European Radiology (2021) 31:1558–1568
https://doi.org/10.1007/s00330-020-07236-3

MUSCULOSKELETAL

Diagnostic accuracy of dual-energy CT for the detection of bone marrow edema in the appendicular skeleton: a systematic review and meta-analysis

Mitchell P. Wilson[1] · Kevin Lui[1] · Dorian Nobbee[1] · Mohammad H. Murad[2] · Matthew D. F. McInnes[3] · Trevor A. McGrath[4] · Prayash Katlariwala[1] · Gavin Low[1]

Received: 8 April 2020 / Revised: 12 July 2020 / Accepted: 27 August 2020 / Published online: 8 September 2020

Abstract

Objectives This meta-analysis evaluated the diagnostic accuracy of dual-energy CT (DECT) for detecting bone marrow edema (BME) in the appendicular skeleton.

Methods A systematic review of MEDLINE, EMBASE, Scopus, the Cochrane Library, and gray literature from inception through January 31, 2020, was performed. Original articles with > 10 patients evaluating the accuracy of DECT for detecting BME in the appendicular skeleton with a reference standard of MRI and/or clinical follow-up were included. Study details were independently extracted by two reviewers. Meta-analysis was performed using a bivariate random-effects model with multivariable meta-regression. Risk of bias (RoB) was evaluated with QUADAS-2.

Results Twenty studies evaluating 790 patients for BME in the appendicular skeleton were included in analysis. The summary sensitivity, specificity, and AUC values for BME in the appendicular skeleton were 86% (95% confidence interval [CI] 82–89%), 93% (95% CI 90–95%), and 0.95, respectively. Quantitative analysis had a higher sensitivity than qualitative analysis on meta-regression ($p = 0.01$), but no difference in specificity ($p = 0.28$). No other covariates contributed to variability in accuracy (etiology (trauma vs non-trauma); location (upper vs lower extremity); and RoB). Studies demonstrated generally low or unclear RoB and applicability. Eight studies included index tests with high RoB when quantitative assessments used a retrospective cut-off value.

Conclusions DECT demonstrates a higher specificity than sensitivity and AUC > 0.9. In scenarios where MRI availability is limited or contraindicated, DECT could be an alternative to MRI for detecting BME in the appendicular skeleton. However, limitations in sources of variability and RoB warrant continued study.

Key Points

- *Twenty studies evaluating 790 patients for bone marrow edema in the appendicular skeleton with dual-energy CT were included in analysis.*
- *The summary sensitivity, specificity, and AUC values for detecting bone marrow in the appendicular skeleton were 86% (95% confidence interval [CI] 82–89%), 93% (95% CI 90–95%), and 0.95, respectively.*
- *In scenarios where MRI availability is limited or is contraindicated, DECT could be an alternative to MRI for detecting bone marrow edema in the appendicular skeleton.*

Keywords Radiography, dual-energy scanned projection · Tomography, X-ray computed · Bone marrow · Bones of upper extremity · Bones of lower extremity

✉ Mitchell P. Wilson
mitch.wilson@ualberta.ca

[1] Department of Radiology and Diagnostic Imaging, University of Alberta, 2B2.41 WMC, 8440-112 Street NW, Edmonton, Alberta T6G 2B7, Canada

[2] Evidence-Based Practice Center, Mayo Clinic, Room 2-54, 205 3rd Ave SW, Rochester, MN 55905, USA

[3] Departments of Radiology and Epidemiology, University of Ottawa/The Ottawa Hospital Research Institute, 501 Smyth Road, Box 232, Ottawa, Ontario K1H-8L6, Canada

[4] Department of Radiology, University of Ottawa, 501 Smyth Road, Box 232, Ottawa, Ontario K1H-8L6, Canada

Abbreviations

BME	Bone marrow edema
DECT	Dual-energy computed tomography
FN	False negative
FN	False positive
MRI	Magnetic resonance imaging
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analysis
QUADAS-2	Quality Assessment of Diagnostic Accuracy Studies-2
ROC	Receiver operator characteristic
TN	True negative
TP	True positive

Introduction

Bone marrow edema (BME) is pathologically associated with microfracture and hemorrhage in the trabecular bone, serving as a biomarker for bone injury [1]. The detection of BME can further elucidate the mechanism and extent of occult bone injuries, while also assisting with management, follow-up, and prognostication [1]. Preferable contrast resolution for BME with MRI has defined MRI as the standard of care for BME detection [2]. However, concerns surrounding both cost and delays in diagnosis related to MR access and safety have been arguments for alternative diagnostic options [3]. Advances in dual-energy computed tomography (DECT) technology including three-material decomposition techniques, which have allowed for the digital subtractions of materials with relevant photoelectric effect such as iodine and calcium, are establishing DECT as a potential imaging alternative to MRI when evaluating for BME [4–6].

Recent meta-analyses show DECT has a high diagnostic accuracy for detecting BME, particularly in the axial skeleton [4–6]. As such, DECT is becoming a recommended imaging modality for BME. However, location is a significant cause for variability with lower accuracies identified in joints of the appendicular skeleton [4]. The purpose of this systematic review and meta-analysis was to evaluate the diagnostic accuracy of DECT for detecting BME in the appendicular skeleton.

Methods

This systematic review and meta-analysis was performed according to current best practices and reported using guidance from the Preferred Reporting Items for Systematic Reviews and Meta-Analysis – Diagnostic Test Accuracy (PRISMA-DTA) guidelines [7–9]. A pre-established protocol was designed and submitted to the PROSPERO database prior to the initiation of the review (CRD42020168477). This analysis was exempt from ethical approval as it included only de-identified data with individual studies acquiring ethical approval from their home institution where necessary.

Literature search

A search of Ovid MEDLINE, Ovid EMBASE, Scopus, and the Cochrane Library (including the Cochrane Central Register of Controlled Trials, Cochrane Central Register of Protocols, and the Cochrane Database of Systematic Reviews) was performed from inception to January 31, 2020, to identify studies which used DECT for the detection of bone marrow edema in the appendicular skeleton. Variations of title/abstract/keywords and medical subject heading terms including "dual energy" AND "computed tomography" AND "bone" AND "edema" were modified dependent on database. Individual search strategies by database are outlined in Appendix Table 6. No language restrictions were applied and translation was performed when required. The search was completed according to best practices for electronic search strategies [10]. Articles from each database were then combined and duplicate articles were removed from the list. Titles and abstracts were screened for relevance and full-text review for potentially relevant studies was then performed by two reviewers with 0 and 6 years of experience with musculoskeletal imaging (K.L. and M.P.W.). Discrepancies in both processes were re-reviewed with consensus achieved between reviewers. A gray literature search was also performed by one author (M.P.W.), evaluating the most recent 2 years of conference proceedings from the Radiological Society of North America (RSNA) and the European Congress of Radiology (ECR) in addition to the most recent annual meeting for the American Roentgen Ray Society (ARRS). Conference abstracts which satisfactorily met the inclusion criteria and were not subsequently published were included in analysis. Finally, reference lists for key studies were checked and forward-searching of these key studies was performed in Google Scholar.

