Coronary artery calcium
detection & quantification
with current and next-generation CT:
it is time for a change

Niels R. van der Werf

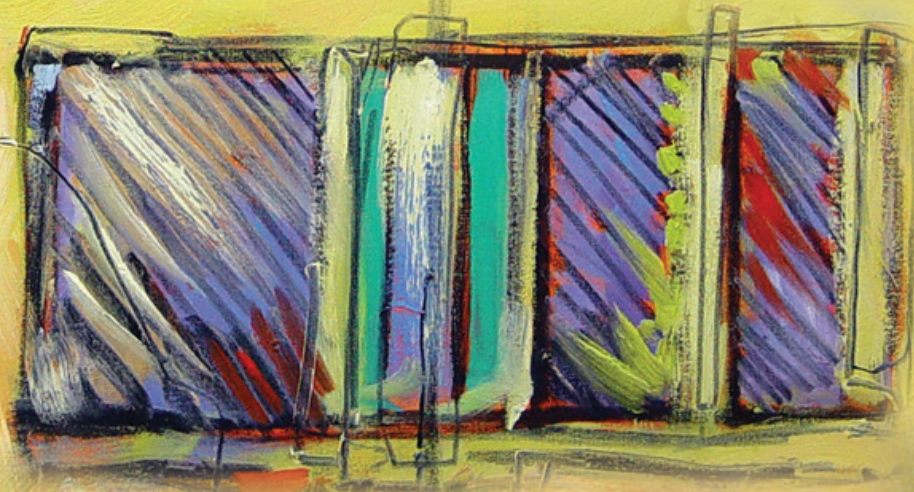# Coronary artery calcium detection & quantification with current and next-generation CT
It is time for a change

Niels R. van der Werf

Coronary artery calcium detection & quantification with current and next-generation CT: *it is time for a change*

Coronary artery calcium detection & quantification with current and next-generation CT
It is time for a change

Kransslagader calcium detectie en kwantificatie met huidige en toekomstige CT technieken
Het is tijd voor verandering
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

woensdag 28 september 2022 des middags te 4.15 uur

door

Niels Roderik van der Werf

geboren op 22 januari 1988
te Groningen

Promotor
   Prof. dr. T. Leiner

Copromotoren
   Dr. M.J.W. Greuter
   Dr. M.J. Willemink

Beoordelingscommissie
   Prof. dr. M.L. Bots
   Prof. dr. N. Buls
   Prof. dr. ir. H.A.W.M. de Jong
   Prof. dr. A. van der Lugt
   Prof. dr. B.K. Velthuis

CONTENTS

Part III – Post-processing parameters

CHAPTER 1

**General introduction
and outline of the thesis**

GENERAL INTRODUCTION

According to the World Health Organization, ischemic heart disease remains the main cause of death worldwide.[1] Moreover, the relative number of deaths from this type of cardiovascular disease is rising, especially in developed countries. Ischemic heart disease is most commonly caused by the occlusion of one or more main coronary arteries. These occlusions occur when atherosclerotic lesions prevent adequate blood flow through the coronary artery.[2] A common feature of atherosclerosis is ectopic bone formation, whereby calcified structures are formed in the coronary arteries.[3]

Early detection of atherosclerosis allows for appropriate medical treatment strategy. For selected asymptomatic individuals, computed tomography (CT) coronary artery calcium (CAC) detection and quantification is recommended by several guidelines to improve clinical risk prediction.[4–7] The reason for this interest in CAC detection and quantification is related to the strong relationship between CAC at CT, total coronary atherosclerosis, and future adverse cardiovascular events.[8] In addition, risks of adverse events increases with increasing CAC quantities over multiple subsequent CT scans, which is closely associated with progression of atherosclerosis.[9]

Clinically, the CAC quantification methodology described by Agatston et al in 1990 remains the reference standard.[9,10] However, patient specific parameters are known to hamper stable CAC quantification using the Agatston score methodology. One of these parameters is patient size, which may reduce Agatston scores by 50% for larger patients.[11] Another parameter which introduces variability in Agatston scores is related to motion of the coronary arteries during the CAC CT acquisition.[12,13] The extent of motion is directly related to the heart rate.[14,15]

To further optimize CAC detection and quantification, many technical advances in CT, which were developed over the last three decades, are available. These technical advances include for example tube voltage reduction, X-ray spectrum shaping with tin filtration, novel reconstruction algorithms like iterative or deep learning reconstruction, thin slice acquisitions or acquisitions with photon counting detectors (PCD).[5,16–18] However, the Agatston scoring method's original parameters (3 mm slice thickness, 120 peak kilovolt (kVp)) – which are in fact not even based on CT data, but on electron beam tomography (EBT) data – are still used in daily practice. In turn, modern CT technology has the potential to reduce variability in CAC scores and reduce radiation dose burden, however this potential is currently not realized.

OUTLINE

The general aim of the current thesis is to further improve diagnostic performance of CAC identification and quantification with CT. First, sources of CAC scoring variability are identified and studied. These sources include patient specific factors, acquisition & reconstruction parameters, and post-processing of reconstructed images (*Figure 1*). For each of these, solutions are presented to reduce CAC scoring variability, improve the prognostic value of CT examinations and/or reduce radiation dose exposure for patients by employing novel technical improvements for CT.
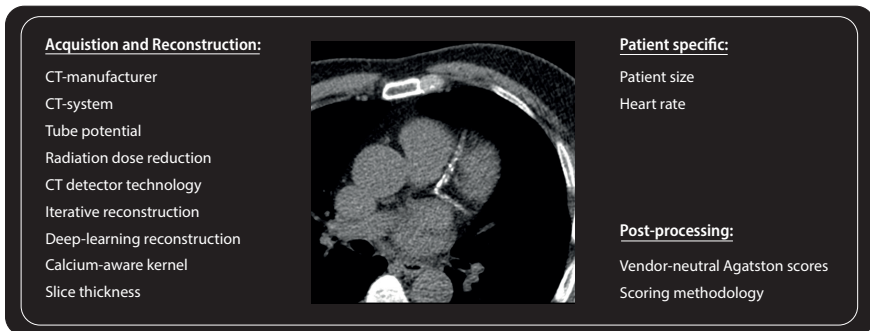
Figure 1 Potential sources of variation in coronary artery calcium (CAC) assessment which were assessed in this thesis. At the center, an example CAC CT image is shown

This thesis consists of three parts, whereby each part is related to one of the before mentioned source of CAC scoring variability. Part 1 is related to patient specific factors. In **chapter 2** the influence of heart rate on CAC scores obtained with state-of-the-art CT systems from all main CT manufacturers is systematically assessed. In **chapters 3 & 4** CT motion artefacts for CAC of different densities are corrected for with the use of a deep learning algorithm.

Part 2 focusses on both acquisition and reconstruction parameters. First, in **chapter 5**, a systematic literature review is presented which explores several dose reduction techniques used for CAC scoring. In **chapter 6**, a technical argument for a new CAC scoring method is presented. In **chapters 7 & 8** the influence of iterative reconstruction on CAC scores is assessed for moving CAC. First, changes in CAC scores for different levels of iterative reconstruction (IR) at different heart rates are

1

studied. Second, the influence of the combination of dose reduction and iterative reconstruction on CAC scores is assessed. In **chapters 9 & 10** a novel photon counting CT (PCCT) detector system is evaluated. First, the CAC scoring potential of this new system is compared with a conventional CT system. This conventional CT scanner uses energy-integrating detector (EID) elements for the detection of incoming x-ray photons. Subsequently, the radiation dose reduction potential of PCCT for CAC scoring is assessed, again in comparison with conventional EID CT. In **chapters 11 & 12**, a calcium-aware reconstruction kernel is assessed. This kernel potentially enables using patient-specific tube voltages to acquire raw data. In the reconstruction process, the resulting CT numbers for CAC are converted from the patient-specific tube voltage value, to the value which would have been obtained at the normal CAC scoring energy of 120 kVp. The potential of this novel reconstruction kernel is assessed with both static and dynamic phantoms. In **chapter 13**, the influence of a novel deep learning-based reconstruction method on detection and quantification of CAC is considered. For this, a comparison is made with filtered back projection (FBP), hybrid IR and model-based IR from the same manufacturer. Finally, in **chapter 14**, an update for the international standard for CAC quantification is presented. The main goal of this new standard is to improve reproducibility of CAC scoring within CT systems as well as between different CT vendors at a reduced radiation dose, by using reduced tube voltage and smaller slice thicknesses.

In part 3, post-processing parameters for CAC scoring are investigated. **Chapter 15** is focused on a vendor neutral Agatston score (vnAS), for which a calculator is developed. In this chapter, the impact of the vnAS on patients' management and outcome is assessed using data from the Multi-Ethnic Study of Atherosclerosis (MESA) study. Lastly, in **chapter 16**, a fully automated quantification method of CAC in an international standardized anthropomorphic phantom is presented.

REFERENCES

1. WHO Fact Sheet - The Top 10 Causes of Death. May 2018. https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

2. Hansson GK. Inflammation, atherosclerosis, and coronary artery disease. The New England journal of medicine. 2005;352(16):1685-1695. doi:10.1056/NEJMra043430

3. Tintut Y, Alfonso Z, Saini T, et al. Multilineage potential of cells from the artery wall. Circulation. 2003;108(20):2505-2510. doi:10.1161/01.CIR.0000096485.64373.C5

4. Greenland P, Alpert JS, Beller GA, et al. 2010 ACCF/AHA guideline for assessment of cardiovascular risk in asymptomatic adults: Executive summary: A report of the American College of cardiology foundation/American Heart association task force on practice guidelines. Circulation. 2010;122(25):2748-2764. doi:10.1161/CIR.0b013e3182051bab

5. Blaha MJ, Mortensen MB, Kianoush S, Tota-Maharaj R, Cainzos-Achirica M. Coronary Artery Calcium Scoring: Is It Time for a Change in Methodology? JACC: Cardiovascular Imaging. 2017;10(8):923-937. doi:10.1016/j.jcmg.2017.05.007

6. Hecht H, Blaha MJ, Berman DS, et al. Clinical indications for coronary artery calcium scoring in asymptomatic patients: Expert consensus statement from the Society of Cardiovascular Computed Tomography. Journal of Cardiovascular Computed Tomography. 2017;11(2):157-168. doi:10.1016/j.jcct.2017.02.010

7. Greenland P, Blaha MJ, Budoff MJ, Erbel R, Watson KE. Coronary Calcium Score and Cardiovascular Risk. J Am Coll Cardiol. 2018;72(4):434-447. doi:10.1016/j.jacc.2018.05.027.

