CrossMark

Diagnostic accuracy of virtual non-contrast enhanced dual-energy CT for diagnosis of adrenal adenoma: A systematic review and meta-analysis

Michael J. Connolly[1] · Matthew D. F. McInnes[1] · Mohamed El-Khodary[2] · Trevor A. McGrath[3] · Nicola Schieda[1]

Abstract

Objective To compare the diagnostic accuracy of dual-energy (DE) virtual non-contrast computed tomography (vNCT) to non-contrast CT (NCT) for the diagnosis of adrenal adenomas.

Methods Search of multiple databases and grey literature was performed. Two reviewers independently applied inclusion criteria and extracted data. Risk of bias was assessed using QUADAS-2. Summary estimates of diagnostic accuracy were generated and sources of heterogeneity were assessed.

Results Five studies (170 patients; 192 adrenal masses) were included for diagnostic accuracy assessment; all used dual-source dual-energy CT. Pooled sensitivity for adrenal adenoma on vNCT was 54% (95% CI: 47–62%). Pooled sensitivity for NCT was 57% (95% CI: 45–69%). Pooling of specificity was not performed since no false positives were reported. There was a trend for overestimation of HU density on vNCT as compared to NCT which appeared related to contrast timing. Potential sources of bias were seen regarding the index test and reference standard for the included studies. Potential sources of heterogeneity between studies were seen in adenoma prevalence and intravenous contrast timing.

Conclusions vNCT images generated from dual-energy CT demonstrated comparable sensitivity to NCT for the diagnosis of adenomas; however the included studies are heterogeneous and at high risk for some types of bias.

Key points
- *Similar sensitivity of vNCT to NCT for diagnosis of adenoma*
- *Heterogeneity could be related to vNCT from early (<=60 sec) CECT studies*
- *Could not pool specificity as there were no false positives*
- *Small number of heterogeneous studies at high risk of bias*

Introduction

Adrenal nodules are commonly encountered incidental imaging findings, identified in approximately 5% of patients undergoing computed tomography (CT) [1, 2]. While the majority of incidental adrenal nodules are benign adenomas, the risk of malignancy in an incidentally detected adrenal nodule increases substantially in patients with a history of cancer [3, 4]; thus, the imaging characterization of incidentally discovered adrenal nodules is important. Adrenal cortical cells normally contain intracytoplasmic lipid which is a precursor for steroid production and approximately 70% of adenomas will also contain intracytoplasmic lipid [5]. The detection of intracytoplasmic lipid in an adrenal nodule using CT or magnetic resonance imaging (MRI) forms the basis of diagnosis for so called "lipid-rich" adenomas. Lipid-poor adenomas can be diagnosed by their high relative washout rate on a CT study with a delayed

✉ Matthew D. F. McInnes
  mmcinnes@toh.on.ca

1 Clinical Epidemiology Program, Ottawa Hospital Research Institute, University of Ottawa Department of Radiology, Room c159 Ottawa Hospital Civic Campus, 1053 Carling Ave., Ottawa, ON K1Y 4E9, USA

2 McMaster University Department of Radiology, Hamilton, ON, Canada

3 Faculty of Medicine, University of Ottawa, Ottawa, Ontario, Canada

contrast phase [3]. Detection of intracytoplasmic lipid can confirm the diagnosis of adrenal adenoma by using a threshold of 10 Hounsfield units (HU) on non-contrast CT (NCT): sensitivity of 71% and specificity of 98% [5]. Due to this, current practice guidelines recommend that for homogeneous adrenal lesions < 4 cm and with density < 10 HU, no further imaging is required [6].

A frequent dilemma is encountered in clinical practice when an incidental adrenal nodule is detected during a single-phase contrast-enhanced CT (CECT) examination because enhanced adrenal nodule attenuation levels will overlap significantly between adenomas and non-adenoma lesions [7–10]. In this instance, a repeat examination using either dedicated NCT (+/- multiphase adrenal washout protocol) or chemical shift MRI is required for further characterization which is generally performed in conjunction with biochemical testing [11]. Dual-energy CT, which has rapidly become incorporated into many abdominal imaging practices, offers the ability to extract a virtual non-contrast CT (vNCT) data set from routinely acquired variable energy single-phase CECT examination [12]. Analysis of the vNCT data using absolute attenuation thresholds has been suggested as a viable method to diagnose adenomas [13]. Diagnosis of adenomas using vNCT is desirable as a 'replacement test', as this could obviate the need for additional follow-up imaging. To date, studies reporting the diagnostic accuracy of vNCT using absolute attenuation thresholds for diagnosis of adrenal adenomata are variable and have been restricted to single-institution retrospective studies [14]. The purpose of this systematic review is to compare the diagnostic accuracy of vNCT to NCT for the diagnosis of adrenal adenomas.

Methods

This systematic review was written using the guidelines outlined by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [15] and the Cochrane Handbook for Diagnostic Test Accuracy Reviews [16](2)(2)(2). The protocol of this study was registered on PROSPERO (International Prospective Register of Systematic Reviews; protocol number, CRD42015025937) on September 4, 2015 [17](3)(3)(3). No significant protocol deviations occurred.

We proposed the following PICO(S) question: **P**atients with an indeterminate adrenal lesion; **I**ndex test of dual-energy CT used to generate a vNCT series and attenuation levels measured on vNCT used for diagnosis; if direct **C**omparison to standard non-contrast CT was performed in the same patients, this data will be extracted and used as a comparator; **O**utcome was assessed with reference standard of surgery, tissue biopsy or clinical follow up of > 6 months for non-adenomas, clinical parameters or previously described

and validated imaging findings compatible with adenoma including: NCT density <10 HU, visual/quantitative signal intensity decrease using chemical shift MRI [4] or adrenal washout protocol CT; any type of **S**tudy that is not a case report or case series was eligible for inclusion [18].