Selection criteria

All original articles evaluating the diagnostic accuracy of DECT for the detection of bone marrow edema in the appendicular skeleton compared with an MRI and/or clinical outcome reference standard were evaluated with full-text review. The sacroiliac joint was included as part of the appendicular skeleton if edema was evaluated in the iliac wing. Studies required sufficient data to reconstruct a 2×2 contingency table and authors were contacted via email when insufficient information was available. Studies were then included if sufficient information was provided by the corresponding author.

Based on a pre-established protocol, studies were excluded from the analysis if (1) less than 10 lesions were included in the assessment, (2) the study did not use DECT as the index test, (3) the target condition was not bone marrow edema, (4) the study evaluated bone marrow edema in the axial skeleton (including the skull, spine, sacrum, and/or coccyx), (5) studies and/or regions without bone marrow edema were not included, (5) a reference standard other than MRI and/or clinical outcome was used, or (6) the article was non-original research (including review articles, guidelines, consensus statements, and letters to the editor).

Data extraction

The relevant data from included studies was extracted independently by two separate reviewers. One reviewer with 6 years of experience in musculoskeletal imaging (M.P.W.) evaluated all included studies. Two additional reviewers with 0 and 6 years of experience in musculoskeletal imaging (K.L. and D.N.) each reviewed half of the included studies independent from the first reviewer. The datasets were then compared and discrepancies re-reviewed by consensus between the two more experienced reviewers (M.P.W. and D.N.). Patient characteristics were recorded including the total number of patients, mean age, number and percentage of male sex, total number of patients with bone edema (or total number of regions with edema where bones were segmentally evaluated), site of pathology, and study indication. Study characteristics included first author, year of publication, country of publication, prospective versus retrospective study design, single versus multicenter data acquisition, reported reference standard, time between trauma and DECT (days), time between DECT and MRI (days), number of readers, presence of consensus reading, and reader experience were recorded. Details of DECT characteristics including brand, low and high tube voltages (kV), pre-defined tube current-time product(s) (mAs), use of a tin filter, collimation (mm), rotation time (seconds), slice width (mm), reconstruction kernel(s), post-processing imaging, and evaluation method were recorded when reported in the primary study. Finally, true positive (TP), false negative (FN), true negative (TN), false positive (FN), sensitivity, specificity, and accuracy were recorded for each assessment. If the study did not report TP/FN/TN/FN results directly but reported the sensitivity and specificity in combination with total patients/regions and total patients/regions with bone marrow edema as determined by the reference standard, the TP/FN/TN/FN results were calculated. Results were averaged when data for multiple reviewers performing qualitative analysis were reported within a single study [11]. Data was collected on Microsoft Excel v15.14.

Risk of bias assessment

The risk of bias and applicability of each study were evaluated using the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool [12]. The QUADAS-2 tool is used in diagnostic test accuracy systematic reviews to evaluate individual studies for potential sources of bias and concerns regarding applicability by assessing four separate domains including (1) patient selection, (2) index test(s), (3) reference standard, and (4) flow and timing. Studies with a high-risk evaluation for any single signaling question in a domain were considered "high risk of bias" for that domain.

Data analysis

At least one 2×2 contingency table was developed for each individual study. For studies where different types of performance were evaluated and reported (ex. qualitative evaluation of BME with a binary or multi-point grading scale and quantitative comparison of average Hounsfield units (HU) in region of interests of areas with and without BME), multiple contingency tables were developed for an individual study. Meta-analysis was performed using a bivariate random-effects model. The level of analysis was based on reporting of the individual study (per lesion or per-region basis) and was included in the same model. Summary sensitivities, specificities, and area under the ROC curve values were evaluated.

Inter-study variability was assumed and multivariable meta-regression was planned to explore for potential causes; variables that were homogeneous at the study level were chosen in order to minimize "ecological bias" [7]. Statistical analysis of the variability and publication bias were not performed based on the current recommendations from the PRISMA-DTA group [8]. Statistical significance for differences between groups was defined as $p < 0.05$. Analysis was conducted using the "mada" package for R version 3.6.2 (The R Project for Statistical Computing).

Results

Literature search

The literature search PRISMA flow diagram is demonstrated in Fig. 1. Forty articles comprising both conference abstracts and full-text articles were reviewed and twenty studies were included in the analysis [13–32] with 18 studies evaluating the qualitative performance [13–18, 20–25, 27–32] and 8 studies evaluating the quantitative performance of DECT [15, 16, 18, 19, 21, 26, 29, 32]. One study was excluded for using a CT reference standard [33].

Patient, study, and DECT characteristics

A total of 790 patients were included in the review. The mean age ranged between 23 and 80 years and male sex ranged from 13 to 79% of patients depending on the study. Two studies evaluated BME in the wrist or hand [22, 25], 1 study evaluated the sacroiliac joint [32], 5 studies evaluated the hip and/or pelvis [21, 24, 27–29], 8 studies evaluated the knee [13, 14,

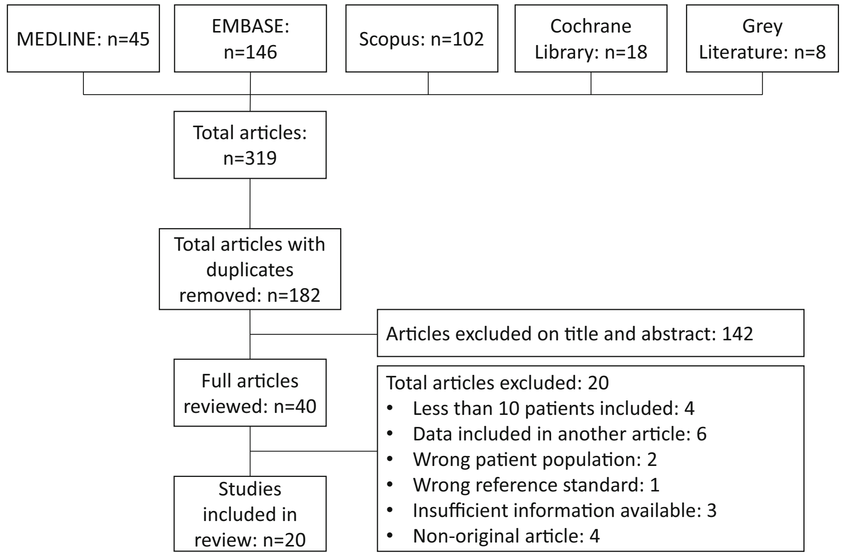

Fig. 1 PRISMA flow diagram showing the screening and selection of studies included in the systematic review