8. Budoff MJ, Achenbach S, Blumenthal RS, et al. Assessment of coronary artery disease by cardiac computed tomography: A scientific statement from the American Heart Association Committee on Cardiovascular Imaging and Intervention, Council on Cardiovascular Radiology and Intervention, and Committee on C. Circulation. 2006;114(16):1761-1791. doi:10.1161/CIRCULATIONAHA.106.178458

9. Alluri K, Joshi PH, Henry TS, Blumenthal RS, Nasir K, Blaha MJ. Scoring of coronary artery calcium scans: History, assumptions, current limitations, and future directions. Atherosclerosis. 2015;239(1):109-117. doi:10.1016/j.atherosclerosis.2014.12.040

10. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte MJ, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. Journal of the American College of Cardiology. 1990;15(4):827-832.

11. Willemink MJ, Abramiuc B, den Harder AM, et al. Coronary calcium scores are systematically underestimated at a large chest size: A multivendor phantom study. Journal of Cardiovascular Computed Tomography. 2015;9(5):415-421. doi:10.1016/j.jcct.2015.03.010

12. Groen JM, Greuter MJ, Schmidt B, Suess C, Vliegenthart R, Oudkerk M. The influence of heart rate, slice thickness, and calcification density on calcium scores using 64-slice multidetector computed tomography: A systematic phantom study. Investigative Radiology. 2007;42(12):848-855. doi:10.1097/RLI.0b013e318154c549

13. Greuter MJW, Groen JM, Nicolai LJ, Dijkstra H, Oudkerk M. A model for quantitative correction of coronary calcium scores on multidetector, dual source, and electron beam computed tomography for influences of linear motion, calcification density, and temporal resolution: A cardiac phantom study. Medical Physics. 2009;36(11):5079-5088. doi:10.1118/1.3213536

14. Husmann L, Leschka S, Desbiolles L, et al. Coronary artery motion and cardiac phases: dependency on heart rate -- implications for CT image reconstruction. Radiology. 2007;245(2):567-576. doi:10.1148/radiol.2451061791

15. Achenbach S, Ropers D, Holle J, Muschiol G, Daniel WG, Moshage W. In-plane coronary arterial motion velocity: measurement with electron-beam CT. Radiology. 2000;216(2):457-463. doi:10.1148/radiology.216.2.r00au19457

16. Symons R, Sandfort V, Mallek M, Ulzheimer S, Pourmorteza A. Coronary artery calcium scoring with photon-counting CT: first in vivo human experience. The international journal of cardiovascular imaging. 2019;35(4):733-739. doi:10.1007/s10554-018-1499-6

17. Vonder M, Pelgrim GJ, Huijsse SEM, et al. Feasibility of spectral shaping for detection and quantification of coronary calcifications in ultra-low dose CT. European Radiology. 2017;27(5):2047-2054. doi:10.1007/s00330-016-4507-z

18. Marwan M, Mettin C, Pflederer T, et al. Very low-dose coronary artery calcium scanning with high-pitch spiral acquisition mode: Comparison between 120-kV and 100-kV tube voltage protocols. Journal of Cardiovascular Computed Tomography. 2013;7(1):32-38. doi:10.1016/j.jcct.2012.11.004
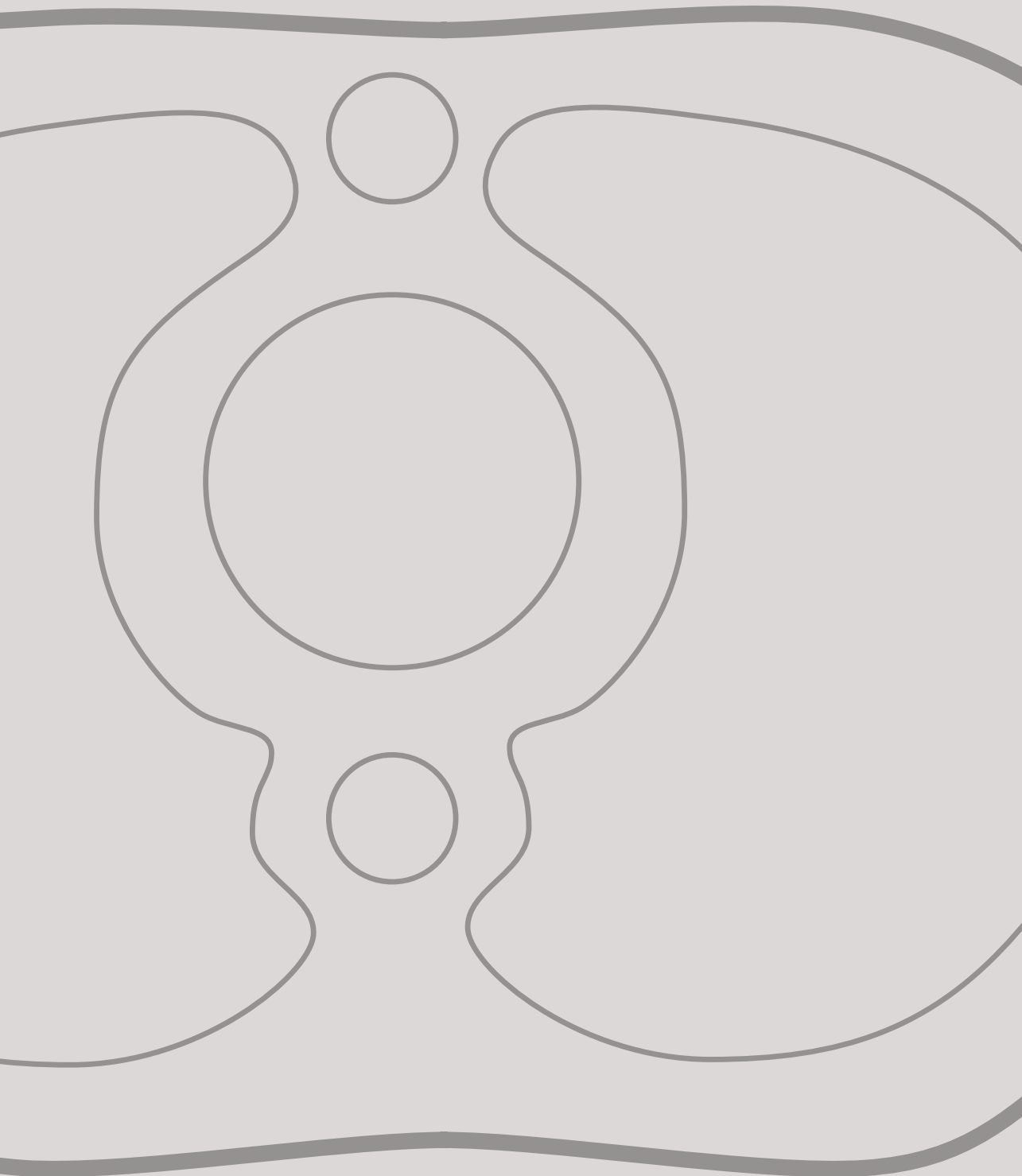
PART I

