grauer, D. L. Cahall, Underdiagnosis of vertebral fractures is a worldwide problem: the IMPACT study, J. Bone Miner. Res. 20 (2005) 557–563, https://doi.org/10.1359/JBMR.041214.
[14] R.Y. van der Velde, S.P.G. Bours, C.E. Wyers, W.F. Lems, P.P.M.M. Geusens, J.P. W. van den Bergh, Effect of implementation of guidelines on assessment and diagnosis of vertebral fractures in patients older than 50 years with a recent non-vertebral fracture, Osteoporos. Int. 28 (2017) 3017–3022, https://doi.org/10.1007/s00198-017-4147-4.
[15] J. Compston, A. Cooper, C. Cooper, N. Gittoes, C. Gregson, N. Harvey, S. Hope, J. A. Kanis, E.V. McCloskey, K.E.S. Poole, D.M. Reid, P. Selby, F. Thompson, A. Thurston, N. Vine, UK clinical guideline for the prevention and treatment of osteoporosis, Arch. Osteoporos. 12 (2017) 43, https://doi.org/10.1007/s11657-017-0324-5.
[16] A. Papaioannou, S. Morin, A.M. Cheung, S. Atkinson, J.P. Brown, S. Feldman, D. A. Hanley, A. Hodsman, S.A. Jamal, S.M. Kaiser, B. Kvern, K. Siminoski, W. D. Leslie, 2010 clinical practice guidelines for the diagnosis and management of osteoporosis in Canada: summary, CMAJ. 182 (2010) 1864–1873, https://doi.org/10.1503/cmaj.100771.
[17] J. Yang, F. Cosman, P.W. Stone, M. Li, J.W. Nieves, Vertebral fracture assessment (VFA) for osteoporosis screening in US postmenopausal women: is it cost-effective? Osteoporos. Int. 31 (2020) 2321–2335, https://doi.org/10.1007/s00198-020-05588-6.
[18] P.M. Camacho, S.M. Petak, N. Binkley, D.L. Diab, L.S. Eldeiry, A. Farooki, S. T. Harris, D.L. Hurley, J. Kelly, E.M. Lewiecki, R. Pessah-Pollack, M. McClung, S. J. Wimalawansa, N.B. Watts, American Association of Clinical Endocrinologists/American College of Endocrinology clinical practice guidelines for the diagnosis and treatment of postmenopausal osteoporosis—2020 update, Endocr. Pract. 26 (2020) 564–570, https://doi.org/10.4158/GL-2020-0524.
[19] S. Derkatch, C. Kirby, D. Kimelman, M.J. Jozani, J. Michael Davidson, W.D. Leslie, Identification of vertebral fractures by convolutional neural networks to predict nonvertebral and hip fractures: a Registry-based Cohort Study of Dual X-ray Absorptiometry, Radiology. 293 (2019) 404–411, https://doi.org/10.1148/radiol.2019190201.
[20] J.T. Schousboe, L.M. Lix, S.N. Morin, S. Derkatch, M. Bryanton, M. Alhrbi, W. D. Leslie, Vertebral fracture assessment increases use of pharmacologic therapy for fracture prevention in clinical practice, J. Bone Miner. Res. 34 (2019) 2205–2212, https://doi.org/10.1002/jbmr.3836.
[21] R. Lindsey, A. Daluiski, S. Chopra, A. Lachapelle, M. Mozer, S. Sicular, D. Hanel, M. Gardner, A. Gupta, R. Hotchkiss, H. Potter, Deep neural network improves fracture detection by clinicians, Proc. Natl. Acad. Sci. U. S. A. 115 (2018) 11591–11596, https://doi.org/10.1073/pnas.1806905115.
[22] W.D. Leslie, P.A. Caetano, L.R. MacWilliam, G.S. Finlayson, Construction and validation of a population-based bone densitometry database, J. Clin. Densitom. 8 (2005) 25–30, https://doi.org/10.1385/JCD:8:1:025.
[23] C. Szegedy, S. Ioffe, V. Vanhoucke, A.A. Alemi, Inception-v4, inception-ResNet and the impact of residual connections on learning, in: 31st AAAI Conf. Artif. Intell. AAAI 2017, AAAI Press, 2017, pp. 4278–4284. https://arxiv.org/abs/1602.07261v2. (Accessed 13 November 2020).
[24] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely connected convolutional networks, in: Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017, Institute of Electrical and Electronics Engineers Inc., 2017, pp. 2261–2269, https://doi.org/10.1109/CVPR.2017.243.