16, 17, 19, 23, 26, 31], 3 studies evaluated the ankle or hindfoot [15, 18, 20], and 1 study evaluated multiple joints [30]. Seventeen studies evaluated only post-traumatic patients and 3 studies evaluated non-traumatic patient populations including patients with rheumatoid arthritis with active clinical synovitis [22], hip pain but no prior trauma [29], and symptomatic axial spondyloarthropathy [32]. One study evaluated both post-traumatic patients and patients with chronic pain [18]. Four studies reported the time between trauma and DECT [13, 14, 17, 30] ranging from acutely after trauma

Table 1 Patient characteristics of individual included studies

Author	Year	No. patients	Age (mean)	Sex (% male)	No. bones with edema (no. regions)*	Site of pathology	Study indication
Ai	2014	14**	25	11 (79%)	NR (36/56 regions)	Knee	Trauma
Bjorkman	2019	48**	23	26 (48%)	52	Knee	Trauma
Booz	2019	62	41	29 (47%)	39	Calcaneus	Trauma
Booz	2020	57	50	27 (47%)	36	Knee	Trauma
Cao	2015	32	40	24 (75%)	32 (127/384 regions)	Knee	Trauma
Foti (knee)	2019	33	54	20 (61%)	17	Knee	Trauma
Foti (ankle)	2019	40	32	29 (73%)	25	Ankle	Trauma and chronic pain
Guggenberger	2012	30	34	15 (50%)	NR (60/300 regions)	Ankle	Trauma
Jang	2019	35	78	25 (71%)	27 fractures	Hip	Trauma
Jans	2018	20	61	9 (45%)	NR	Hand	Non-trauma***
Juhng	2013	23	47	11 (48%)	NR (84/378 regions)	Knee	Trauma
Kellock	2017	118	74	39 (33%)	22 fractures	Hip (proximal femur)	Trauma
Kim	2019	49	NR	NR	NR	Wrist	Trauma
Pache	2010	21	36	16 (76%)	NR (59/236 regions)	Knee	Trauma
Palm	2019	46	80	6 (13%)	31	Pelvis	Trauma
Reddy	2015	25	77	7 (28%)	20 fractures	Hip	Trauma
Son	2019	40**	58	16 (40%)	40	Hip	Non-trauma***
Tu	2016	11	49	5 (45%)	NR (78/253 regions)	Shoulder/knee/ankle	Trauma
Wang	2019	39	36	22 (56%)	NR (43/195 regions)	Knee	Trauma
Wu	2019	47**	27	28 (57%)	NR (55/89 regions)	Sacroiliac joint	Non-trauma***

*Studies which evaluate bone regions, shown with parentheses in the table, evaluated on a per-region rather than per-bone basis

**Indicates that bilateral joint imaging was performed

***Indications for non-trauma imaging included: patients with rheumatoid arthritis with active clinical synovitis [22], patients with hip pain but no prior trauma [29], and patients with symptomatic axial spondyloarthropathy [32]

[23] to within 100 days of trauma [13]. Patient characteristics of individual included studies are shown in Table 1.

Study characteristics of individual studies included are detailed in Table 2. Three studies were performed in North America, 9 studies were performed in Europe, and 7 studies were performed in Asia. Studies were equally mixed between prospective and retrospective study designs. All studies were performed at a single academic center. MRI alone was the reference standard for 17 studies. Two studies used a mix of MRI and clinical/surgical follow-up in suspected hip fractures following trauma [21, 28], and one study evaluated patients with suspected hip fractures with a 30-day surgical and/or clinical outcome reference standard [24]. All but one study used multiple readers for evaluation. Studies primarily used board-certified radiologists with several years of experience for evaluation.

DECT characteristics of individual studies included are shown in Table 3. All but one reporting study used a Siemens DECT. Most studies used 80 kV for low voltage and 140 kV for high voltage. Variable current-time products were used dependent on the study. Most studies used a tin filter for higher voltage imaging. Studies used a mix of soft tissue and bone kernels with most studies reporting the use of a three-material decomposition algorithm in post-processing. Thirteen of the reporting studies used a binary evaluation method, with 7 studies reporting the use of a multi-point scale.

Diagnostic accuracy of DECT for bone marrow edema in the appendicular skeleton

Performance results for individual studies by the site of pathology are presented in Appendix Table 7. The pooled and

Table 2 Study characteristics of individual included studies

Author	Year	Country	Study design	Reference standard	Time interval (days)*	No. readers	Consensus reading	Reader experience (years)
Ai	2014	USA	Retrospective	MRI	≤ 90 days	2	Yes	MSK radiologists (11 + 19)
Bjorkman	2019	Sweden	Prospective	MRI	< 7	2	Yes	Resident (3) + radiologist (> 7)
Booz	2019	Germany	Retrospective	MRI	< 7	5	No	Qualitative: 4 residents (3–5) + radiologist (8) Quantitative: medical student (2)
Booz	2020	Germany	Retrospective	MRI	< 7	6	No	Qualitative: 4 residents (3–5) + 2 radiologists (7–10) Quantitative: medical student (2)
Cao	2015	China	Prospective	MRI	< 7	2	Yes	Radiologists ("vastly experienced")
Foti (knee)	2019	Italy	Prospective	MRI	< 7	2	No	Radiologists (11 + 25)
Foti (ankle)	2019	Italy	Prospective	MRI	< 7	2	Yes	Radiologists (15 + 35)
Guggenberger	2012	Switzerland	Prospective	MRI	≤ 1	2	No	Radiologists (3 + 4)
Jang	2019	South Korea	Retrospective	MRI + Clinical	< 30	2	No	Qualitative: resident (3) + radiologist (14) Quantitative: trainees (3 + 6)
Jans	2018	Belgium	Prospective	MRI	NR	2	Yes	Radiologists (NR)
Juhng	2013	South Korea	Prospective	MRI	NR	2	No	Radiologists (5 + 25)
Kellock	2017	Canada	Retrospective	Clinical**	< 30	3	No	Fellow (6) + 2 radiologists (8 + 19)
Kim	2019	South Korea	Retrospective	MRI	NR	2	No	NR
Pache	2010	Germany	Prospective	MRI	≤ 5	2	No	Radiologists (5 + 7)
Palm	2019	Germany	Retrospective	MRI	≤ 14	1	No	Radiologist (NR)
Reddy	2015	Canada	Retrospective	MRI + Clinical	< 30	2	Yes	Radiologists (3 + 10)
Son	2019	South Korea	Retrospective	MRI	≤ 31	2	No	Resident (0.75) + radiologist (9)
Tu	2016	China	Prospective	MRI	≤ 1	2	No	MSK radiologists ("senior")
Wang	2019	China	Prospective	MRI	< 6	2	No	Radiologists (> 5)
Wu	2019	China	Prospective	MRI	≤ 1	2	No	Radiologists (17 + 25)