Patient specific factors
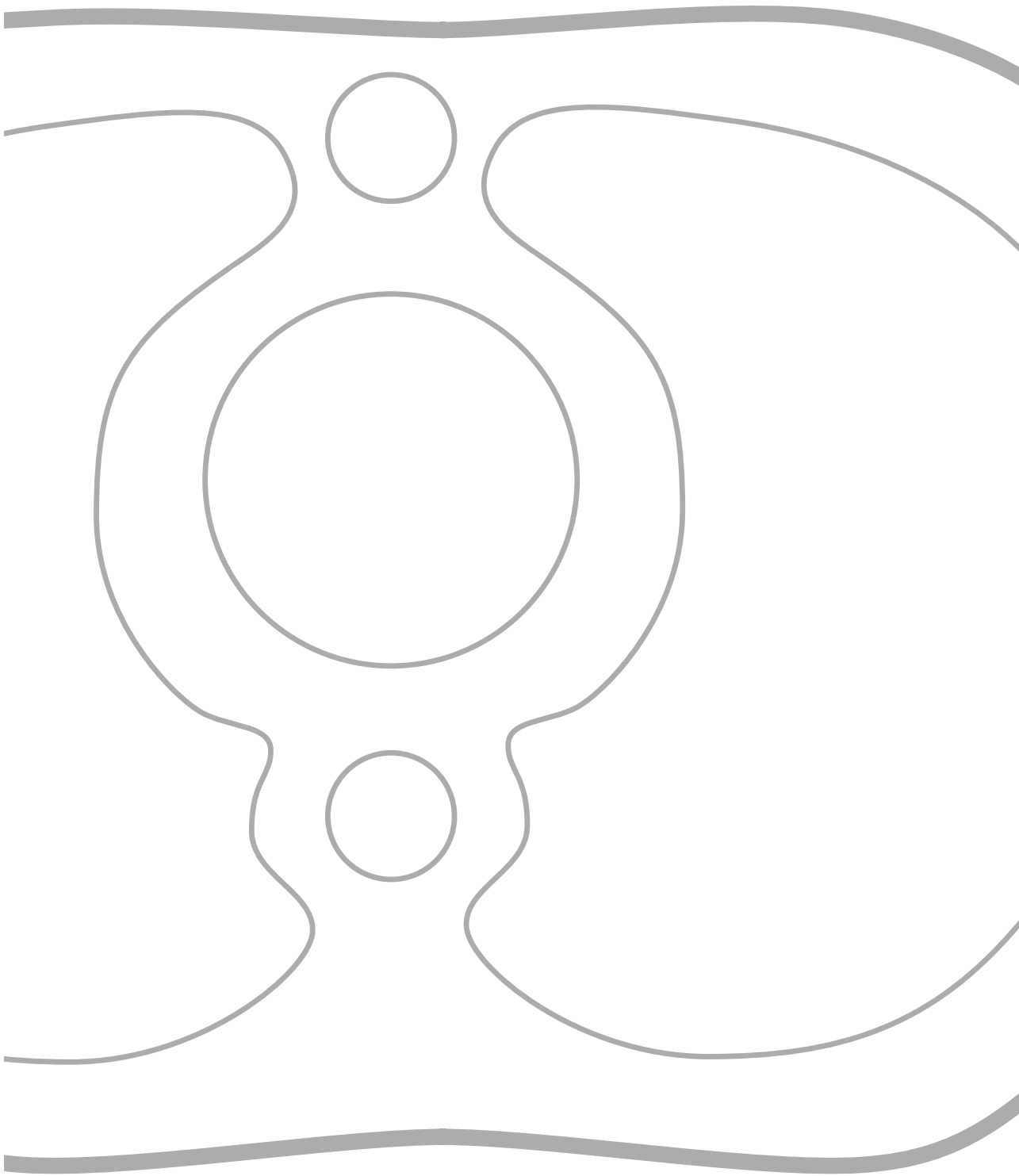
CHAPTER 2

Influence of heart rate
on coronary calcium scores:
a multi-manufacturer phantom study

Niels R. van der Werf, MSc

Martin J. Willemink, MD PhD

Tineke P. Willems, MD PhD

Rozemarijn Vliegenthart, MD PhD

Marcel J.W. Greuter, PhD

Tim Leiner, MD PhD

ABSTRACT

2

Purpose
To evaluate the influence of heart rate on coronary calcium scores (CCS) using a dynamic phantom on four high-end computed tomography (CT) systems from different manufacturers.

Methods
Artificial coronary arteries were moved in an anthropomorphic chest phantom at linear velocities, corresponding to <60, 60-75 and >75 beats per minute (bpm). Data was acquired with routinely used clinical protocols for CCS on four high-end CT systems (CT1-CT4). CCS, quantified as Agatston and mass scores were compared to reference scores at <60 bpm. Influence of heart rate was assessed for each system with the cardiac motion susceptibility (CMS) index.

Results
At increased heart rates (>75 bpm), Agatston scores of the low mass calcification were similar to the reference score, while Agatston scores of the medium and high mass calcification increased significantly up to 50% for all CT systems. Threefold CMS increases at >75 bpm in comparison with <60 bpm were shown. For medium and high mass calcifications, significant differences in CMS between CT systems were found.

Conclusion
Heart rate substantially influences CCS for high-end CT systems of four major manufacturers, but CT systems differ in motion susceptibility. Follow-up CCS CT scans should be acquired on the same CT system and protocol, and preferably with comparable heart rates.

INTRODUCTION

Despite advances in treatment, atherosclerotic cardiovascular disease remains the main cause of mortality worldwide.[1] For the determination of the prognosis of future cardiovascular disease, risk categories based on the amount of coronary calcium are increasingly used.[2] Coronary calcium can be quantified with computed tomography (CT) as coronary calcium scores (CCS), including Agatston scores, mass scores and volume scores.[3] Correct and reproducible CCS are essential considering their use in the guidelines.[1,4] If risk-based treatment decision is uncertain after quantitative risk assessment, the 2013 guidelines recommend repeated CCS measurements. Also, CCS progression in repeated CCS measurements has recently been associated with heart failure.[5]

Several studies have demonstrated the dependence of CCS on motion.[6–9] Motion artifacts can result in incorrect Agatston scores, which may thereby lead to incorrect risk classification. Depending on heart rate and anatomical location, coronary arteries move at velocities of 10 to 30 mm/s during the acquisition phase, whereas coronary arteries are often erroneously assumed to be stationary during CT acquisition.[10–12]

Previously published studies on the influence of heart rate on CCS have only focused on CT systems from a single manufacturer.[9,13,14] Recently, substantial differences in CCS between new generation CT systems were demonstrated. Willemink et al found differences in Agatston scores between CT systems of up to 43.9% for static calcifications.[15] However, the effect of heart rate on CCS for different state-of-the-art CT systems remains unknown. The aim of the current study was therefore to determine the influence of heart rate on CCS for the high-end CT systems from four major manufacturers at routinely used clinical protocols with a dynamic phantom.

MATERIAL AND METHODS

Phantom
A computer-controlled lever (QRM-Sim2D, QRM, Möhrendorf, Germany) moved an artificial coronary artery in a water container in the center of an anthropomorphic chest phantom (QRM-Chest, QRM, Möhrendorf, Germany) (*Figure 1*). The chest phantom consisted of artificial lungs, a spine insert and a shell of tissue equivalent material. An extension ring (QRM-Extension Ring, QRM, Möhrendorf, Germany), made of fat equivalent material, was used to increase the size of the phantom to
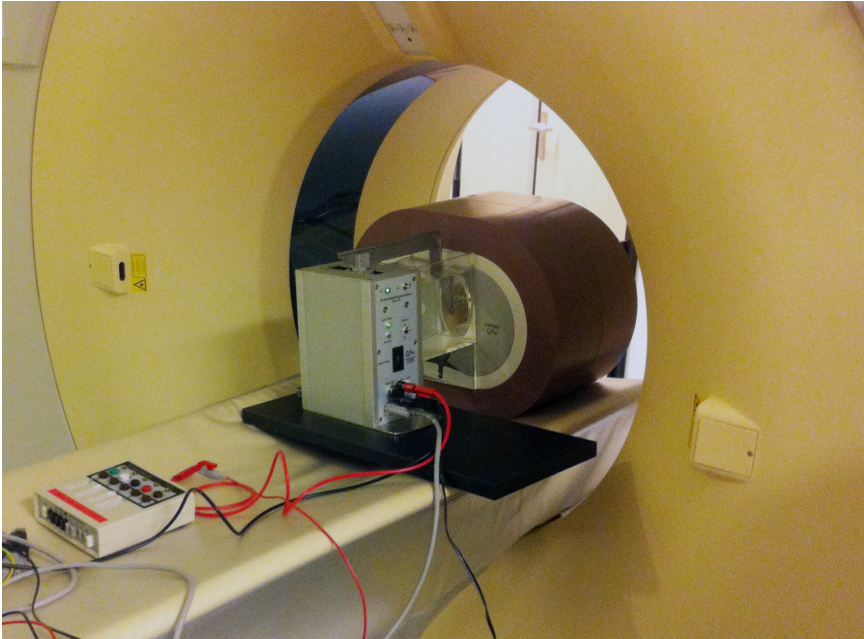
2

Figure 1 Anthropomorphic chest phantom with extension ring and motion controller (QRM, Moehrendorf, Germany). An artificial coronary artery moved in the horizontal plane inside the water tank in the center of the chest phantom

400 x 300 mm, in order to mimic an average sized patient. The artificial artery was moved at constant linear velocities of 0, 10, 20 and 30 mm/s in the horizontal plane, perpendicular to the scan direction. Artificial calcified coronary arteries were used, which contained cylindrical calcifications of calcium hydroxyapatite (HA) (*Figure 2*). The dimensions of the calcifications were 5.0±0.1 mm in diameter and 10.0±0.1 mm in length. Three calcification densities were used: 196±3, 408±2 and 800±2 mg HA/cm^3. These densities were categorized as low (38.5±1.7 mg), medium (80.1±3.3 mg) and high (157.1±6.5 mg), corresponding to mild, moderate and severe calcified coronary plaque burden, respectively.

Image acquisition and evaluation
In order to assess the influence of heart rate on CCS in a clinical setting, routinely used clinical CT CCS protocols were used (*Table 1*). The protocols were equal to the manufacturer recommended protocol if available or were adapted from the factory settings based on recommendations by the specific manufacturer consultants. Four high-end CT systems from the main four CT manufacturers were used
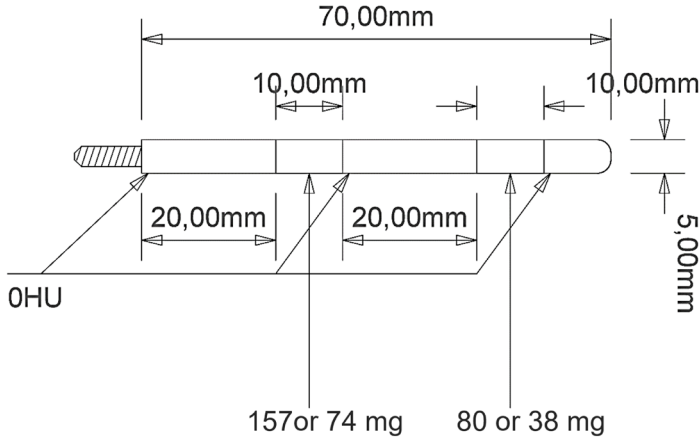
Figure 2 The cylindrical artificial coronary artery contained two calcified inserts with a diameter of 5.0±0.1 mm and a length of 10.0±0.1 mm as indicated

Table 1 Acquisition and reconstruction parameters used on CT system CT1 to CT4.

CT system	CT1	CT2	CT3	CT4
Tube voltage [kV]	120	120	120	120
Tube current per rotation [mA]	175	50	80	80
Collimation [mm]	64 x 0.625	128 x 0.625	128 x 0.6	320 x 0.5
Rotation time [ms]	350	270	280	350
Temporal resolution* [ms]	175	135	75	175
Slice thickness [mm]	2.5	3.0	3.0	3.0
Increment [mm]	2.5	3.0	3.0	3.0
Kernel	Standard	XCA	B35f	FC12
Calcium scoring software	Smartscore 4.0	Heartbeat-CS	Syngo.via	Vitrea FX 6.5.0

 * As defined in the isocenter

(CT1-CT4): Discovery CT 750 HD (GE Healthcare, Waukesha, Wisconsin, USA), Brilliance iCT (Philips Healthcare, Best, The Netherlands), Somatom Definition Flash (Siemens Healthcare, Forchheim, Germany) and Aquilion One (Toshiba Medical Systems, Otawara, Japan), respectively.