Search strategy

A literature search using Medline, Embase and the Cochrane Central Register of Controlled Trials (or CCTR) was performed with the assistance of a hospital librarian (A.D.). The final search was performed on October 23, 2015. The search strategy was developed by an abdominal radiologist (M.D.F.M., 9 years of post-residency experience and 5 years of experience conducting systematic reviews). Peer review of the search strategy was performed by an additional abdominal radiologist (N.S., 4 years of post-residency experience) and a hospital librarian (A.D.). No language restrictions were placed on search results. The search strategy is provided in Appendix 1. Retrieved titles and abstracts were independently reviewed by two authors (M.J.C., M.D.F.M.) for relevance (M.J.C. is a resident with 3 years of radiology experience). Full text versions of relevant studies were retrieved for further evaluation. Inclusion criteria were applied to the full-text articles independently by two reviewers (M.J.C., M.D.F.M.). Studies were excluded if any of the inclusion criteria were not met. Reference lists of included studies were checked manually to identify other relevant papers. A search of the grey literature was performed by one of the authors (M.J.C.), up to date as of March 5, 2016, including conference proceedings in 2013–2015 Radiological Society of North America (RSNA) Scientific Meetings, 2015 American Roentgen Ray Society (ARRS) meeting, 2013–2015 European Congress of Radiology (ECR) meetings as well as National Institute of Health (NIH) Clinical Trials database and World Health Organization (WHO) International Clinical Trials Registry Platform. Grey literature search strategy is also provided in Appendix 1.

Inclusion criteria

1. Human patients with an adrenal lesion.
2. Dual-energy CT was performed and a vNCT series was generated as part of the protocol.
3. The data is retrievable to calculate a 2 x 2 contingency table (or sufficient data to calculate either sensitivity or specificity).
4. An acceptable reference standard (described in the 'Outcome' section above) was used for the majority of patients.
5. The study is not a case report, case series, review article or commentary.

6. The study patients are not a subset of patients from another included paper. If studies contain overlapping samples, the duplicate patients will be removed when possible (either by information from the paper or contacting the authors).
7. vNCT applying absolute HU attenuation values is the index test used.

Data extraction and quality assessment

Two authors (M.J.C., M.E-K.) independently extracted data by using data extraction sheets. Data extracted included: Study first author, journal and year of publication, number of patients, patient age (mean, range), number of patients excluded (because of study overlap, insufficient test, no reference standard), technical parameters of dual-energy CT imaging (vendor, dual-energy CT technique [e.g. two consecutive scans at different energies, dual-source/dual-energy, rapid kilovoltage switching or energy-resolving detector] and CT protocol), absolute attenuation threshold used to differentiate adenomas from other adrenal lesions, number of true positive (TP), true negative (TN), false positive (FP) and false negative (FN), number of adenomas, metastases, pheochromocytomas, other lesions and reference standard applied. If standard non-contrast CT was used as a direct comparison in the study, the data relevant to this comparator index test was extracted: technical parameters of CT imaging, absolute attenuation threshold used to differentiate adenomas from other adrenal lesions, number of TP, TN, FP and FN. Disagreements were resolved by consensus. The following definitions were used: TP is a positive index test (vNCT) and diagnosis of adrenal adenoma is confirmed by the reference standard. TN is a negative index test and diagnosis other than adenoma is confirmed by the reference standard. FP is a positive index test for adenoma not confirmed by the reference standard. FN is a negative index test and diagnosis of adenoma is confirmed by the reference standard. Adrenal lesion was defined using the American College of Radiology definition of an adrenal mass ≥ 1 cm [11]. An adrenal lesion containing macroscopic fat such as a myelolipoma measuring < 10 HU was considered to be true negative (not false positive) based on the appearance and density measurement of macroscopic fat. Otherwise, no assumptions were made for HU density for a given pathological diagnosis unless the density was explicitly stated by the authors. Studies with multiple readers with different numbers of TP, FN, FP or TN for adrenal lesions were averaged to obtain 'study-level' data and rounded to the nearest whole number.

Non-diagnostic tests were excluded from the analysis on the assumption that they occur randomly [19]. Samples to which no reference standard was applied were excluded, when possible. One of the authors (M.J.C.) extracted the following additional data: mean, standard deviation and range for adrenal lesion density (HU) on vNCT and NCT, where available. If statistical tests were performed for differences in attenuation values between vNCT and NCT, statistical test type and p-values were reported.

A customized QUADAS-2 (a tool for quality assessment of diagnostic accuracy studies in systematic reviews) tool based on the four domains of study selection, initial reading, reference standard, and 'flow and timing' was used [20]. The customized tool was applied to all studies by two of the authors (M.J.C and M.D.F.M.). Disagreements were resolved by consensus discussion. Appendix 3 provides the QUADAS-2 tool used in this review. The risk for bias across studies (publication bias) was not assessed since there is no generally accepted method for diagnostic accuracy studies and the number of included studies was low [21, 22]. We attempted to minimize publication bias primarily via our search strategy which included multiple databases, conference abstracts and trial registries.

Data analysis

Data extraction was performed using Microsoft Excel (Redmond, WA, USA). 2 x 2 data was summarized in forest plots of sensitivity and specificity for each study, where available, using Meta-DiSc software (Madrid, Spain) [23]. Pooling of data to generate summary estimates was performed for sensitivity only, using the random effects model. The justification for this is that the reported specificity for the included studies was 1; therefore, use of a univariate random effects model (rather than a bivariate one) was felt to be appropriate [24]. If comparator test data was available (standard non-contrast CT), this was displayed and analyzed in a similar manner.

Heterogeneity, in the form of between-studies variance, was quantified using the I^2 index; use of I^2 (a univariate measure) was felt to be appropriate for the same reasons indicated above [21, 22, 24]. Levels of heterogeneity (low <25, moderate 25–75 and high >75) were defined as per Higgins et al. [25]. In addition, sources for heterogeneity were explored, including evaluating the effect of the following (where possible, based on provided data) on diagnostic test accuracy: 1) adenoma prevalence; 2) technical parameters; 3) threshold applied; 4) risk of bias.