NR, not reported

*Time interval between the dual-energy CT and reference standard in days

**Study evaluated patients with suspected hip fracture using surgery with visualized hip fracture within 30 days of presentation and/or clinical follow-up as the reference standard

Table 3 Dual-energy CT characteristics of individual included studies

Author	Year	DECT brand	kV 1	kv 2	mAs	Tin filter	Pitch	No. detector rows	Collimation (mm)	Rotation time (s)	Slice width (mm)	Evaluation method
Ai	2014	Siemens	80	140	340 + 238	NR	0.7	40	0.6	NR	0.5	Binary
Bjorkman	2019	Siemens	80	150	125	Yes (150 kV)	0.7	128	0.6	0.5		Binary
Booz	2019	Siemens	90	150	180	Yes (150 kV)	0.6	192	0.6	0.5	0.75	Binary
Booz	2020	Siemens	90	150	180	Yes (150 kV)	0.6	192	0.6	0.5	0.75	Binary
Cao	2015	Siemens	80	140	255 + 60	NR	0.9	20	0.6	NR	NR	4 pt. Scale
Foti (knee)	2019	Siemens	80	150	NR	Yes (150 kV)	NR	NR	NR	NR	NR	Binary
Foti (ankle)	2019	Siemens	80	150	220 + 138	Yes (150 kV)	NR	NR	0.6	NR	0.75	Binary
Guggenberger	2012	Siemens	80	140	360 + 180	Yes (140 kV)	0.7	40	0.6	1	2	4 pt. Scale
Jang	2019	Siemens	100	140	160	Yes	0.6	32	0.6	0.5	2	Binary
Jans	2018	Siemens	80	140	NR	NR	0.7	20	0.6	1	2	Binary
Juhng	2013	Siemens	80	140	50 + 213	NR	0.7	20	0.6	0.5	NR	4 pt. scale
Kellock	2017	Siemens	100	140	160	Yes	0.7	40	0.6	0.5	0.75	Binary + diagnostic confidence (10 pt. scale)
Kim	2019	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	Binary
Pache	2010	Siemens	80	140	234 + 56	NR	0.7	20	0.6	0.5	2	3 pt. scale
Palm	2019	Siemens	80	150	NR	NR	NR	NR	NR	NR	NR	Binary
Reddy	2015	Siemens	100	140	160	Yes	0.7	40	0.6	0.33	1	Binary
Son	2019	GE	80	140	405	NR	0.5	NR	NR	0.5	NR	3 pt. scale
Tu	2016	Siemens	80	140	234 + 55		0.7	NR	0.6	NR	NR	4 pt. scale
Wang	2019	Siemens	80	150	250 + 150	Yes (150 kV)	0.6	128	0.6	0.5	2	Binary
Wu	2019	Siemens	80	140	220 + 138	Yes (140 kV)	0.7	40	0.6	1	1	3 pt. scale

DECT, dual-energy computed tomography; *NR*, not reported; *3 pt.*, three point; *4 pt.*, four point; *10 pt.*, ten point

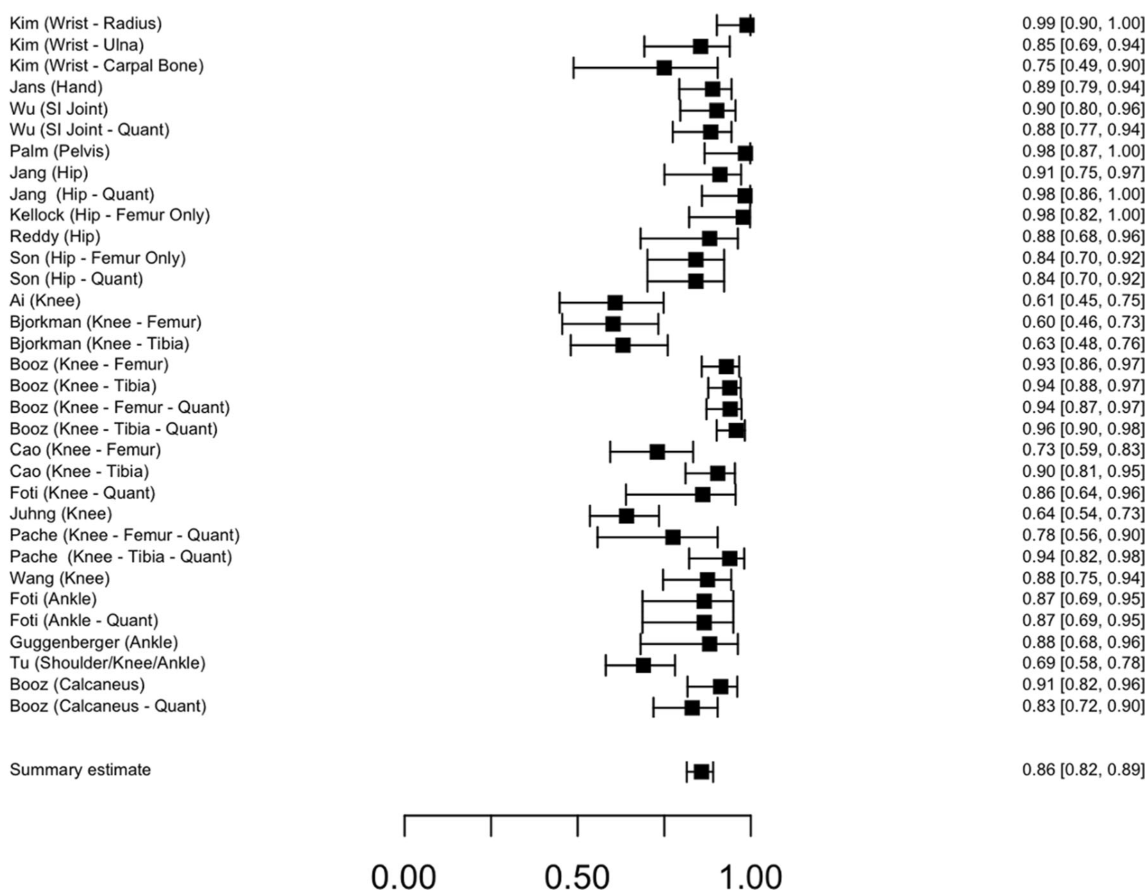

Fig. 2 Forest plots of primary study sensitivity of dual-energy CT for detecting bone marrow edema in the appendicular skeleton