For each velocity the phantom was scanned five times, with a small random translation (2 mm) and rotation (2 degrees) between the scans by repositioning of the phantom. An electrocardiographic trigger output of the motion controller

was used for electrocardiographic triggering of the CT systems to ensure that data was acquired only during linear motion of the phantom (*Figure 3*). Data were reconstructed with filtered back projection (FBP) on each CT system using the kernels listed in *Table 1*. CCS was quantified as Agatston score and mass score using the routinely used manufacturers' software with a default threshold for calcium scoring of 130 Hounsfield units (HU). The mass score calibration factor was calculated as described by McCollough et al.[16] One observer performed the measurements with a semi-automatic method by selecting the calcification. For each individual calcification, median and range from the five measurements were calculated for both the Agatston score and mass score.
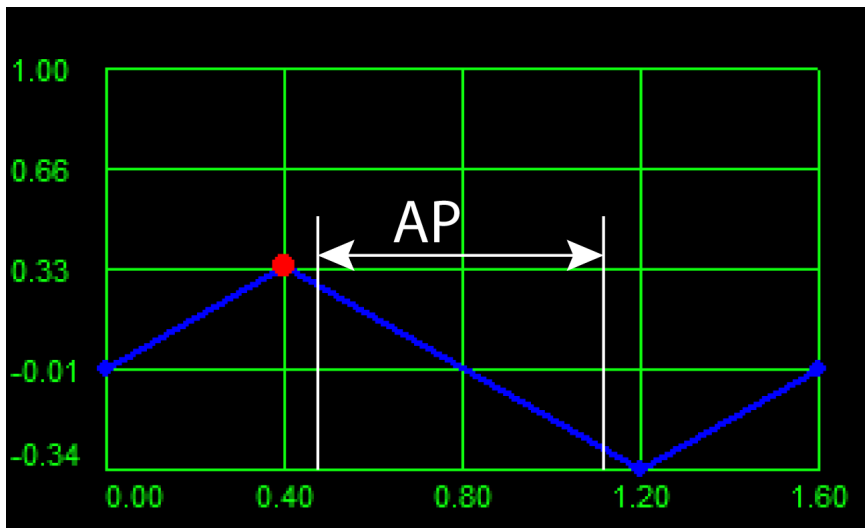
Figure 3 Example profile for 30 mm/s (>75 bpm) movement of the Sim2D motion controller. The red dot resembles the ECG triggering point, after which acquisition was started. For all CT systems, acquisition was in the acquisition phase (AP), as indicated in the figure. Therefore, all acquisitions were during constant motion of the motion controller

In order to interpret the results in terms of heart rate dependent CCS, the linear velocities of the artificial coronary arteries were converted to corresponding heart rates. Averaged over all coronary arteries at a cardiac cycle phase of 70% of the RR-interval, the average velocity of the coronary arteries during the acquisition phase is approximately 10, 20 and 30 mm/s for heart rates of <60, 60-75 and >75 beats per minute (bpm), respectively.[11]

Statistical analysis

Wilcoxon signed rank tests were used to compare Agatston and mass scores to the reference scores at <60 bpm, and to compare mass scores to the physical mass. A p-value smaller than 0.05 was used to indicate significant differences. Motion susceptibility was assessed with the cardiac motion susceptibility (CMS) index, as described by Groen et al.[17] This index gives a measure of the median deviation of the calcium scores over all heart rates from the CCS at rest. The CMS index was calculated with the following equation:

$$CMS = \frac{1}{N-1} \sqrt{\sum_{i=1}^{N} (x_0 - x_i)^2 \frac{1}{x_0}}$$

In this equation, x_0 is the CCS at 0 bpm. The total number of heart rates is given by N, and the CCS at heart rate i by x_i. A smaller CMS value signifies a lower motion susceptibility of CCS to cardiac motion. Kruskal Wallis tests were used to assess differences between calcium masses for each CT system and between CT systems, with a p-value smaller than 0.05 indicating significant differences. All statistical analyses were performed with SPSS for Windows, version 22.0 (IBM corp., Armonk, NY, USA).

RESULTS

CCS for the reference heart rate of <60 bpm are listed in *Tables 2* and *3*. At the reference heart rate, mass scores underestimated physical mass significantly for all calcification masses and all CT systems, except for the high (p=0.500) and medium (p=0.221) mass calcification for CT 2 and CT4 respectively. Example CT images for the high mass calcification at >75 bpm for all CT systems are shown in *Figure 4*.

Influence of heart rate on Agatston score

The influence of increasing heart rates on Agatston scores for low, medium and high mass calcifications is shown in *Table 2* and *Figure 5A*. For the low mass calcification and all CT systems, Agatston scores at increasing heart rates were not significantly different from the reference Agatston score at <60 bpm, regardless of the CT system. For the medium mass calcification at >75 bpm, significant increases in Agatston scores were shown for all CT systems. Significant increases up to 50% were also found for the high mass calcification. The influence of heart rate on the Agatston scores was different depending on the particular CT system.

2

For example, at increased heart rate, CT3 showed an increase in Agatston score of 22% for the high mass calcification, while differences of 45% to 50% were found for the other CT systems.
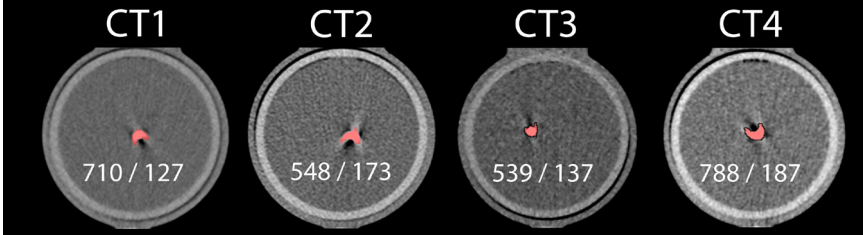
Figure 4 Example images for all CT systems of the high mass (157 mg) calcification at >75 bpm. Agatston / mass scores are given for each CT system

Table 2 Reference Agatston scores at <60 bpm and percentage deviation from this reference as a function of heart rate for low mass (38 mg), medium mass (80 mg) and high mass (157 mg) calcifications. Results are indicated as median and total range values between brackets. Significant deviations from the reference Agatston scores are indicated by asterisks

CT	Calcification mass	Reference	60 – 75 bpm		>75 bpm	
		Median (range)	Median % (range)	p-value	Median % (range)	p-value
CT1	Low	99 (95-136)	-7 (-31; 2)	0.144	-21 (-31; 6)	0.078
	Medium	351 (344-375)	4 (-4; 30)	0.176	32 (11; 33)	0.043*
	High	515 (450-553)	-3 (-11; 41)	0.686	50 (41; 58)	0.043*
CT2	Low	105 (79-120)	-6 (-21; 30)	0.893	-9 (-38; 11)	0.279
	Medium	315 (267-347)	11 (1; 39)	0.043*	33 (7; 67)	0.043*
	High	421 (387-459)	14 (2; 48)	0.043*	49 (19; 70)	0.043*
CT3	Low	103 (89-117)	-10 (-20; 28)	0.686	6 (-18; 23)	0.893
	Medium	304 (295-340)	4 (-2; 9)	0.225	9 (3; 19)	0.043*
	High	424 (406-443)	4 (0; 12)	0.043*	22 (9; 28)	0.042*
CT4	Low	108 (105-132)	-5 (-17; 15)	0.345	-10 (-21; 1)	0.080
	Medium	367 (320-377)	33 (23; 34)	0.042*	26 (14; 45)	0.043*
	High	505 (472-544)	38 (36; 53)	0.043*	45 (44; 54)	0.043*

Influence of heart rate on mass score

Mass scores as a function of heart rate for the low, medium and high mass calcification are shown in *Table 3*. At 60-75 bpm, significantly different mass scores in comparison with the reference at <60 bpm were only shown for CT1 with the low mass calcification. At >75 bpm, CT1 showed significantly different mass scores for the low mass and high mass calcification. The low mass and high mass calcification also showed significantly different mass scores for CT2. For CT3, all mass scores at >75 bpm were similar to the reference scores. CT4 showed significantly different mass scores for the low mass and high mass calcification.

Table 3 Reference Mass scores at <60 bpm and percentage deviation from this reference as a function of heart rate for low mass (38 mg), medium mass (80 mg) and high mass (157 mg) calcifications. Results are indicated as median and total range values between brackets. Significantly different mass scores, in comparison with the physical mass, are indicated by an †. Significant deviations from the reference mass scores are indicated by asterisks

CT	Calcification mass	Reference		60 – 75 bpm		>75 bpm	
		Median (range)	p-value	Median % (range)	p-value	Median % (range)	p-value
CT1	Low	31 (28-35)	0.078	-26 (-33; -18)	0.043*	-27 (-33; -4)	0.041*
	Medium	76 (74-78)	0.043†	-6 (-21; 3)	0.080	-9 (-22; 1)	0.078
	High	145 (127-153)	0.043†	-8 (-25; -6)	0.042*	-19 (-21; -7)	0.043*
CT2	Low	25 (23-26)	0.279	0 (-12; 13)	0.785	-13 (-22; -4)	0.043*
	Medium	75 (68-79)	0.043†	1 (0; 10)	0.102	8 (0; 19)	0.068
	High	158 (152-167)	0.043†	3 (0; 15)	0.068	15 (7; 21)	0.043*
CT3	Low	20 (18-23)	0.893	-11 (-16; 5)	0.221	-11 (-13; 0)	0.066
	Medium	61 (59-64)	0.043†	0 (-2; 5)	0.461	-2 (-5; 8)	0.892
	High	128 (127-132)	0.042†	1 (0; 6)	0.066	4 (-2; 13)	0.078
CT4	Low	28 (26-30)	0.080	-7 (-18; 12)	0.276	-18 (-20; -12)	0.042*
	Medium	81 (77-84)	0.043†	4 (2; 12)	0.042	1 (-5; 5)	0.683
	High	170 (167-175)	0.043†	12 (6; 12)	0.039*	9 (4; 11)	0.043*