Results

Search

Appendix 1 provides details on the search strategy. Figure 1 is a study flow diagram detailing search results and study inclusion. Appendix 2 provides a list of studies excluded at the level of full-text search, with the reason for exclusion. The initial search identified 60 articles, 28 trials and 2 abstracts;

Fig. 1 Flow diagram illustrating search results, study review and study inclusion. RSNA = Radiological Society of North America, ARRS = American Roentgen Ray Society, ECR = European Congress of Radiology, WHO = World Health Organization
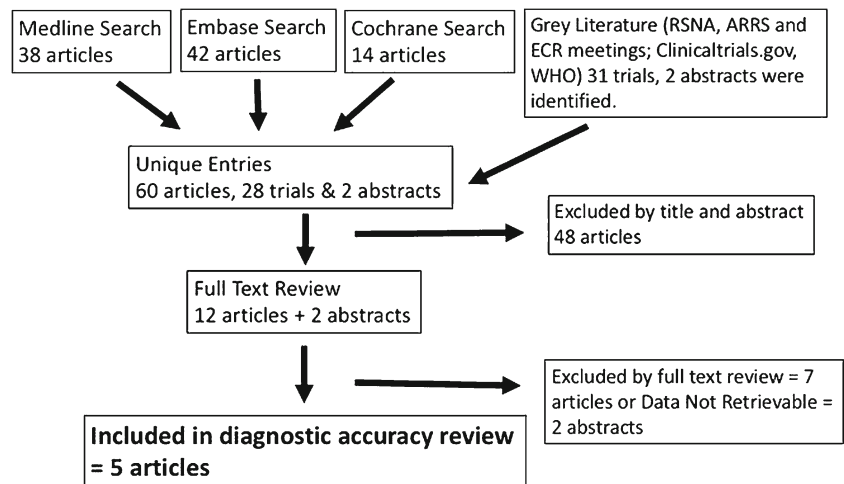
these were screened by title and abstract to 12 articles and 2 abstracts for full-text review. After application of the inclusion criteria, five articles (170 patients; 192 adrenal lesions) were included. All of these studies used dual-source dual-energy CT scanners. [13, 14, 26–28]. All included articles were in English. Retrievable data from Kim et al. [14] only permitted calculation of the sensitivity for the diagnosis of adrenal adenomas.

Data extraction

Table 1 summarizes the characteristics of the included studies. Table 2 summarizes the extracted data from the studies included in the diagnostic accuracy portion of the study. Table 3 summarizes the extracted data and relevant study characteristics from the studies comparing the density measurements of the adrenal lesions on vNCT and NCT with the comparative statistical tests performed and their results. There were statistically significant differences reported in measured HU densities between vNCT and NCT for the 60-second contrast delays in Botsikas et al. and Kim et al. studies. No significant differences were seen for 150 minute delays for Botsikas et al. and Kim et al. or for Gnannt et al. and Ho et al.

Data analysis

Figure 2 is a forest plot of sensitivity data for the diagnosis of adrenal adenoma using vNCT and NCT for the five included studies. Kim et al. and Botsikas et al. performed vNCT using both 60-second and 15-minute intravenous contrast delays, which are plotted separately. The pooled sensitivity for vNCT is 54% (95% CI: 47 to 62%) and heterogeneity I^2 index = 86.4%, which is suggestive of high levels of heterogeneity. NCT data for the adrenal lesions was available for three of the five studies [13, 14, 26]. The pooled sensitivity for NCT is 57% (95% CI: 45 to 69%) and I^2 = 10.1%, consistent with low levels of heterogeneity. Figure 3 is a forest plot of

sensitivities for the subgroup of vNCT where there is corresponding NCT data for each adrenal lesion. The pooled sensitivity for these three studies is 42% (95% CI: 33 to 51%) and I^2 = 74.9%, which may represent moderate heterogeneity. Based on the overlapping 95% CIs, there are no significant differences between the vNCT, subgroup vNCT or NCT sensitivities. Pooling of specificity was not performed since the reported specificity for all studies was 100% (no false positives, see Table 2).

Table 4 depicts the results of the risk of bias assessment of individual studies in the diagnostic accuracy stream when using the modified QUADAS-2 tool. Two areas of concern regarding risk of bias were identified: the index test and reference standard. The risk of bias for the index test was high for Ho et al. since one of the index test interpreters had also reviewed the cases and interpreted the reference standard [13]. For Botsikas et al., Helck et al. and Kim et al., it was unclear if the index test was interpreted without knowledge of the results of the reference standard. The second area of concern was the reference standard. The risk of bias was unclear for four studies due to heterogeneity of the reference standards applied within studies and/or suboptimal reference standards such as clinical follow-up or unenhanced CT only rather than histologic reference standard. The study by Gnannt et al. was felt to be at high risk of bias because NCT (<10 HU defined adenoma) was used as the reference standard for all cases, representing a form of incorporation bias [19, 27].

The I^2 index demonstrated high levels of heterogeneity between studies for pooled sensitivity of vNCT. There was low heterogeneity of importance between studies for pooled sensitivity of NCT. There are several possible reasons for high levels of heterogeneity for the vNCT group. Two important areas could be disease prevalence and technical parameters. The percentage of adrenal masses which were adrenal adenomas in the study population varied considerably from 56–83%. The percentage of adrenal adenomas which measured < 10 HU on NCT varied considerably from 47–70%. From a

Table 1 Characteristics of included studies

Background		Patients				Number of adrenal lesions	Date of acquisition	Study center
Study	Journal of publication (year of publication)	No. of included patients (no. of excluded patients)	Age of patients included: mean (range)/no. of female patients	Study indication for dual-energy CT.				
Botsikas et al.	Eur J Radiol (2014)	19 (1*)	Not available/not available	Clinically suspected adrenal lesion or incomplete characterization of adrenal lesion on CT or ultrasound.	24	Sept. 2010 – March 2012	Geneva University Hospital, Department of Imaging and Medical Information Sciences, Genève, Switzerland
Gnannt et al.	AJR (2012)	32 (0)	Not available/Not Available	Follow up after endovascular aortic aneurysm repair.	39	Not stated	Institute of Diagnostic and Interventional Radiology, University Hospital Zurich, Zurich, Switzerland.
Helck et al.	Eur Radiol (2014)	51 (0)	70 y (24-85 y)/23 F	Patient database search for dual-energy AND adrenal lesion OR adenoma OR incidentaloma OR adrenal metastasis.	57	April 2007 – Sept. 2012	Institute for Clinical Radiology University of Munich Grosshadern Campus, Munich, Germany
Ho et al.	AJR (2012)	19 (2**)	65 y (38-84 y)/8 F	Renal cell carcinoma follow-up, indeterminate renal mass, hematuria, ureteral stricture, recurrent urinary infections, prostate cancer, hypertension and renal angiomyolipoma.	23	Nov. 26, 2007 – June 22, 2010	Dept. of Radiology, Duke University, Durham, NC, USA
Kim et al.	Radiology (2013)	49 (0)	56.8 y (35-74 y)/16 F	Assess suspicious adrenal masses detected on routine abdominal CT images.	49	December 2009 – June 2010	Dept. of Radiology and Center for Imaging Science, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea