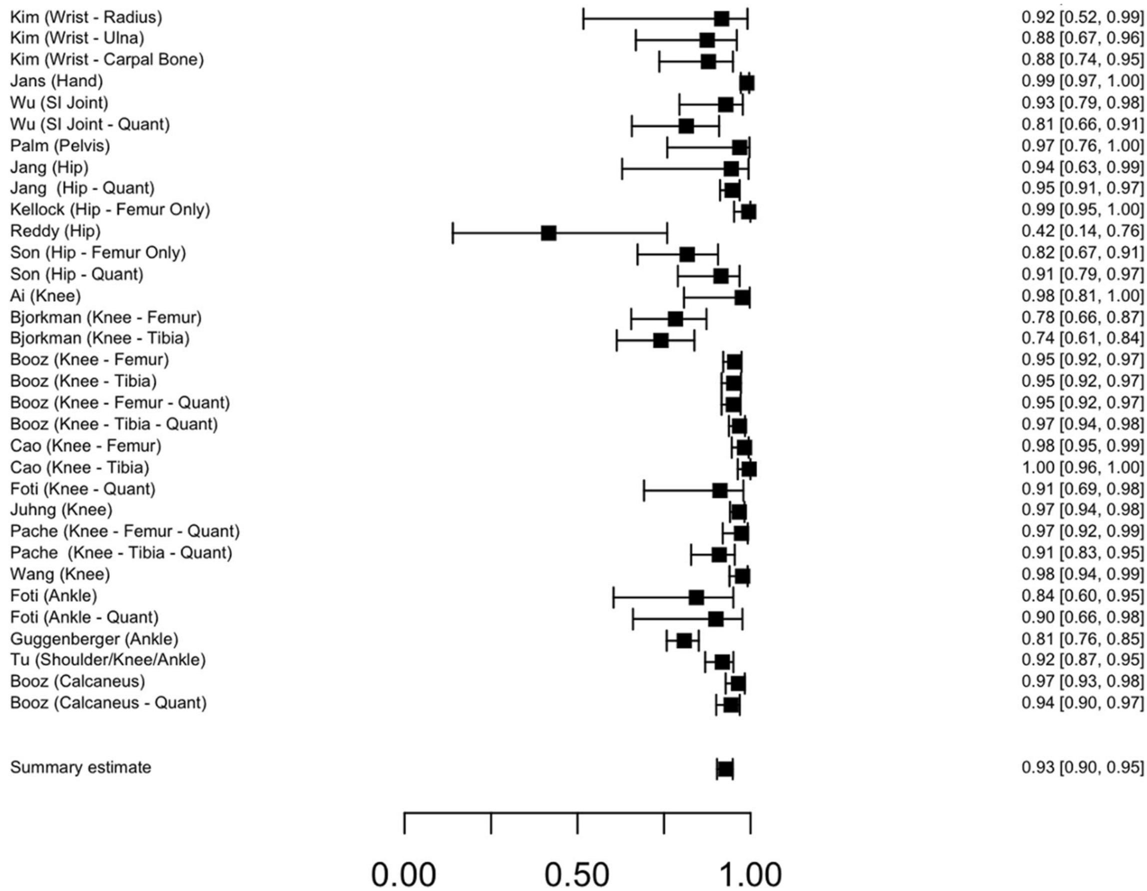

Fig. 3 Forest plots of primary study specificity of dual-energy CT for detecting bone marrow edema in the appendicular skeleton

weighted sensitivity and specificity of DECT for detecting BME in the appendicular skeleton were 86% (95% confidence interval [CI] 82–89%), 93% (95% CI 90–95%), and 0.95, respectively (Figs. 2, 3, and 4). Nearly all studies reporting inter-observer and intra-observer agreement included kappa-statistic values within the substantial to perfect agreement range [34]. Two studies demonstrated lower inter-observer agreement when evaluating the knee joint [14] and carpal bones specifically [25].

Meta-regression was performed evaluating qualitative versus quantitative analysis, trauma versus non-trauma study indication, upper versus lower extremity location, and low versus high or unclear risk of bias studies (Table 4). Quantitative analysis had a statistically higher sensitivity ($p = 0.01$) but no difference in specificity ($p = 0.28$) compared with qualitative analysis. No other differences between sensitivity and specificity were identified on meta-regression.

Risk of bias assessment

The results of QUADAS-2 assessment for risk of bias and applicability in individual studies are shown in Table 5. Studies were predominantly low risk or unclear risk of bias across domains for both risk of bias and applicability. Studies were a mix of low risk and/or high risk of bias for index test dependent on the evaluation method. Where qualitative evaluation methods were pre-specified, risk of bias for index test was deemed low. Where performance of quantitative evaluation of continuous variables was assessed with a retrospective cut-off value, studies were deemed high risk of bias for index test. One study was deemed high risk of bias in the reference standard domain for using 3 different MRI magnet strengths (1 T, 1.5 T,

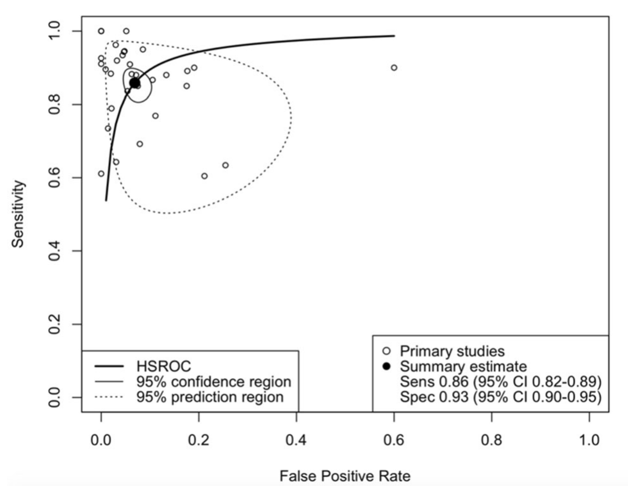

Fig. 4 Summary area under the ROC curve of dual-energy CT for detecting bone marrow edema in the appendicular skeleton