In comparison with the physical mass, for all CT systems, mass scores for the low mass calcification were underestimated significantly (p<0.05) by 19% to 48% averaged over all heart rates (*Figure 5B*). For the medium mass calcification at 60-75 bpm, CT4 showed an accurate approximation of the physical mass (deviation 7%, p=0.080), whereas the other CT systems showed significant underestimations (p<0.05) up to 23%. At >75 bpm, both CT2 and CT4 showed an accurate approximation of the physical mass (deviations 1% (p=0.686)

2
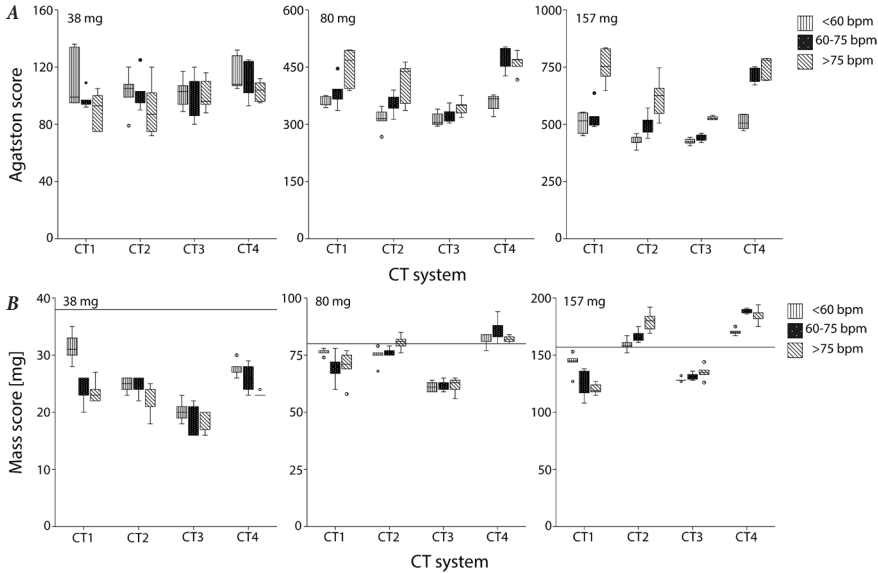
Figure 5 A) Agatston scores of the four artificial coronary arteries as found on CT-systems CT1 to CT4 for calcification masses 38 (left), 80 (middle) and 157 mg (right) **B)** Mass scores of the four artificial coronary arteries as found on CT-systems CT1 to CT4 for calcification masses 38 (left), 80 (middle) and 157 mg (right). Physical mass of the calcification is represented with solid lines

and 2% (p=0.080), respectively), whereas CT1 and CT3 underestimated the physical mass by 11% (p=0.043) and 21% (p=0.043). For the high mass calcification, CT1 and CT3 underestimated the physical mass significantly (p<0.05) up to 24% for 60-75 and >75 bpm. Significant overestimations were found for CT2 and CT4 at increased heart rates. At 60-75 bpm, mass scores deviated by 5% (p=0.043) and 20% (p=0.043) for CT2 and CT4 respectively, while differences of 15% (p=0.043) and 19% (p=0.042) were found at >75bpm.

CMS analysis

Susceptibility of the Agatston and mass scores to motion, as expressed by CMS values, were smaller for mass scores, which indicates a smaller motion susceptibility of the mass score in comparison with the Agatston score (*Figure 6*). Averaged over all calcification masses, the CMS value of CT3 was the smallest. Comparison of the CMS values of the three calcification masses per CT system showed significant differences between these values for CT1 (p=0.009) and CT4 (p=0.002). The motion susceptibility of the Agatston score for CT1 and CT4 therefore depends on the calcification density. For CT2 (p=0.114) and CT3 (p=0.054) CMS values of the
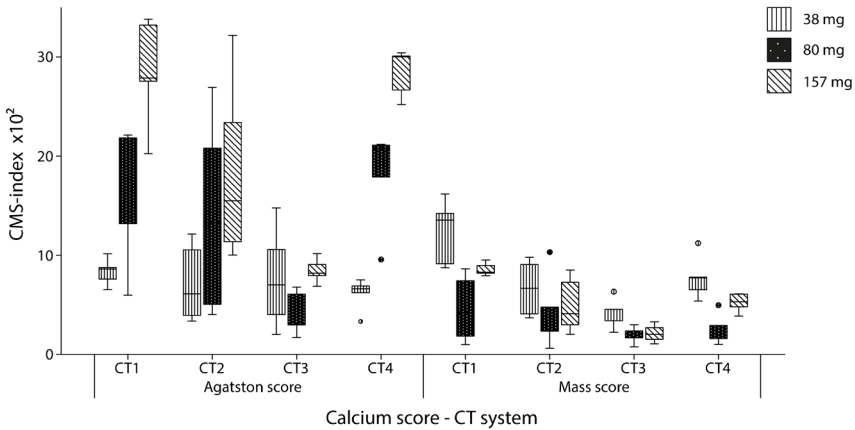
Figure 6 CMS index for Agatston and mass scores for all CT systems and for the low (38 mg), medium (80 mg) and high (157 mg) mass calcification. Lower values indicate a smaller susceptibility to differences in motion, and therefore a smaller susceptibility to differences in heart rate

three masses were not significantly different. Statistical analysis of the mass score CMS values showed similar results. Significant differences between CMS values of the different calcification masses were found for CT1 (p=0.008) and CT4 (p=0.005, whereas CT2 (p=0.432) and CT3 (p=0.050) did not show significant differences.

Differences in motion susceptibility between CT systems were significant. For the low mass calcification, susceptibility to motion as expressed in CMS values of the Agatston scores were comparable (p=0.564) between CT systems, whereas CMS values of the mass scores showed significant differences (p=0.009). The medium mass calcification showed significantly different (p=0.043) Agatston score CMS values, while mass score CMS values were comparable (p=0.687). For the high mass calcification, significant differences were shown for both the Agatston (p=0.006) and mass score (p=0.004) CMS values.

DISCUSSION

This dynamic phantom study showed that for high-end CT systems of four major CT manufacturers coronary calcium scores were substantially, but not equally, influenced (up to 50%) by heart rate. Motion susceptibility depends on both calcification mass and CT system. At increased heart rates, the stability of the mass score is superior to Agatston score stability.

2

The most important findings in this study were as follows. First, we found that no modern CT system is capable of completely mitigating the influence of heart rate on Agatston score or calcification mass. Second, CMS index analysis showed significant differences for both Agatston and mass scores between four high-end CT systems. Also, different heart rates were found to differentially affect calcification masses for two CT systems. Thus, for these CT systems the influence of heart rate on CCS depends on calcification density.

For follow-up imaging a valid and precise scoring method is essential, therefore subsequent scans should be made with the same CT system and heart rates should be comparable. For low mass calcifications, voxels above the 130HU threshold at low heart rate might fall below this threshold as a result of motion blurring. As a result of this effect, CCS will decrease for these calcifications. For medium and high mass calcifications, the number of voxels above the 130HU threshold will increase as a result of motion blurring, thereby increasing CCS. The registered calcification size is smaller for CT systems with a higher temporal resolution, which is demonstrated in *Figure 4* for CT3 in comparison with the other CT systems. However, many other parameters, including acquisition and reconstruction parameters, also influence the registered calcification size.[18] In addition, technical advances, including dual-energy CT in combination with mono-energetic reconstructions, may influence the registered calcification quantification. With these mono-energetic reconstructions, it has been shown that blooming artefacts can be reduced.[19] Although especially for high density calcifications improved size registration is expected, the influence of these reconstructions on motion artefacts is unknown at this time. Third, the stability of the mass score is clearly superior to the stability of the Agatston score. This can be explained from the different way of calculating the mass score in comparison with the Agatston score. While the Agatston score considers the highest HU value in the lesion, the mass score can be thought of as a weighted average, which is therefore more robust to the effects of motion blurring. We found that the use of modern, state-of-the-art CT systems still results in substantial deviation from the true physical mass.

Our results are in line with a study conducted by Groen et al, who found decreased CCS for low mass calcifications and increased CCS for high mass calcifications at increased heart rates.[7] However, this study focused on only one CT system, while we used four high-end CT systems in the current study. Our findings are in agreement with the studies of Greuter et al and Groen et al, who also showed the superior stability of mass scores in comparison to Agatston scores.[6,7] These studies, however, were only performed on CT systems from one manufacturer. Tigges et

al used explanted coronary arteries mounted on a cardiac motion phantom to examine the influence of heart rate on CCS.[8] The explanted coronary arteries were subdivided into categories based on their respective Agatston score. The study showed a decrease in CCS with increasing heart rates, independent of calcification mass. For low mass calcifications these results correspond well to our results, whereas the high mass calcification results differ from our results. This difference in results probably originates from a difference in calcification density, because a high Agatston score can be the result of either a large calcification or a high density. The density of the calcification was not specified in the study of Tigges et al.

The current study has limitations. First, the artificial coronary arteries used in this study were translated at constant linear motion, whereas complex 3D motions are observed in-vivo. This 3D motion is patient and artery specific. Nevertheless, the scan times were relatively short as a result of fast rotation times, whereby the constant linear motion of our phantom was deemed sufficient as a model of the complex in vivo motion of coronary arteries. Second, linear velocity was converted to heart rates based on only one available study.[11] This limitation was accepted because of large variation in heart rates within the chosen heart rate groups (<60, 60-75 and >75 bpm). Third, slice thickness and increment were not the same for all CT systems. Previous studies demonstrated increased CCS for decreased slice thicknesses.[20,21] However, the current study focusses on CCS from clinically used CCS protocols to evaluate the influence of heart rate within each CT system. Also, the deviating slice thickness and increment for CT1 in comparison with the other systems is in line with protocol recommendations as described previously by McCollough et al.[16] With the clinically used CCS protocols, CCS were acquired with routine CT manufacturers' software. However, it was demonstrated by Weininger et al that different types of CCS scoring software resulted in similar CCS.[22]

In conclusion, the current dynamic phantom study showed that heart rate substantially influences CCS for the high-end CT systems of the four major CT manufacturers. Therefore, this study suggests that heart rate should be taken into account when interpreting the clinical impact of patients' calcium scores. Moreover, it is essential to acquire CCS scans on similar CT systems between follow-up studies, preferably at comparable heart rates. Also, mass scores were found to be more stable then Agatston scores.