Background		Technique				
Study	CT Vendor/model(s)	Dual-energy technique used	IV Contrast type, dose and injection rate	Scan delay(s) for dual-energy acquisitions	Dual-Energy CT Postprocessing and Image Reconstruction Software	Non-contrast technique used
Botsikas et al.	Siemens SOMATOM Definition Flash (Siemens Healthcare, Erlangen, Germany)	Dual-source (tube A: 140 kVP, 4D dose modulation (Siemens Medical Solutions, Forcheim, Germany),	Iohexol 350 mg/mL (Accupaque 350, GE Healthcare, Little Chafford, UK), 1.5 mL/kg at 3.5 mL/s	1. 60-second delay 2. 15-minute delayed	Syngo liver VNC (Siemens Medical Solutions, Erlangen, Germany)	Dual-energy 120-kVp weighted average

Table 1 (continued)

Background Study	Technique CT Vendor/model(s)	Dual-energy technique used	IV Contrast type, dose and injection rate	Scan delay(s) for dual-energy acquisitions	Dual-Energy CT Postprocessing and Image Reconstruction Software	Non-contrast technique used
Gnannt et al.	Siemens SOMATOM Definition (Siemens Healthcare)	Dual-source (tube A: 140 kVp, 95 mAs, tube B: 80 kVp, 400 mAs)	Iopromide 300 mg/mL (Ultravist 300, Bayer Schering Pharma), 100 mL at 4 mL/s.	70-second delay	Syngo liver VNC (Siemens Medical Solutions, Forcheim Germany)	Single-energy 120 kVp, 350 mAs
Helck et al.	1. Siemens SOMATOM Definition 2. Siemens SOMATOM Definition Flash	1. Dual-source (tube A: 140 kVp, CareDOSE 4D (Siemens Medical Solutions, Forcheim, Germany), tube B: 80 kVp, CareDOSE 4D) 2. Dual-source (tube A: Sn140 kVp, CareDOSE 4D, tube B:100 kVp, CareDOSE 4D)	Iopromide 370 mg/mL (Ultravist 370, Bayer Schering Diagnostics, Berlin, Germany), 1.35 mL/kg, injection rate not stated.	70-second delay	Syngo liver VNC (Siemens Medical Solutions, Forcheim Germany)	Dual-energy 120 kVp weighted average [1. 70:30 tube ratio using CareDOSE 4D (Siemens), 2. 50:50 using CareDOSE 4D]
Ho et al.	Siemens SOMATOM Definition	Dual-source (tube A: 140 kVp, 126 mAs, tube B: 80 kVp, 499 mAs)	150 mL of iopamidol, Isovue 300, (Bracco Diagnostics) at 3 mL/sec	100-second delay	Syngo liver VNC (Siemens Medical Solutions, Forcheim Germany)	Single-energy 120 kVp, 250 mAs.
Kim et al.	Siemens SOMATOM Definition Flash	Dual-source (tube A: 140 kVp, 150 mAs, tube B: 80 kVp 281 mAs).	120 mL of Ultravist 370 (Bayer) at 3 mL/sec	1 minute and 15 minutes	Syngo liver VNC (Siemens Medical Solutions, Forcheim Germany)	Single-energy 120 kVp, 180mAs. Care Dose4D

Tube B: 100 kVp, 250 mAs)

* Non-diagnostic CT due to motion

** 1 patient non-diagnostic CT due to morbid obesity and reconstruction artifacts, 1 patient lost dual-energy data from improper data recording

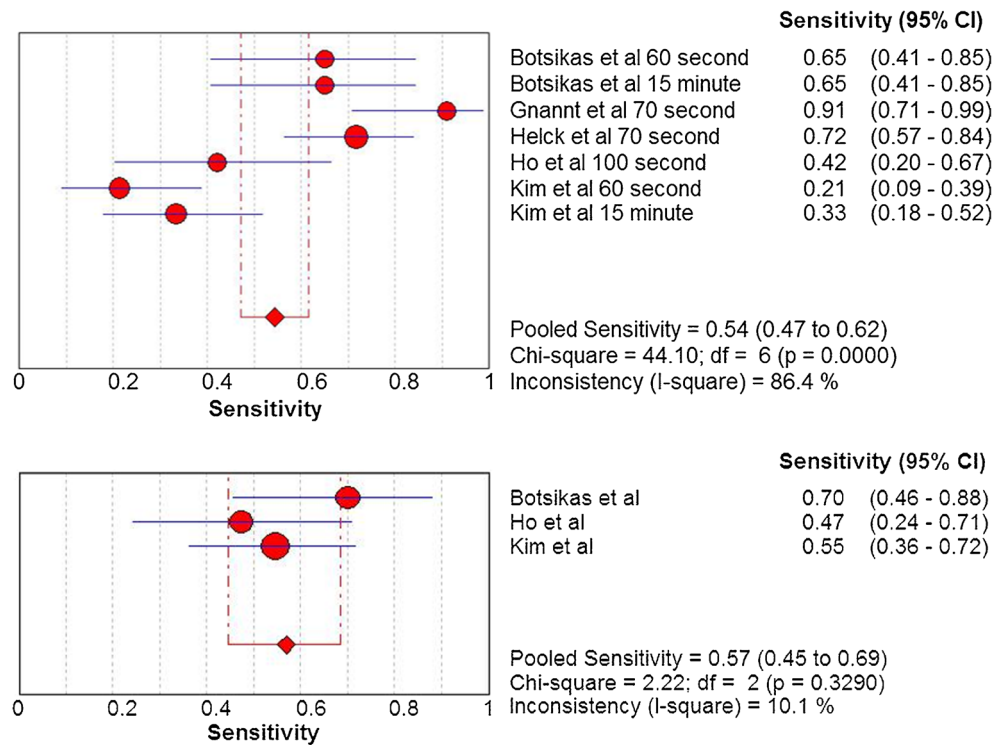**Table 2** Results for studies included in the diagnostic accuracy analysis