Table 4 Bivariate random-effects multivariable meta-regression evaluating for causes of variability among studies

	Sensitivity (*p* value)	Specificity (*p* value)
Qualitative versus quantitative analysis	0.01*	0.28
Trauma versus non-trauma study indication	0.95	0.70
Upper versus lower extremity location	0.10	0.52
Low versus high or unclear risk of bias**	0.05	0.17

*Quantitative analysis has a statistically higher sensitivity than qualitative analysis

**Studies were only considered "low risk" if they met criteria for low risk in each of the four QUADAS-2 domains

Table 5 Results of QUADAS-2 assessment for risk of bias and applicability in individual studies. Studies using quantitative assessment with a retrospective cut-off value were deemed high risk of bias in index test. Studies using a prospective qualitative analysis and retrospective cut-off for quantitative analysis were deemed both low risk and high risk, respectively, for index test dependent of type of analysis [15, 16, 18, 21, 29, 32]

	Risk of Bias				*Applicability Concerns*		
	Patient Selection	*Index Test*	*Reference Standard*	*Flow and Timing*	*Patient Selection*	*Index Test*	*Reference Standard*
Ai 2014	+	+	+	−	+	+	+
Bjorkman 2019	+	+	+	+	+	+	+
Booz 2019	+	+ −	+	+	+	+	+
Booz 2020	+	+ −	+	+	+	+	+
Cao 2015	+	+	+	+	+	+	+
Foti (knee) 2019	?	−	?	?	?	?	+
Foti (ankle) 2019	+	+ −	+	+	+	+	+
Guggenberger 2012	+	+	+	+	+	+	+
Jang 2019	+	+ −	+	?	+	+	+
Jans 2018	+	?	?	?	+	?	+
Juhng 2013	?	+	+	?	+	+	+
Kellock 2017	+	+	+	+	+	+	+
Kim 2019	?	?	?	?	?	?	+
Pache 2010	+	−	+	+	+	+	+
Palm 2019	+	+	−	+	+	+	+
Reddy 2015	+	+	+	+	+	+	+
Sun 2019	+	+ −	+	−	+	+	+
Tu 2016	−	+	+	+	+	+	+
Wang 2019	+	+	+	+	+	+	+
Wu 2019	+	+ −	+	+	+	+	+

and 3 T) [27]. Two studies were deemed high risk of bias in the flow and timing domain for using maximum intervals between DECT and MRI studies over 2 weeks' time [13, 29].

Discussion

This meta-analysis demonstrates that the sensitivity (86% [95% CI 82–89%]), specificity (93% [95% CI 90–95%]), and AUC values (0.95) of DECT for detecting BME in the appendicular skeleton compared with a MRI and/or clinical outcome are similar or better than prior meta-analyses largely evaluating the axial skeleton [4–6]. These findings counter a prior meta-regression analysis by Li et al indicating that bone position in the appendicular skeleton, notably the ankle in their analysis, was likely to contribute to variability when compared with studies including the axial skeleton [4]. This meta-analysis therefore contributes to the growing evidence suggesting that DECT can be a viable alternative imaging modality to MRI for BME with predominantly substantial to perfect agreement, despite an appendicular location.

This meta-analysis used a rigorous methodological design and is not subject to the small number of included studies seen previously [5, 6]. Meta-regression of key features felt to potentially account for presumed variability amongst studies was performed, but only identified one potential source of variability. A higher sensitivity and similar specificity were identified with quantitative rather than qualitative analysis. If this difference represents a true result, this may suggest that quantitative analysis offers a better ability to exclude regions without BME while preserving specificity, compared with qualitative assessment. However, given that a similar effect was not identified for specificity, and numerous potential confounding variables were not included in the regression model, this result should be interpreted with caution. Some alternative patient- and study-specific causes which were not explored but may contribute to variability include age, specific location or region of bone marrow edema, bone size, differences in reader experience, and/or differences in the reference standard. The studies included in the analysis did use many similar DECT imaging characteristics, decreasing the likelihood that index test parameters contribute to variability.

There may also be some limitations in the generalizability of the findings of this review to non-expert readers. Most included studies evaluated only board-certified radiologists with several years of experience for qualitative interpretation of BME which could result in higher sensitivity and specificity values than is seen in general practice. In a study with multiple reviewers evaluating DECT for the presence of BME in patients with osteoporotic vertebral compression fractures, the most experienced reader demonstrated near MRI accuracy for BME while less experienced readers did not [35].

Despite the limitations regarding unexplored confounding variables impacting accuracy, as well as risk of bias in several studies, this meta-analysis indicates that DECT may be an alternative to MR imaging for evaluating BME in the appendicular skeleton when MR is not available or contraindicated. Further exploration of sources for variability, potentially through an individual patient data meta-analysis, would be helpful to increase the confidence of the generalizability of these results.

Funding Studentship funding for this systematic review was provided through an Office of the Provost and VP (Academic) Summer Student Award from the University of Alberta Faculty of Medicine and Dentistry, and through the Radiology Endowment Fund from the University of Alberta Department of Radiology and Diagnostic Imaging.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Mitchell P Wilson.

Conflict of interest The authors of this manuscript declare no relationships with any companies whose products or services may be related to the subject matter of the article.

Statistics and biometry Three of the authors have significant statistical expertise (MHM, MDFM, TAM).

Informed consent Written informed consent was not required for this study because it included only de-identified data with individual studies acquiring ethical approval from their home institution where necessary.

Ethical approval Institutional review board approval was not required because it included only de-identified data with individual studies acquiring ethical approval from their home institution where necessary.