2

REFERENCES

1. Greenland P, Alpert JS, Beller GA, et al. 2010 ACCF/AHA guideline for assessment of cardiovascular risk in asymptomatic adults: Executive summary: A report of the American College of cardiology foundation/American Heart association task force on practice guidelines. Circulation. 2010;122(25):2748-2764. doi:10.1161/CIR.0b013e3182051bab

2. Wilson PWF, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation 1998. Circulation. 1998;97:1837-1847.

3. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte MJ, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. J Am Coll Cardiol. 1990;15(4):827-832.

4. Goff DC, Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: A report of the American college of cardiology/American heart association task force on practice guidelines. Circulation. 2014;129(25 SUPPL. 1):49-76. doi:10.1161/01.cir.0000437741.48606.98

5. Bakhshi H, Ambale-Venkatesh B, Yang X, et al. Progression of Coronary Artery Calcium and Incident Heart Failure: The Multi-Ethnic Study of Atherosclerosis. J Am Heart Assoc. 2017;6(4). doi:10.1161/JAHA.116.005253

6. Greuter MJW, Groen JM, Nicolai LJ, Dijkstra H, Oudkerk M. A model for quantitative correction of coronary calcium scores on multidetector, dual source, and electron beam computed tomography for influences of linear motion, calcification density, and temporal resolution: A cardiac phantom study. Med Phys. 2009;36(11):5079-5088. doi:10.1118/1.3213536

7. Groen JM, Greuter MJ, Schmidt B, Suess C, Vliegenthart R, Oudkerk M. The influence of heart rate, slice thickness, and calcification density on calcium scores using 64-slice multidetector computed tomography: A systematic phantom study. Invest Radiol. 2007;42(12):848-855. doi:10.1097/RLI.0b013e318154c549

8. Tigges S, Arepalli CD, Tridandapani S, et al. A phantom study of the effect of heart rate, coronary artery displacement and vessel trajectory on coronary artery calcium score: Potential for risk misclassification. J Cardiovasc Comput Tomogr. 2012;6(4):260-267. doi:10.1016/j.jcct.2012.01.005

9. Funabashi N, Irie R, Namihira Y, et al. Influence of tube voltage and heart rate on the Agatston calcium score using an in vitro, novel ECG-gated dual energy reconstruction 320 slice CT technique. Int J Cardiol. 2015;180:218-220. doi:10.1016/j.ijcard.2014.11.164

10. Achenbach S, Ropers D, Holle J, Muschiol G, Daniel WG, Moshage W. In-plane coronary arterial motion velocity: measurement with electron-beam CT. Radiology. 2000;216(2):457-463. doi:10.1148/radiology.216.2.r00au19457

11. Husmann L, Leschka S, Desbiolles L, et al. Coronary artery motion and cardiac phases: dependency on heart rate -- implications for CT image reconstruction. Radiology. 2007;245(2):567-576. doi:10.1148/radiol.2451061791

12. Rutten A, Krul SPJ, Meijs MFL, De Vos AM, Cramer MJM, Prokop M. Variability of coronary calcium scores throughout the cardiac cycle: Implications for the appropriate use of electrocardiogram-dose modulation with retrospectively gated computed tomography. Invest Radiol. 2008;43(3):187-194. doi:10.1097/RLI.0b013e31815cdd56

13. Hong C, Bae KT, Pilgram TK, Zhu F. Coronary artery calcium quantification at multi-detector row CT: influence of heart rate and measurement methods on interacquisition variability initial experience. Radiology. 2003;228(1):95-100. doi:10.1148/radiol.2281020685

14. Hong C, Bae KT, Pilgram TK. Coronary Artery Calcium: Accuracy and Reproducibility of Measurements with Multi–Detector Row CT—Assessment of Effects of Different Thresholds and Quantification Methods. Radiology. 2003;227(3):795-801. doi:10.1148/radiol.2273020369

15. Willemink MJ, Vliegenthart R, Takx RAP, et al. Coronary Artery Calcification Scoring with State-of-the-Art CT Scanners from Different Vendors Has Substantial Effect on Risk Classification. Radiology. 2014;273(3):695-702. doi:10.1148/radiol.14140066

16. McCollough CH, Ulzheimer S, Halliburton SS, Shanneik K, White RD, Kalender WA. Coronary Artery Calcium: A Multi-institutional, Multimanufacturer International Standard for Quantification at Cardiac CT. Radiology. 2007;243(2):527-538. doi:10.1148/radiol.2432050808

17. Groen JM, Greuter MJW, Vliegenthart R, et al. Calcium scoring using 64-slice MDCT, dual source CT and EBT: A comparative phantom study. Int J Cardiovasc Imaging. 2008;24(5):547-556. doi:10.1007/s10554-007-9282-0

18. van Ooijen PMA, Vliegenthart R, Witteman JC, Oudkerk M. Influence of scoring parameter settings on Agatston and volume scores for coronary calcification. Eur Radiol. 2005;15(1):102-110. doi:10.1007/s00330-004-2479-x

19. Hickethier T, Baeßler B, Kroeger JR, et al. Monoenergetic reconstructions for imaging of coronary artery stents using spectral detector CT: In-vitro experience and comparison to conventional images. J Cardiovasc Comput Tomogr. 2017;11(1):33-39. doi:10.1016/j.jcct.2016.12.005

20. Muhlenbruch G, Thomas C, Wildberger JE, et al. Effect of varying slice thickness on coronary calcium scoring with multislice computed tomography in vitro and in vivo. Invest Radiol. 2005;40(11):695-699.

21. Van Der Bijl N, De Bruin PW, Geleijns J, et al. Assessment of coronary artery calcium by using volumetric 320-row multi-detector computed tomography: Comparison of 0.5 mm with 3.0 mm slice reconstructions. Int J Cardiovasc Imaging. 2010;26(4):473-482. doi:10.1007/s10554-010-9581-8

22. Weininger M, Ritz KS, Schoepf UJ, et al. Interplatform Reproducibility of CT Coronary Calcium Scoring Software. Radiology. 2012;265(1):70-77. doi:10.1148/radiol.12112532
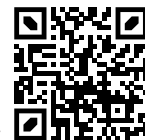
Online material is available via:
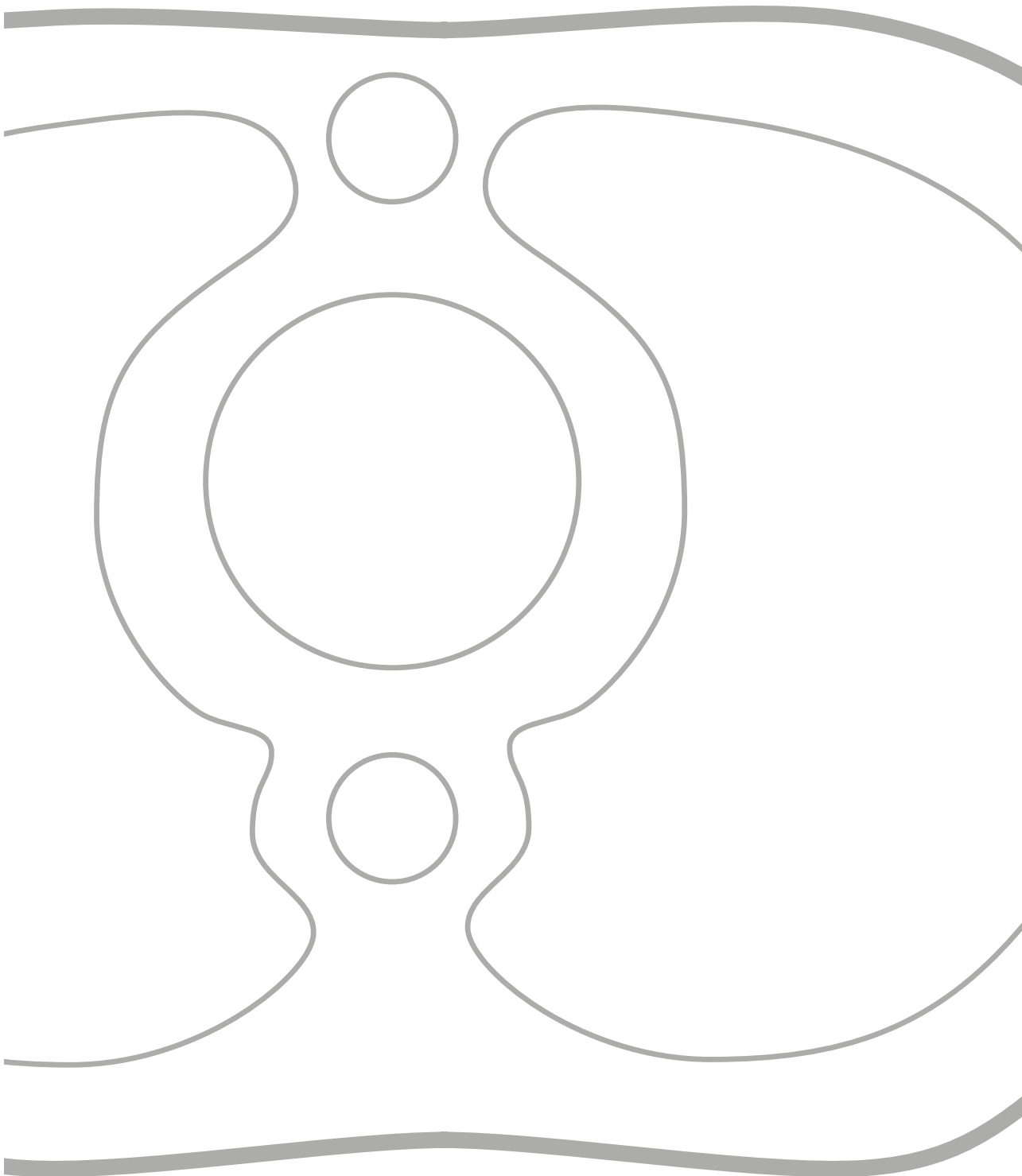
CHAPTER 3

Motion-corrected coronary calcium scores by a convolutional neural network: a robotic simulating study

Yaping Zhang, MD

Niels R. van der Werf, MSc

Beibei Jiang, MD

Robbert W. van Hamersvelt, MD PhD

Marcel J.W. Greuter, PhD

Xueqian Xie, PhD

ABSTRACT

Objective

To classify motion-induced blurred images of calcified coronary plaques so as to correct coronary calcium scores on nontriggered chest CT, using a deep convolutional neural network (CNN) trained by images of motion artefacts.