Study	No. of adrenal lesions	No. of TP on vNCT	No. of FP on vNCT	No. of FN on vNCT	No. of TN on vNCT	No. of TP on NCT	No. of FP on NCT	No. of FN on NCT	No. of TN on NCT	Reference standard(s) [no. of lesions applied to]
Botsikas et al. 1. 60 seconds 2. 15 minutes	24	13 13	0 0	7 7	4 4	14	0	6	4	-NCT<0 HU (5) -Size stability (14) -Histology (5)
Gnannt et al.	39	20*	0	2*	17	N/A	N/A	N/A	N/A	NCT (39)
Helck et al.	57	33	0	13	11	N/A	N/A	N/A	N/A	-Clinical follow-up ≥ 18 months(12) -NCT (26) -PET/CT (5) -IP-OOP MRI (5) -Histology (9)
Ho et al.	23	8	0	11*	4*	9	0	10	4	-Size stability (13) -NCT≤10 HU (9) -IP-OOP MRI (1)
Kim et al. 1. 60 seconds 2. 15 minutes	49	7 11	N/A N/A	26 22	N/A N/A	18	N/A	15	N/A	-Histology (10) -Size stability, CTrotocol, biochemical tests (39)

Notes: vNCT = virtual non-contrast CT, NCT = non-contrast CT, TP = true positive, FP = false positive, FN = false negative, TN = true negative. IP-OOP = In phase-out of phase. *Study which had multiple readers with different numbers of TP, FN, FP or TN for adrenal lesions were averaged and rounded to nearest whole number.

Table 3 Summary of reported statistical comparisons of vNCT with respect to contrast timing

Study		Botsikas et al. EJR 2014	Gnannt et al. AJR 2012	Helck et al. Eur Radiol 2014	Ho et al. AJR 2012	Kim et al. Radiology 2013
vNCT Contrast delay(s)		1. 60 seconds 2. 15 minutes	1. 70 seconds 2. N/A	1. 70 seconds 2. N/A	1. 100 seconds 2. N/A	1. 60 seconds 2. 15 minutes
Mean (+/- std. dev.; HU)	NCT	8.6 (+/- 15.2)	6.6 (+/- 18.5)	N/A	12.9 (+/- 13.4)	13.3 (+/- 22)
	vNCT Contrast delay 1	12.6 (+/- 14.9)	7.9 (+/- 18.3)	N/A	14.7 (+/- 15.1)	21 (+/- 22)
	vNCT Contrast delay 2	8.5 (+/- 13.6)	N/A	N/A	N/A	15.9 (+/-20.6)
Range (HU)	NCT	-15.7 to 38.6	-82.8 to 30.9	N/A	N/A	-103 to 40
	vNCT Contrast delay 1	-15 to 41.6	-76.6 to 30.2	-67.4 to 48.1	N/A	-93 to 49
	vNCT Contrast delay 2	-13.1 to 36	N/A	N/A	N/A	-94 to 45
Contrast type, volume & rate		GE Accupaque 350, 1.5 mL/kg at 3.5 mL/s	Bayer Ultravist 300, 100 mL at 4 mL/ second	Bayer Ultravist 370, 1.35 mL/kg. Rate not stated.	Isovue 300, 150 mL at 3 mL/second	Bayer Ultravist 370, 120 mL at 3 mL/second
No. of patients (adrenal lesions)		19 (24)	32(39)	51(57)	19(23)	49 (49)
Statistical test used to compare NCT with vNCT HU values		One-way ANOVA with Dunnett's multiple comparisons test.	Paired student's t test with post hoc Bonferroni correction	N/A	Paired student's t test	Paired t test with Bonferroni correction
P value	NCT versus vNCT contrast delay 1	0.020	0.87	N/A	0.2	0.001
	NCT versus vNCT contrast delay 2	0.9974	N/A	N/A	N/A	0.119

Fig. 2 Forest plot of sensitivity for vNCT (*top*) and NCT (*bottom*) for diagnosis of adrenal adenoma. Included studies are listed on the right with summary accuracy measurements, 95% CI and pooled sensitivity listed on the right

Sensitivity (95% CI)

Botsikas et al 60 second	0.65	(0.41 - 0.85)
Botsikas et al 15 minute	0.65	(0.41 - 0.85)
Gnannt et al 70 second	0.91	(0.71 - 0.99)
Helck et al 70 second	0.72	(0.57 - 0.84)
Ho et al 100 second	0.42	(0.20 - 0.67)
Kim et al 60 second	0.21	(0.09 - 0.39)
Kim et al 15 minute	0.33	(0.18 - 0.52)

Pooled Sensitivity = 0.54 (0.47 to 0.62)
Chi-square = 44.10; df = 6 (p = 0.0000)
Inconsistency (I-square) = 86.4 %

Sensitivity (95% CI)

Botsikas et al	0.70	(0.46 - 0.88)
Ho et al	0.47	(0.24 - 0.71)
Kim et al	0.55	(0.36 - 0.72)

Pooled Sensitivity = 0.57 (0.45 to 0.69)
Chi-square = 2.22; df = 2 (p = 0.3290)
Inconsistency (I-square) = 10.1 %

technical standpoint, there were differences in contrast timing between studies with Botsikas et al. and Kim et al. using both 60-second and 15-minute post contrast administration; Gnannt et al. and Helck et al. used 70 seconds and Ho et al. used 100 seconds. Technical parameters were otherwise comparable between studies. All studies used a cut-off of < 10 HU on vNCT for the diagnosis of adrenal adenoma.