Methodology
• Systematic review and meta-analysis using a bivariate random-effects model with multivariable meta-regression

Appendix

Table 6 Search strategy by database

MEDLINE: (exp Tomography, X-Ray Computed/ AND dual energy.mp.) AND (edema.mp. OR exp. Edema/) AND bone.mp. OR exp. "Bone and Bones"/)
EMBASE: (dual energy.mp. AND exp. computer assisted tomography/) AND (edema.mp. OR exp. edema/) AND (bone.mp. OR exp. bone/ OR exp. bone injury/)
COCHRANE REVIEW: TITLE-ABS-KEY(dual energy AND edema AND bone)
SCOPUS: TITLE-ABS-KEY(dual AND energy AND (computed AND tomography OR CT)) AND TITLE-ABS-KEY(edema) AND TITLE-ABS-KEY(bone)

Table 7 Performance results for individual studies by site of pathology

Author	Year	Site of pathology	No. regions	No. Marrow edema	TP	FP	FN	TN	Sensitivity	Specificity	Accuracy	Inter-observer agreement*	Intra-observer agreement*	Cut-off value**
Ai	2014	Knee	56	36	22	0	14	20	61	100	75			Qualitative
Bjorkman	2019	Knee - femur	95	43	26	11	17	41	60	79	71	0.37	0.64–0.66	Qualitative
Bjorkman	2019	Knee - tibia	96	41	26	14	15	41	63	75	70	0.41	0.54–0.59	Qualitative
Booz	2019	Calcaneus	248	62	57	6	5	180	92	97	96	0.84		Qualitative
Booz	2019	Calcaneus	247	62	51	10	10	176	82	95	92			− 53 HU
Booz	2020	Knee - femur	342	91	85	11	6	240	93	96	95	0.81		Qualitative
Booz	2020	Knee - tibia	342	106	100	11	6	225	94	95	95	0.84		Qualitative
Booz	2020	Knee - femur	342	91	86	12	5	239	95	95	95			− 42 HU
Booz	2020	Knee - tibia	342	106	102	7	4	229	96	97	97			− 51 HU
Cao	2015	Knee - femur	192	49	36	2	13	141	73	99	92	0.66		Qualitative
Cao	2015	Knee - tibia	192	67	61	0	6	125	91	100	97	0.8		Qualitative
Foti	2019	Knee	33	17	15	1	2	15	90	93	91	0.82	0.86	− 15 HU
Foti	2019	Ankle	40	25	22	2	3	13	88	87	88	0.83	0.87	Qualitative
Foti	2019	Ankle	39	25	22	1	3	13	88	93	88			− 20 HU
Guggenberger	2012	Ankle	292	52	38	32	15	208	73	87	84	0.66		Qualitative
Guggenberger	2012	Ankle	292	52	18	52	2	221	90	81	82			Qualitative
Jang	2019	Hip	35	27	25	0	2	8	71	100	94	0.67–0.77		Qualitative
Jang	2019	Hip – femur + pelvis	282	29	29	13	0	240	100	95	95	0.95		− 55 HU
Jans	2018	Hand	400	67	60	3	7	330	90	99	98			Qualitative
Juhng	2013	Knee - femur	162	29	17	7	12	126	59	95	88			Qualitative
Juhng	2013	Knee - tibia	162	44	33	3	11	116	75	97	92			Qualitative
Juhng	2013	Knee – femur +tibia +patella	378	84	54	9	30	285	64	97	90			
Kellock	2017	Hip - femur	118	22	22	0	0	96	100	100	100	0.96		Qualitative
Kim	2019	Wrist - radius	49	44	44	0	0	5	100	100	100	1		Qualitative
Kim	2019	Wrist - ulna	49	30	26	2	4	17	88	88	88	1		Qualitative
Kim	2019	Wrist - carpals	49	13	10	4	3	32	79	89	86	0.43		Qualitative
Pache	2010	Knee - femur	114	19	15	2	4	93	79	98	95	0.78		− 33 HU
Pache	2010	Knee - tibia	122	40	38	7	2	75	95	92	93	0.87		− 60 HU
Palm	2019	Pelvis	46	31	31	0	0	15	100	100	100			Qualitative
Reddy	2015	Hip	25	20	18	3	2	2	90	40	80			Qualitative
Son	2019	Hip - femur	80	40	34	7	6	33	85	83	84			Qualitative
Son	2019	Hip - femur	80	40	34	3	6	37	85	93	89			951 mg/cm^3
Son	2019	Hip - femur	80	40	34	40	6	29	85	73	79			957 mg/cm^3
Tu	2016	Shoulder +knee +ankle	255	78	54	14	24	163	69	92	85	0.61		Qualitative
Tu	2016	Shoulder +knee +ankle	253	78	62	15	53	123	54	89	73	0.42		Qualitative
Wang	2019	Knee	195	43	38	3	5	149	88	98	96	0.88		Qualitative
Wu	2019	Sacroiliac joint	89	55	50	2	5	32	91	94	92	0.81		Qualitative
Wu	2019	Sacroiliac joint	89	55	49	6	6	28	90	83	87			− 33 HU
Wu	2019	Sacroiliac joint	89	55	52	7	3	27	94	80	87			− 42 HU

Results were averaged for studies with multiple reviewers evaluating bone marrow edema with qualitative analysis [11]

TP, true positive; *FP*, true negative; *FN*, false negative; *TN*, true negative; *HU*, Hounsfield unit

*Agreement is reported as a kappa statistic where 0.61–0.80 is "substantial agreement," 0.81–0.99 is "almost perfect agreement," and 1 is "perfect agreement" [34]

**Cut-off value for quantitative studies determined as a region of interest measuring at least 3 mm placed at least 1 mm from the cortical margin. One study evaluated water mass density measured as mg/cm^3 [29]

