Seminars in Arthritis and Rheumatism ▮ (2017) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Seminars in Arthritis and Rheumatism

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

Diagnostic accuracy of dual-energy computed tomography in patients with gout: A meta-analysis☆,☆☆

Young Ho Lee, MD, PhD*, Gwan Gyu Song, MD, PhD

Department of Internal Medicine, Korea University Medical Center, Korea University College of Medicine, Seoul, South Korea

ARTICLE INFO

Keywords:
Gout
DECT
Diagnostic accuracy
Meta-analysis

ABSTRACT

Objective: This study aimed to evaluate the diagnostic performance of dual-energy computed tomography (DECT) for patients with gout.
Methods: We searched the Medline, Embase, and Cochrane Library databases, and performed a meta-analysis on the diagnostic accuracy of DECT in patients with gout.
Results: A total of eight studies including 510 patients with gout and 268 controls (patients with non-gout inflammatory arthritis) were available for the meta-analysis. The pooled sensitivity and specificity of DECT were 84.7% (95% confidence interval [CI]: 81.3–87.7) and 93.7% (93.0–96.3), respectively. The positive likelihood ratio, negative likelihood ratio, and diagnostic odds ratio were 9.882 (6.122–15.95), 0.163 (0.097–0.272), and 78.10 (31.14–195.84), respectively. The area under the curve of DECT was 0.956 and the Q^* index was 0.889, indicating a high diagnostic accuracy. Some between-study heterogeneity was found in the meta-analyses. However, there was no evidence of a threshold effect (Spearman correlation coefficient = 0.419; $p = 0.035$). In addition, meta-regression showed that the sample size, study design, and diagnostic criteria were not sources of heterogeneity, and subgroup meta-analyses did not change the overall diagnostic accuracy.
Conclusions: Our meta-analysis of published studies demonstrates that DECT has a high diagnostic accuracy and plays an important role in the diagnosis of gout.

Introduction

Gout is an inflammatory disorder characterized by hyperuricemia, and by the deposition of monosodium urate (MSU) crystals in intra-articular and peri-articular locations, resulting in episodic gout flares, gouty arthropathy, tophi formation, and urolithiasis [1]. The gold standard for diagnosis of gout is microscopic analysis of synovial fluid aspirate, which reveals negatively birefringent needle-shaped MSU crystals in polarized light microscopy [2]. However, joint aspiration can be technically challenging in patients with small amounts of joint fluid, and identification via joint aspiration is not always possible [3]. In addition, synovial fluid aspiration may not reveal uric acid crystals in up to 25% of patients with gout [4]. Owing to these limitations, synovial aspiration has been performed in a small percentage of gout cases, and patients are often diagnosed based on hyperuricemia, and clinical and radiographic findings. In case of non-visualization of MSU crystals in the joint aspirate, clinical, radiographic, and laboratory criteria may be helpful in diagnosis [5]. However, the accuracy of detection of uric acid crystals varies considerably among laboratories; the serum level of uric acid can be within normal limits in acute gouty arthritis, and most patients with hyperuricemia do not experience gout [6].

Dual-energy computed tomography (DECT) is a new diagnostic tool for gout; it is equipped with two X-ray tubes, allowing simultaneous acquisition at two different energy levels [6]. Unlike conventional computed tomography (CT) scans, DECT can characterize the composition of the material according to the differential X-ray photon energy-dependent attenuation of the compounds being examined at the two different energy levels [7]. Thus, DECT visualizes uric acid crystal deposits and bone structures using different display colors, and specifically identifies MSU crystals and differentiates them from other types of crystals.

The diagnostic accuracy of DECT has been studied in the context of gout, and it showed controversial results in patients with gout [8–15]. This may be attributed to the small sample sizes, low statistical power, and/or the presence of clinical heterogeneity. To overcome the limitations of the individual studies, resolve the inconsistencies, and evaluate DECT for its assessment, a systematic analytical approach is warranted [16]. Therefore, we performed a

☆The authors have no financial or non-financial conflict of interest to declare.
☆☆This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.
* Correspondence to: Division of Rheumatology, Department of Internal Medicine, Korea University Anam Hospital, Korea University College of Medicine, 73, Inchon-ro, Seongbuk-gu, Seoul 02841, South Korea.
E-mail address: lyhcgh@korea.ac.kr (Y.H. Lee).

http://dx.doi.org/10.1016/j.semarthrit.2017.03.002
0049-0172/

meta-analysis on the sensitivity and specificity of DECT for the diagnosis of gout in order to assess the diagnostic accuracy of DECT, using published data.