Methods

Three artificial coronary arteries containing nine calcified plaques of different densities (high, medium, and low) and sizes (large, medium, and small) were attached to a moving robotic arm. The artificial arteries moving at 0–90 mm/s were scanned to generate nine categories (each from one calcified plaque) of images with motion artefacts. An inception v3 CNN was fine-tuned and validated. Agatston scores of the predicted classification by CNN were considered as corrected scores. Variation of Agatston scores on moving plaque and by CNN correction were calculated using the scores at rest as reference.

Results

The overall accuracy of CNN classification was 79.2±6.1% for nine categories. The accuracy was 88.3±4.9%, 75.9±6.4% and 73.5±5.0% for the high-, medium- and low-density plaques, respectively. Compared to the Agatston score at rest, the overall median score variation was 37.8% (1st, 3rd quartile: 10.5%, 68.8%) in moving plaques. CNN correction largely decreased the variation to 3.7% (1.9%, 9.1%) ($p<0.001$, Mann-Whitney U test), and improved the sensitivity (percentage of non-zero scores among all the scores) from 65% to 85% for detection of coronary calcifications.

Conclusion

In this experimental study, CNN showed the ability to classify motion-induced blurred images and correct calcium scores derived from nontriggered chest CT. CNN correction largely reduces the overall Agatston score variation and increases the sensitivity to detect calcifications.

3

INTRODUCTION

Coronary artery calcium is well known as a specific marker for atherosclerosis.[1] The burden of coronary calcium is traditionally expressed as an Agatston score derived from computed tomography (CT) and is generally regarded as an independent risk factor for coronary artery disease.[2,3] Therefore Agatston scores are implied to evaluate cardiovascular risk in asymptomatic populations at intermediate risk.[4] However, because the heart rhythmically moves, the CT images are blurred by motion artefacts which hamper an accurate determination of the amount of coronary calcium.[5] Prospective electrocardiographic (ECG)-triggering or retrospective ECG-gating techniques are generally used to minimize motion of the coronary arteries during CT acquisition.[6] Although minimized, coronary arteries have been shown to move at typical velocities of 10-30 mm/s during acquisition with peaks up to 90 mm/s.[7,8]

Compared with ECG-triggering or -gating CT, nontriggered chest CT is extensively used worldwide. With the implementation of lung cancer screening, large datasets of patients are acquired with nontriggered chest CT.[9] As heavy smoking is, in addition to lung disease, also associated with cardiovascular disease, quantification of coronary artery calcification plays an active role in the identification of individuals at risk for adverse cardiovascular events.[10] Because this information is obtained from already acquired CT scans, radiation dose and costs might be effectively deduced. Recent guidelines provide a Class I indication for the evaluation and report of at least qualitative coronary calcium scores on nontriggered chest CT examinations.[11] However, in a nontriggered chest scan, it is difficult to avoid motion artefacts of calcified plaques, which might result in considerable variation in calcium scores. And although proven to be efficient to identify high risk patients with high calcium scores in nontriggered CT, motion artifacts play a more important role in patients with lower calcium scores as changes in calcium score more easily lead to patient risk reclassification to different calcium score categories.[9,12] Efforts have been made to correct calcium scores of calcified plaques in motion. Ma et al. suggested that motion artefact metrics may be useful for developing and evaluating methods to reduce motion in cardiac CT images.[13]

Convolutional neural networks (CNN) allow for computational models composed of multiple processing layers which can learn image features and classify images.[14] This might drastically improve the performance in computer vision and many other domains. Recently, CNN algorithms have been applied to recognize medical images by extracting internal features of lesions, such as classification of giant cell

tumor of bone, skin cancer and lung nodules.[15–17] Šprem et al. proposed a CNN-based method to enable identification of scans affected by severe cardiac motion.[18] Before correcting calcium scores in a clinical scenario, we intend to perform an experimental study to evaluate the CNN's classification accuracy of motion-induced blurred images of calcified coronary plaques. The CNN architecture was trained by a large dataset of image motion artefacts generated by a moving robotic arm.

METHODS

Robotic arm and phantom

The moving cardiac phantom (Sim2D, QRM GmbH) consisted of a computer-controlled motion unit and a water container. The cardiac phantom was placed inside a chest phantom (Thorax, QRM GmbH) to simulate a chest environment (*Figure 1a*). A lever on the motion unit moved at a linear velocity programmed at 0-90 mm/s (*Figure 1b*), which was within the velocity range of in-vivo coronary arteries.[7,8] An ECG signal was simulated by the motion unit, and synchronized with the lever movement cycle.

Three artificial coronary arteries with calcified lesions made of hydroxyapatite were investigated (*Figure1c and Figure 1d*). Three densities of the calcified lesions were simulated, including high density of 800 Hounsfield Units (HU), medium density of 400HU and low density of 200HU, all at 120kV. Lesions of the same density were assembled to one artificial artery, containing three sizes (*Online Table 1*). These densities and sizes resembled the density and size of in-vivo coronary calcifications.[19,20] The background CT density of the artificial arteries was 50HU at 120kV, made of polyurethane resin, resembling in-vivo CT attenuation of blood. The lever with one attached artificial artery was placed inside the water container which was subsequently inserted into the chest phantom. The artificial artery was positioned parallel to the central axis of the CT examination bed (z-axis) and moved in the x-y plane (perpendicular to the z-axis). Each artificial artery was imaged one by one to avoid any overlapping artefacts.

CT imaging

Chest CT examinations were performed using two 64-slice CT systems (CT-A: Somatom Sensation 64, Siemens; CT-B: Brilliance 64, Philips). The acquisition protocol was derived from the Dutch-Belgian randomized lung cancer screening

3
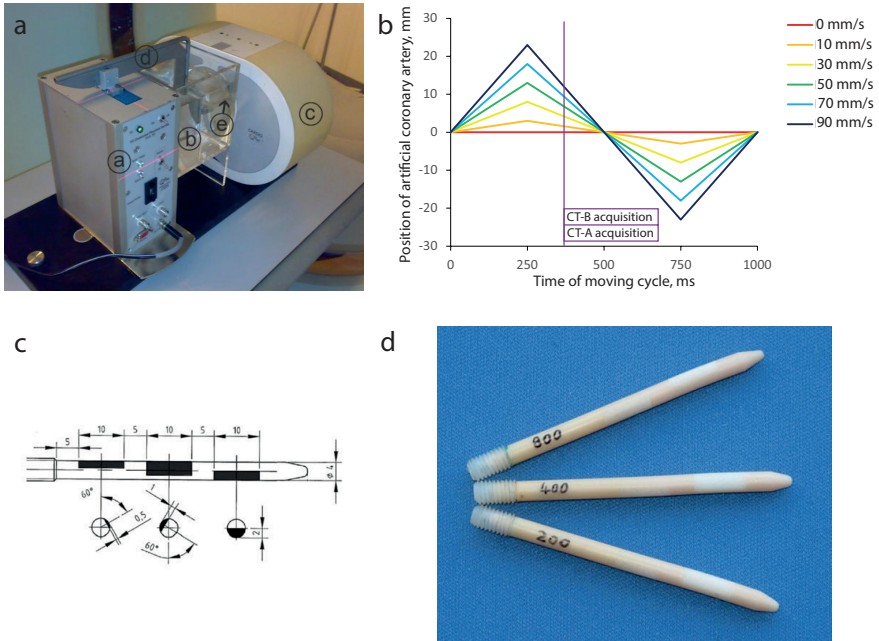
Figure 1. **A.** The moving robotic arm with a chest phantom, comprising of (a) a computer controlled motion unit, (b) a water container, (c) a chest phantom, (d) a lever, and (e) an artificial coronary artery **B.** Position of the artificial coronary artery as a function of time, for velocities of 10, 30, 50, 70 and 90 mm/s. The image acquisition windows of CT-A and CT-B in ECG-triggered acquisition mode are indicated, starting at 370ms and with a width of 330ms and 267ms, respectively, ensuring acquisition without turning points of the artificial coronary artery **C.** Artificial arteries with calcified lesions of high (800 Hounsfield Units [HU]), medium (400 HU) and low density (200 HU) in large, medium and small size. Schematic drawing of an artificial artery, the black shadows indicate the artificial calcified lesions, the size (in mm) and orientation of the artificial calcified lesions are indicated. **D.** Photograph of the three artificial arteries

trial (NELSON).[21] In order to scan the artificial coronary arteries at a constant motion velocity to avoid lever backtracking, the data acquisition was started at 37% of the moving cycle (*Figure 1b*). The two CT systems were routinely calibrated each working day. Details about image acquisition protocol are shown in *Online Table 2*. Each examination was repeated six times on both CT systems. Between each acquisition, the phantom with a robotic arm was moved by approximately 5 mm and rotated by approximately 5 degrees to simulate interscan variability.