Due to the small number of included studies, subgroup analysis or meta-regression for such things as contrast timing or adenoma prevalence could not be performed.
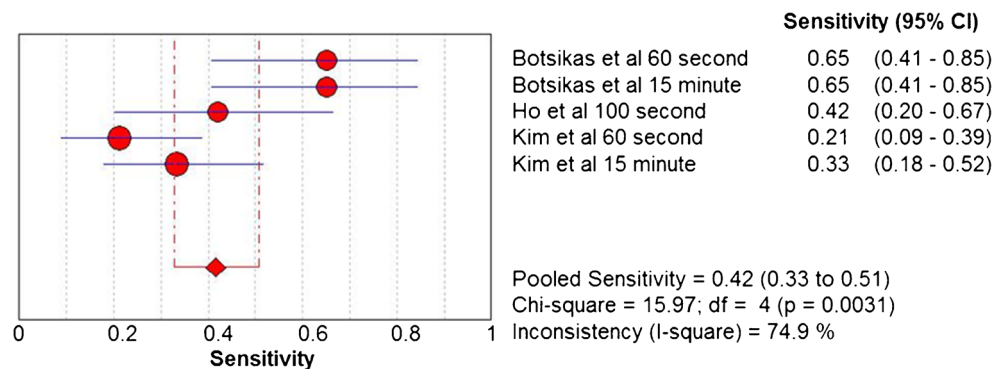
Discussion

In this systematic review and meta-analysis, we found that vNCT images generated from dual-energy CT demonstrated

comparable sensitivity to NCT for the diagnosis of adenomas. These findings are of potential importance because diagnosis of adenomas using vNCT alone could prevent additional follow-up imaging studies and reduce cost and cumulative radiation dose (if NCT is pursued as the next imaging test) to the patient. However, some important limitations should raise caution when interpreting the data, including a high risk of some forms of bias and the number of included studies is quite low; as such, our study may be underpowered to detect differences in sensitivity (there was a potential trend for lower sensitivity for vNCT). In addition, planned pooling of specificity data could not be performed since no FP cases were reported.

High heterogeneity for sensitivity of vNCT was identified (and not for NCT). This may be due to potential overestimation of the density in HU on vNCT as compared to NCT. The

Fig. 3 Subgroup forest plot of sensitivities for vNCT for diagnosis of adrenal adenoma on studies that have corresponding NCT data

Sensitivity (95% CI)

Botsikas et al 60 second	0.65	(0.41 - 0.85)
Botsikas et al 15 minute	0.65	(0.41 - 0.85)
Ho et al 100 second	0.42	(0.20 - 0.67)
Kim et al 60 second	0.21	(0.09 - 0.39)
Kim et al 15 minute	0.33	(0.18 - 0.52)

Pooled Sensitivity = 0.42 (0.33 to 0.51)
Chi-square = 15.97; df = 4 (p = 0.0031)
Inconsistency (I-square) = 74.9 %

Table 4 Results of QUADAS-2 assessment for risk of bias assessment

Risk of bias

	Patient selection	Index test	Reference standard	Flow and timing
Botsikas et al. EJR 2014	Low	Unclear	Unclear	Low
Gnannt et al. AJR 2012	Low	Low	High	Low
Helck et al. Eur Radiol 2014	Low	Unclear	Unclear	Low
Ho et al. AJR 2012	Low	High	Unclear	Low
Kim et al. Radiology 2013	Low	Unclear	Unclear	Low

degree of overestimation may depend on the intravenous contrast timing of the vNCT. There were statistically significant differences between measured density in HU of adrenal lesions for Botsikas et al. and Kim et al. when comparing the 60-second intravenous contrast delay vNCT and NCT but not between the 15-minute delay vNCT and NCT. There were no statistically significant differences between Gnannt et al. and Ho et al. for vNCT and NCT which had contrast delays of 70 seconds and 100 seconds, respectively. Adrenal adenomas characteristically show early enhancement followed by washout of contrast compared to surrounding tissues which forms the basis of adrenal CT washout studies [10, 29]. A previous study by Szolar and Kammerhuber demonstrated adrenal adenoma enhancement peaking around 60 seconds following intravenous contrast administration [29]. Peak adenoma enhancement at 60 seconds followed by rapid washout could explain why there was overestimation of attenuation values for vNCT derived from 60-second acquisition data that becomes less pronounced with derivation of vNCT from more delayed contrast timing.

Previous studies have consisted of single-center, retrospective data and compared vNCT to NCT using a variety of methods [13, 14, 26–28]. We studied what we believe to be the most important measures of vNCT for adrenal adenomas: diagnostic accuracy and head-to-head comparison of HU density measurements. By combining the data from these studies to create a larger patient population and focusing on these important metrics, we were able to provide better insight into whether vNCT on incidentally discovered adrenal masses has the potential to reduce the number of follow-up chemical shift MRI or NCT studies performed to further characterize nodules detected incidentally at CECT. We have demonstrated that differences measured in HU for vNCT compared to NCT exist and may be attributed to timing of the contrast enhancement series from which the vNCT was derived.

Our study had several limitations. Studies utilizing rapid-switching or detector-based dual-energy CT scanners were unable to be included due to the absence of any studies evaluating adrenal lesions and using HU values measured on vNCT. In the case of rapid-switching dual-energy CT, measurement of HU on vNCT has only recently been described using multimaterial decomposition and has only been studied in normal anatomic structures [13, 14, 30]. Only five studies, all of which were retrospective and utilized dual-source dual-energy CT scanners, met the inclusion criteria. Heterogeneity between the studies' intravenous contrast timing and adenoma prevalence (specifically the percentage of adenomas measuring < 10 HU on NCT) could affect the measured sensitivity. For the 3 studies which provided NCT values, only 57% of their adrenal adenomas measured < 10 HU on NCT [13, 14, 26]. This is less than the widely quoted value of 70% of adrenal adenomas being 'lipid-rich' or measuring < 10 HU on NCT [5] and leads to a lower sensitivity for vNCT on those same studies due to lipid-rich adrenal adenoma prevalence. This could be due to a selection bias in these studies (i.e. indeterminate lesions were more likely to be investigated). The QUADAS-2 assessment identified potential or unclear risks of bias in the index test and reference standard which could affect the calculated specificity, sensitivity and measured HU density. Lastly, with the exception of Botsikas et al., we were reliant on published values and unable to access the raw data for the included studies which could have led to more data for pooling of specificity of vNCT, NCT and head-to-head comparison of measured HU of adrenal masses for vNCT and NCT.

Our study demonstrates that virtual NCT has comparable sensitivity to true NCT when using attenuation thresholds for the diagnosis of adrenal adenoma. We were not able to pool specificity results across studies as there were no reported false positive diagnoses of adenoma using vNCT (or NCT) in the evaluated studies. Because we identified only a small number of studies with concerns about high risk of some forms of bias, we suggest that additional prospective, multi-center trials with an acceptable reference standard to validate the diagnostic accuracy of vNCT for diagnosis of adrenal adenoma are required. The heterogeneity with regard to the sensitivity of vNCT observed in our study may relate to the timing of the post-contrast acquisition from which the vNCT data set was derived. Studies performed at 60 seconds resulted in higher vNCT attenuation values compared to their true NCT attenuation values. This important observation will require further analysis but may suggest a limitation of vNCT for diagnosis of adrenal adenoma using data derived from early post-contrast enhanced acquisitions.