References

1. Boks SS, Vroegindeweij D, Koes BW, Hunink MG, Bierma-Zeinstra SM (2006) Follow-up of occult bone lesions detected at MR imaging: systematic review. Radiology 238(3):853–862
2. Mandalia V, Fogg AJ, Chari R, Murray J, Beale A, Henson JH (2005) Bone bruising of the knee. Clin Radiol 60(6):627–636
3. Komlosi P, Wintermark M (2017) Dual energy computed tomography applications for the evaluation of the spine. Neuroimaging Clin N Am 27(3):483–487
4. Li M, Qu Y, Song B (2017) Meta-analysis of dual-energy computed tomography virtual non-calcium imaging to detect bone marrow edema. Eur J Radiol 95:124–129
5. Suh CH, Yun SJ, Jin W, Lee SH, Park SY, Ryu CW (2018) Diagnostic performance of dual-energy CT for the detection of bone marrow oedema: a systematic review and meta-analysis. Eur Radiol 28(10):4182–4194
6. Yang P, Wu G, Chang X (2018) Diagnostic accuracy of dual-energy computed tomography in bone marrow edema with vertebral compression fractures: a meta-analysis. Eur J Radiol 99:124–129
7. Macaskill P, Gatsonis C, Deeks JJ, Harbord R, Takwoingi Y (2010) Chapter 10: analyzing and presenting results. In: Deeks JJ, Bossuyt PM, Gatsonis C (eds) Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy version 1.0. The Cochrane Collection Available via http://srdta.cochrane.org/. Accessed 24 Jan 2020
8. McInnes MDF, Moher D, Thombs BD et al (2018) Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. JAMA 319(4):388–396
9. McGrath TA, Bossuyt PM, Cronin P et al (2019) Best practices for MRI systematic reviews and meta-analyses. J Magn Reson Imaging 49:e51–e64
10. Sampson M, McGowan J, Cogo E, Grimshaw J, Moher D, Lefebvre C (2009) An evidence-based practice guideline for the peer review of electronic search strategies. J Clin Epidemiol 62: 944–952
11. McGrath TA, McInnes MDF, Langer FW, Hong J, Korevaar DA, Bossuyt PMM (2017) Treatment of multiple test readers in diagnostic accuracy systematic reviews-meta-analyses of imaging studies. Eur J Radiol 93:59–64
12. Whiting PF, Rutjes AW, Westwood ME et al (2011) QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 155:529–536
13. Ai S, Qu M, Glazebrook KN et al (2014) Use of dual-energy CT and virtual non-calcium techniques to evaluate post-traumatic bone bruises in knees in the subacute setting. Skeletal Radiol 43:1289–1295
14. Bjorkman AS, Koskinen SK, Lindblom M, Persson A (2020) Diagnostic accuracy of dual-energy CT for detection of bone marrow lesions in the subacutely injured knee with MRI as reference method. Acta Radiol. https://doi.org/10.1177/0284185119877343
15. Booz C, Noske J, Albrecht MH et al (2019) Traumatic bone marrow edema of the calcaneus: evaluation of color-coded virtual non-calcium dual-energy CT in a multi-reader diagnostic accuracy study. Eur J Radiol 118:207–214
16. Booz C, Noske J, Lenga L (2020) Color-coded virtual non-calcium dual-energy CT for the depiction of bone marrow edema in patients with acute knee trauma: a multireader diagnostic accuracy study. Eur Radiol 30(1):141–150
17. Cao JX, Wang YM, Kong XQ, Yang C, Wang P (2015) Good interrater reliability of a new grading system in detecting traumatic bone marrow lesions in the knee by dual energy CT virtual non-calcium images. Eur J Radiol 84(6):1109–1115
18. Foti G, Catania M, Caia S et al (2019) Identification of bone marrow edema of the ankle: diagnostic accuracy of dual-energy CT in comparison with MRI. Radiol Med 124(10):1028–1036
19. Foti G, Catania M, Romano L, Caia S, Carbognin G (2019) Identification of bone marrow edema and osteochondral injuries of the knee: diagnostic accuracy of dual-energy CT and virtual non-calcium techniques. Semin Musculoskelet Radiol 23(S02): S1–S18
20. Guggenberger R, Gnannt R, Hodler J et al (2012) Diagnostic performance of dual-energy CT for the detection of traumatic bone marrow lesions in the ankle: comparison with MR imaging. Radiology 264(1):164–173
21. Jang SW, Chung BM, Kim WT, Gil JR (2019) Nondisplaced fractures on hip CT: added value of dual-energy CT virtual non-calcium imaging for detection of bone marrow edema using visual and quantitative analyses. Acta Radiol 60(11):1465–1473
22. Jans L, De Kock I, Herregods N et al (2018) Dual-energy CT: a new imaging modality for bone marrow oedema in rheumatoid arthritis. Ann Rheum Dis 77(6):958–960
23. Juhng SK, Seo SH, Kang EJ, et al (2013) Clinical usefulness of dual energy CT to detect posttraumatic bone bruise of the knee joint. ECR 2013:C-0349. Accessed 17 Feb 2020
24. Kellock TT, Nicolaou S, Kim SSY et al (2017) Detection of bone marrow edema in nondisplaced hip fractures: utility of a virtual noncalcium dual-energy CT application. Radiology 284(3):798–805
25. Kim J, Yoo H, Chae H et al (2019) Diagnostic accuracy of dual-layer detector CT using calcium-supressed images for the detection of bone marrow edema in the wrist. RSNA 2019:archive.rsna.org/2019/19017350.html. Accessed 17 Feb 2020
26. Pache G, Krauss B, Strohm P et al (2010) Dual-energy CT virtual noncalcium technique: detecting posttraumatic bone marrow lesions – feasibility study. Radiology 256(2):617–624
27. Palm HG, Lang P, Hackenbroch C, Sailer L, Friemert B (2020) Dual-energy CT as an innovative method for diagnosing fragility fractures of the pelvic ring: a retrospective comparison with MRI as the gold standard. Arch Orthop Trauma Surg 140(4):473–480
28. Reddy T, McLaughlin PD, Mallinson PI et al (2015) Detection of occult, undisplaced hip fractures with a dual-energy CT algorithm targeted to detection of bone marrow edema. Emerg Radiol 22:25–29
29. Son W, Park C, Jeong HS, Song YS, Lee IS (2020) Bone marrow edema in non-traumatic hip: high accuracy of dual-energy CT with water-hydroxyapatite decomposition imaging. Eur Radiol 30(4): 2191–2198
30. Tu N, Li Z, Yang J et al (2016) Study on assessment of bone marrow lesions by dual-energy computed tomography. Chin Imaging J Integr Tradit West Med 1:1–4
31. Wang MY, Zhang XY, Xu L et al (2019) Detection of bone marrow oedema in knee joints using a dual-energy CT virtual non-calcium technique. Clin Radiol 74(10):815.e1–815.e7
32. Wu H, Zhang G, Shi L et al (2019) Axial spondyloarthritis: dual-energy virtual noncalcium CT in the detection of bone marrow edema in the sacroiliac joints. Radiology 290(1):157–164
33. Ali IT, Wong WD, Liang T et al (2018) Clinical utility of dual-energy CT analysis of bone marrow edema in acute wrist fractures. AJR Am J Roentgenol 210:842–847
34. Viera AJ, Garrett JM (2005) Understanding interobserver agreement: the kappa statistic. Fam Med 37(5):360–363
35. Kaup M, Wichmann JL, Scholtz JE et al (2016) Dual-energy CT-based display of bone marrow edema in osteoporotic vertebral compression fractures: impact on diagnostic accuracy of radiologists with varying levels of experience in correlation to MR imaging. Radiology 280:510–519

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.