Materials and methods

Identification of eligible studies and data extraction

We used the Medline, Embase, and Cochrane Library databases to identify articles published until October 2016, in which DECT was performed in patients with gout and in control subjects. In addition, all references cited in the selected articles were reviewed to identify studies not indexed by the electronic databases. PICO stands for Patient, Intervention, Comparison, and Outcome. The PICO of this study was patients with gout (P), DECT (I), classification criteria or joint aspirates for gout (C), and the sensitivity, specificity (O). To incorporate the concept from the PICO analysis in the search strategy, the following keywords and subject terms were used in the search: “dual-energy computed tomography OR dual-energy CT OR DECT”, “sensitivity”, and “gout”. Studies were selected for the analysis if they included (i) case–control, cross-sectional, or cohort studies, (ii) sufficient data to calculate the sensitivity and specificity of DECT, and (iii) patients with gout diagnosed on the basis of the classification criteria [17] or the demonstration of MSU crystals on a joint aspirate. No language or race restrictions were applied. Studies with overlapping or insufficient data, and review studies were excluded. Two independent reviewers extracted data about the methods and results of meta-analysis from the original studies. Any discrepancies between the reviewers were resolved by consensus. The meta-analysis was conducted in accordance with PRISMA guidelines [18]. We extracted information on author(s), the publication year, and the demographic characteristics of the participants (age, disease duration, country, and diagnostic criteria) from each study. DECT data were extracted from all primary studies to fill the four cell values of a diagnostic 2 × 2 table (true positives, false positives, true negatives, and false negatives). The Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) was used to assess the quality of each study [19].

Evaluation of statistical associations

Within- and between-study variations and heterogeneities were assessed using Cochran’s *Q*-statistic. Cochran's *Q*-statistic test assesses the null hypothesis that all studies evaluated the same effect. The effect of heterogeneity was quantified using I^2 with a range from 0% to 100%, representing the proportion of between-study variability attributable to heterogeneity rather than to chance [20]. I^2 values of 25%, 50%, and 75% were nominally assigned as low, moderate, and high estimates, respectively. The fixed-effects model assumes that a genetic factor has a similar effect on disease susceptibility across all studies investigated, and that observed variations among studies are caused by chance alone [21]. The random-effects model assumes that different studies show substantial diversity, and assesses both within-study sampling error and between-study variance [22]. The random-effects model is the most appropriate to use in the presence of significant between-study heterogeneity [22]. We used a random-effects model to combine the sensitivity, specificity, positive and negative likelihood ratio (PLR, NLR), and diagnostic odds ratio (DOR) estimates due to heterogeneity, and analyzed the summary receiver-operating characteristic (SROC) curves. DOR is a unitary measure of diagnostic performance that encompasses both sensitivity and specificity, or both PLR and NLR, and it is considered a suitable global measure of accuracy for comparing the overall diagnostic accuracies of different tests [23]. As sensitivity and specificity are interdependent, independent calculations may sometimes underestimate both variables. SROC curve analysis is more appropriate, because it accounts for this mutual dependence. The area under the curve (AUC) (in this case, area under the SROC curve) presents an overall summary of test performance and displays the trade-off between sensitivity and specificity, and an AUC of 1.0 (100%) indicates perfect discriminatory ability for a diagnostic test (13). In addition, the Q^* index is another useful global estimate of test accuracy for comparing SROC curves. The Q^* index is defined as the point where sensitivity equals specificity on an SROC curve, and is the point on an SROC curve intersected by the anti-diagonal. A Q^* value of 1.0 indicates 100% accuracy (i.e., sensitivity and specificity of 100%) (13). Statistical manipulations for this meta-analysis were performed using Meta-DiSc version 1.4 (Hospital Universitario Ramon y Cajal, Madrid, Spain) [24], a Comprehensive Meta-Analysis computer program (Biostat, Englewood, NJ, USA), and a RevMan software version 5.3 (The Nordic Cochrane Centre, The Cochrane Collaboration, Copenhagen, Denmark).

Evaluation of heterogeneity, meta-regression, and publication bias

A between-study heterogeneity observed in a meta-analysis indicates variability in results across studies. A threshold effect is the most important cause of heterogeneity. Different sensitivities and specificities due to various study conditions cause different threshold effects. We checked the Spearman correlation coefficient between the logit of sensitivity and the logit of 1-specificity to assess the presence of a threshold effect. To examine the potential source of heterogeneity observed in the meta-analysis, subgroup analysis and meta-regression were performed with the following covariates: (i) sample size, (ii) study design, and (iii) diagnostic criteria. Funnel plots are often used to detect publication bias. However, owing to the limitations of funnel plotting, which requires a range of studies of varying sizes involving subjective judgments, publication bias was evaluated using Egger's linear regression test [25], which measures funnel plot asymmetry using a natural logarithm scale of odds ratios (ORs). When asymmetry was indicated, we used the “trim and fill” method to adjust summary estimates for observed bias [26]. This method removes small studies until funnel plot symmetry is achieved by recalculating the center of the funnel before removed studies are replaced with their missing mirror-image counterparts. A revised summary estimate was then calculated using all original studies and hypothetical “filled” studies.

Results

Studies included in the meta-analysis

We identified 82 studies through electronic and manual searching, and 11 studies were selected for a full-text review on the basis of the title and abstract. Three of these studies were excluded because of insufficient data (Supplementary Figure). Thus, eight studies that reported on the diagnostic accuracy of DECT met our inclusion criteria, including a total of 510 patients with gout and 268 controls [8–15]. Of these studies, four studies used the ACR criteria for the diagnosis of gout, and four studies employed the MSU deposition criteria. The characteristic features of the participants in the studies included in the meta-analysis are given in Table 1. The quality assessments of the diagnostic accuracy of the studies showed good results (Supplementary Figure).