Appendix 1. Search strategies

Embase: 44
Medline: 38
Cochrane: 14
Total with duplicates removed, up to date as of October 23, 2015: 61 articles
NIH Clinical Trials Database: 27
WHO Clinical Trials Registry Platform: 4
Total with duplicates removed, up to date as of March 6, 2016: 28 trials
2013-2015 RSNA, 2015 ARRS, and 2013-2015 ECR Scientific Meetings, up to date as of March 6, 2016: 2 abstracts

Cochrane - Date Run:	23/10/15 13:05:04.121	
ID	Search	Hits
#1	adrenal:ti,ab,kw (Word variations have been searched)	4908
#2	MeSH descriptor: [Tomography, Emission-Computed] explode all trees	2678
#3	(ct or computed tomograph*):ti,ab,kw (Word variations have been searched)	66064
#4	#2 or #3	66711
#5	dual energy:ti,ab,kw	2568
#6	#4 and #5	644
#7	dect:ti,ab,kw	10
#8	#6 or #7	644
#9	#1 and #8	14

Cochrane Central Register of Controlled Trials : Issue 9 of 12, September 2015 = 14 results
Database: Embase <1974 to 2015 October 22 > Search Strategy:

1	adrenal.mp. (170288)
2	exp computer assisted tomography/(688233)
3	(ct or computed tomograph$).tw. (493269)
4	2 or 3 (838139)
5	dual energy.tw. (21990)
6	4 and 5 (4187)
7	dual energy computer assisted tomography/or dect.tw. (757)
8	6 or 7 (4277)
9	1 and 8 (44)

Pubmed (October 23, 2015)
(adrenal[Text Word]) AND ((((((dual energy) OR Dual-Energy Scanned Projection[MeSH Terms])) AND (((CT[Title/Abstract]) OR computed tomograph*[Title/Abstract]) OR Tomography, X-Ray Computed[MeSH Terms])) OR dect[Title/Abstract]) = 38

Clinical trials

Up to date as of March 5, 2016
ClinicalTrials.gov, a service of the U.S. National Institutes of Health
https://www.clinicaltrials.gov/ct2/home
Search: "adrenal AND computed tomography"
27 studies found.

WHO International Clinical Trials Registry Platform
http://apps.who.int/trialsearch/
Search: "adrenal AND computed tomography"
4 trials found.

Scientific abstracts

Up to date as of March 5, 2016. Titles and abstracts highlighted by search criteria below were screened; abstracts potentially meeting inclusion criteria were counted.

RSNA 2015 MEETING- search "dual-energy"
https://rsna2015.rsna.org/program/?PAGE=7
2 abstracts

RSNA 2014 MEETING- search "dual-energy"
http://rsna2014.rsna.org/program/?PAGE=8
0 abstracts

RSNA 2013 MEETING- search "dual-energy"
http://rsna2013.rsna.org/program/?PAGE=6
0 abstracts

ARRS 2015 MEETING- search "dual-energy"
http://cf.arrs.org/abstracts/
0 abstracts

ECR 2015 MEETING-
https://www.myesr.org/cms/website.php?id=/en/ecr_2015/ecr_2015_boa.htm
Search "Adrenal"-
0 abstracts

ECR 2014 MEETING-
https://www.myesr.org/cms/website.php?id=/en/ecr_2014/ecr_2014_book_of_abstracts.htm
Search "Adrenal"
0 abstracts

ECR 2013 MEETING-
https://www.myesr.org/cms/website.php?id=/30047/en/ecr_2013/book_of_abstracts.htm
Search "Adrenal"
0 abstracts

Appendix 2

Table 5 Exclusions made at the full text review level (including abstracts) with reason for exclusion

First author surname	Journal/Meeting	Publication/Meeting year	Reason for exclusion
Bonci	RSNA	2015	Data not retrievable for a 2 x 2 table.
Glazer	AJR	2014	vNECT did not use Hounsfield unit attenuation values.
Gupta	AJR	2010	vNECT series was not generated (no IV contrast material administered).
Ju	Acad Radiol	2015	vNECT did not use Hounsfield unit attenuation values.
Khagani	J Med Imaging Radiat Oncol	2013	Data not retrievable for a 2 x 2 table.
Mileto	Radiology	2015	vNECT did not use Hounsfield unit attenuation values.
Morgan	J Comput Assisted Tomogr	2013	vNECT did not use Hounsfield unit attenuation values.
Shi	Acta Radiol	2014	vNECT series was not generated (no IV contrast material administered).
Shitong	RSNA	2015	Data not retrievable for a 2 x 2 table.

Appendix 3

Table 6 Quality assessment tool for diagnostic accuracy studies 2

Domain 1: Patient selection	
Was a consecutive or random sample of patients enrolled?	Yes/No/Unclear
Was a case-control design avoided?	Yes/No/Unclear
Did the study avoid inappropriate exclusions?	Yes/No/Unclear
Could the selection of patients have introduced bias?	RISK: LOW/HIGH/UNCLEAR
Domain 2: Index test	
Were the index test results interpreted without knowledge of the results of the reference standard?	Yes/No/Unclear
If a threshold was used, was it pre-specified?	Yes/No/Unclear
Could the conduct or interpretation of the index test have introduced bias?	RISK: LOW/HIGH/UNCLEAR
Domain 3: Reference standard	
Is the reference standard likely to correctly classify the target condition?	Yes/No/Unclear
Were the reference standard results interpreted without knowledge of the results of the index test?	Yes/No/Unclear
Could the reference standard, its conduct or its interpretation have introduced bias?	RISK: LOW/HIGH/UNCLEAR
Domain 4: Flow and timing	
Was there an appropriate interval between index test(s) and reference standard?	Yes/No/Unclear
Did all patients receive a reference standard?	Yes/No/Unclear
Did patients receive the same reference standard?	Yes/No/Unclear
Were all patients included in the analysis?	Yes/No/Unclear
Could the patient flow have introduced bias?	RISK: LOW/HIGH/UNCLEAR

References

1. Song JH, Chaudhry FS, Mayo-Smith WW (2008) The incidental adrenal mass on CT: prevalence of adrenal disease in 1,049 consecutive adrenal masses in patients with no known malignancy. AJR Am J Roentgenol 190(5):1163–1168

2. Glazer HS, Weyman PJ, Sagel SS, Levitt RG, McClennan BL (1982) Nonfunctioning adrenal masses: incidental discovery on computed tomography. AJR Am J Roentgenol 139(1):81–85

3. Schieda N, Al Dandan O, Kielar AZ, Flood TA, McInnes MD, Siegelman ES (2014) Pitfalls of adrenal imaging with chemical shift MRI. Clin Radiol. doi:10.1016/j.crad.2014.06.020

4. Siegelman ES (2012) Adrenal MRI: techniques and clinical applications. J Magn Reson Imaging 36(2):272–285

5. Boland GW, Lee MJ, Gazelle GS, Halpern EF, McNicholas MM, Mueller PR (1998) Characterization of adrenal masses using unenhanced CT: an analysis of the CT literature. AJR Am J Roentgenol 171(1):201–204

6. Fassnacht M, Arlt W, Bancos I, Dralle H, Newell-Price J, Sahdev A et al (2016) Management of adrenal incidentalomas: European society of endocrinology clinical practice guideline in collaboration with the European network for the study of adrenal tumors. Eur J Endocrinol 175(2):G1–G34

7. Patel J, Davenport MS, Cohan RH, Caoili EM (2013) Can established CT attenuation and washout criteria for adrenal adenoma accurately exclude pheochromocytoma? AJR Am J Roentgenol 201(1):122–127

8. Park BK, Kim B, Ko K, Jeong SY, Kwon GY (2006) Adrenal masses falsely diagnosed as adenomas on unenhanced and delayed contrast-enhanced computed tomography: pathological correlation. Eur Radiol 16(3):642–647

9. Park BK, Kim CK, Kwon GY, Kim JH (2007) Re-evaluation of pheochromocytomas on delayed contrast-enhanced CT: washout enhancement and other imaging features. Eur Radiol 17(11): 2804–2809

10. Schieda N, Alrashed A, Flood TA, Samji K, Shabana W, McInnes MD (2016) Comparison of Quantitative MRI and CT washout analysis for differentiation of adrenal pheochromocytoma from adrenal adenoma. AJR Am J Roentgenol 206(6):1141–1148

11. Choyke PL, Criteria ACoA (2006) ACR appropriateness criteria on incidentally discovered adrenal mass. J Am Coll Radiol 3(7):498–504

12. Marin D, Boll DT, Mileto A, Nelson RC (2014) State of the art: dual-energy CT of the abdomen. Radiology 271(2):327–342

13. Ho LM, Marin D, Neville AM, Barnhart HX, Gupta RT, Paulson EK et al (2012) Characterization of adrenal nodules with dual-energy CT: can virtual unenhanced attenuation values replace true unenhanced attenuation values? AJR Am J Roentgenol 198(4): 840–845

14. Kim YK, Park BK, Kim CK, Park SY (2013) Adenoma characterization: adrenal protocol with dual-energy CT. Radiology 267(1): 155–163

15. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 6(7), e1000097. doi:10.1371/journal.pmed.1000097

16. Deeks J, Bossuyt P, Gatsonis C (2013) Cochrane handbook for systematic reviews of diagnostic test accuracy 1.0.0 ed: The cochrane collaboration

17. York Uo. International prospective register of systematic reviews (PROSPERO) 2014 [November 14, 2014]. Available from: http://www.crd.york.ac.uk/PROSPERO/.

18. Staunton M (2007) Evidence-based radiology: steps 1 and 2–asking answerable questions and searching for evidence. Radiology 242(1):23–31

19. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J (2003) The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol 3:25

20. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB et al (2011) QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 155(8): 529–536

21. McInnes MD, Bossuyt PM (2015) Pitfalls of systematic reviews and meta-analyses in imaging research. Radiology 277(1):13–21

22. Macaskill P GC, Deeks JJ, Harbord RM, Takwoingi Y (2010) Chapter 10: Analysing and Presenting Results. In: Deeks JJ, Bossuyt PM, Gatsonis C (editors), Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy The Cochrane Collaboration

23. Zamora J, Abraira V, Muriel A, Khan K, Coomarasamy A (2006) Meta-DiSc: a software for meta-analysis of test accuracy data. BMC Med Res Methodol 6:31

24. McGrath T, McInnes M, Korevaar D, Bossuyt P (2016) Meta-analyses of diagnostic accuracy in imaging journals: analysis of pooling techniques and their effect on summary estimates of diagnostic accuracy. Radiology 281(1):78–85

25. Higgins JP, Thompson SG, Deeks JJ, Altman DG (2003) Measuring inconsistency in meta-analyses. BMJ 327(7414):557–560

26. Botsikas D, Triponez F, Boudabbous S, Hansen C, Becker CD, Montet X (2014) Incidental adrenal lesions detected on enhanced abdominal dual-energy CT: can the diagnostic workup be shortened by the implementation of virtual unenhanced images? Eur J Radiol 83(10):1746–1751

27. Gnannt R, Fischer M, Goetti R, Karlo C, Leschka S, Alkadhi H (2012) Dual-energy CT for characterization of the incidental adrenal mass: preliminary observations. AJR Am J Roentgenol 198(1): 138–144

28. Helck A, Hummel N, Meinel FG, Johnson T, Nikolaou K, Graser A (2014) Can single-phase dual-energy CT reliably identify adrenal adenomas? Eur Radiol 24(7):1636–1642

29. Szolar DH, Kammerhuber FH (1998) Adrenal adenomas and nonadenomas: assessment of washout at delayed contrast-enhanced CT. Radiology 207(2):369–375

30. Kaza RK, Raff EA, Davenport MS, Khalatbari S (2016) Variability of CT attenuation measurements in virtual unenhanced images generated using multimaterial decomposition from fast kilovoltage-switching dual-energy CT. Acad Radiol. doi:10.1016/j.acra.2016.09.002