For each calcified lesion, the Agatston, mass and volume score were semi-automatically quantified using commercially available software (Aquarius iNtuition, TeraRecon), with a default threshold of 130HU to define a calcification.[2]

3

Data preparation and image processing

The CT images were categorized into nine classes, each corresponding to one calcified plaque. We converted the DICOM images to grayscale Joint Photographic Experts Group (JPEG) format using a mediastinum window setting (window width 350HU, window level 40HU), cropped the images around the calcified plaque in a square shape of 20mm in side length, and resized them to a resolution of 299×299 pixels in order to make them compatible to the original Inception v3 network architecture (*Figure 2*).

Before training the CNN, we firstly augmented the images to increase the number of training images (*the Appendix*). By this process, of each image 30 images were generated.
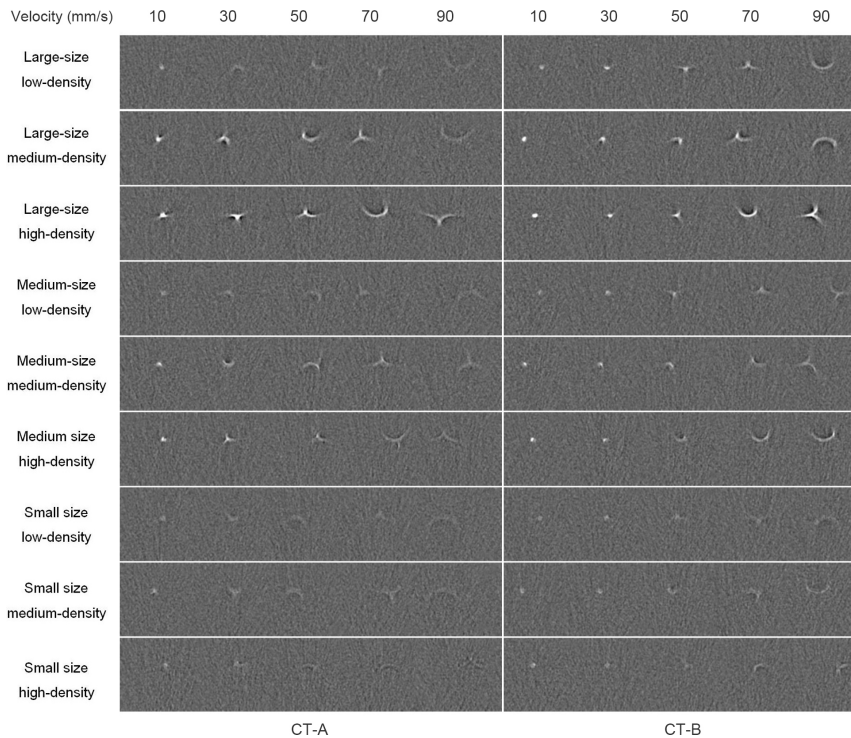
Figure 2. Representative images of nine classes of calcified plaques

Grouping outline

In order to evaluate the performance of the CNN model, a 6-fold cross validation method was used which is widely-applied in machine learning to test the generalization ability of the model.[16,22] In six subsequent deep learning procedures, the images of the five repetitive scans out of six were used as the training (90% of the images) and the validation (10%) dataset. The images of the other one scan were considered as the test dataset. The classification performance was calculated as the pooled result of six non-overlapping test datasets.

After image augmentation, each training dataset included 43,740 images [9(plaques)×6(speeds)×2(CTs)× 5(repetitive scans)×3(image slices)×30(by image augmentation)×90%]. Each corresponding validation dataset included 4,860 images. The test dataset was not augmented when evaluating the classification performance. The six test datasets included 1,944 images [9(plaques)×6(speeds)×2(CTs)×6 (by 6-fold cross validation)×3 (image slices)].

Training algorithm and environment

The deep learning was established based on the pre-trained Inception v3 CNN architecture, consisting of 316 layers (*Figure 3*).[16] We replaced the classifier layer from 1,000 to nine categories, and subsequently fine-tuned the parameters with our training dataset based on a back propagation method across all layers (*online Figure 1*). We performed the procedure using the deep learning toolbox (Matlab R2018b, MathWorks). The program worked on a workstation with a graphics processing unit (RTX2080Ti, nVidia). The details of deep learning algorithm and environment were showed in the Appendix.
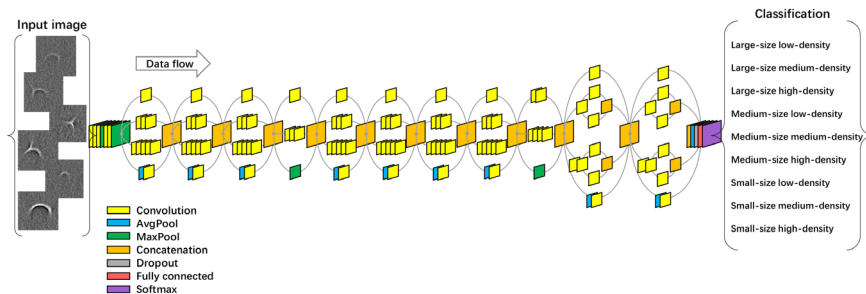
Figure 3. Architecture of the GoogLeNet Inception v3 deep convolutional neural network. The network consists of 316 layers, including convolution, average pool, max pool, concatenation, dropout, fully connection and softmax layer. An important feature of Inception v3 is factorization, which factorizes a big kernel into small ones, resulting in less parameters and higher speed in the training process to avoid overfitting

3

Inference method

The CNN inference result for each image of calcified plaque consisted of a one-dimensional numeric series with a length of nine, representing the matching probability for nine classifications. Since the length of each artificial calcified plaque was 10mm (*Figure 1c*), it encompassed three 3mm CT slices. The classification associated with the maximum matching probability of these three CT images was considered as the CNN corrected category.

Subsequently, the CNN corrected calcium score was defined as the score of each calcified plaque at rest, instead of the one with motion artefacts. A calcification by CNN correction was considered if the uncorrected score was zero, but the corrected score was non-zero.

Statistics

The normality of data was assessed by the Kolmogorov–Smirnov test. Data conforming to a normal distribution was expressed as mean ± standard deviation, otherwise expressed as median (1st, 3rd quartile). The ability of the CNN algorithm to correctly classify the calcified plaques with motion artefacts on the CT images was presented in a two-dimensional confusion matrix. This matrix consisted of actual (y axis) and predicted labels (x axis), and visualized the agreement between the actual and predicted category. The mean calcium score (Agatston, mass and volume score) of multiple measurements for each calcified plaque at rest were considered as the reference score. Measurement variation was calculated as: variation = $|(s-r)/r| \times 100\%$, where s is the calcium score of the calcified plaque at motion or the CNN corrected score, and r is the score at rest.

A paired-samples t test was used to compare the calcium scores of calcified plaques at rest and those in motion. A Wilcoxon signed-rank test was used to compare the calcium scores at rest and those by CNN correction, and to compare the score variations in motion and those by CNN correction. Sensitivity to detect calcifications with a CT density threshold of 130HU and by CNN correction was calculated by the percentage of non-zero scores among all the scores of an artificial plaque in the velocity range (0-90mm/s). A p-value of <0.05 was considered statistically significant. Statistical analyses were performed using a software package (SPSS 17.0, IBM).

3

RESULTS

Images of the nine calcified lesions were blurred by motion. The shapes of the motion artefact differed depending on the size and density of the calcified plaques. Increased blurring was observed when the velocity increased.

Agreement between actual and predicted label

The confusion matrix showed the agreement of the CNN classification between the actual and predicted category (*Figure 4*). The overall accuracy of the CNN classification was 79.2±6.1% for all nine artificial plaques. With respect to plaque size, the accuracy to predict the large-, medium- and small-sized plaques was 86.1±5.7%, 72.2±6.6% and 79.3±7.2%, respectively. Regarding plaque density, the accuracy for the high-, medium- and low-density plaques was 88.3±4.9%, 75.9±6.4% and 73.5±5.0%, respectively. The large-size high-density plaque showed the highest accuracy of 100%, whereas the lowest accuracy was 61.1±3.8% for the medium-size medium-density plaque.
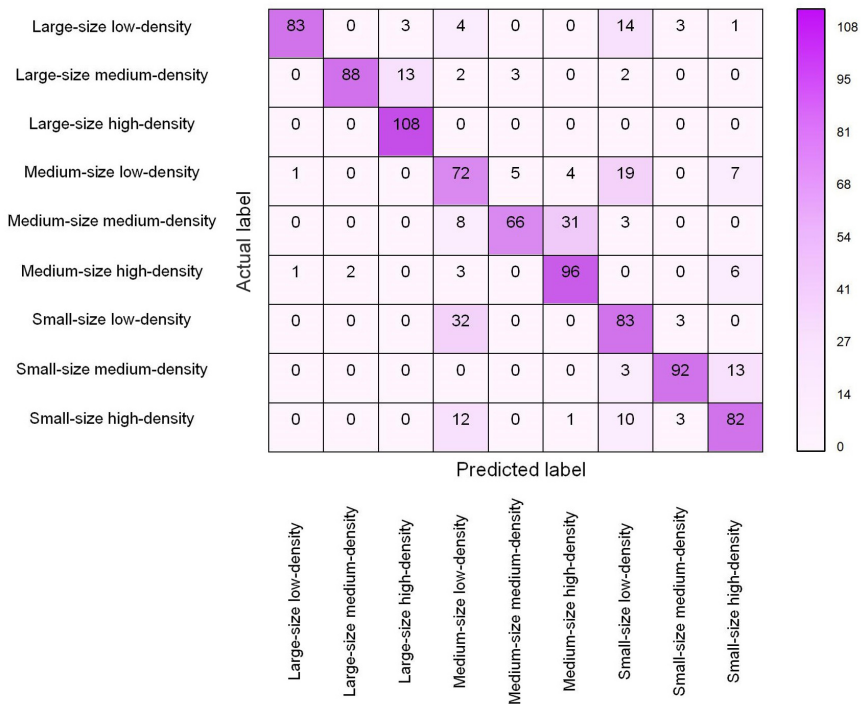
Figure 4. Confusion matrix of the CNN classification algorithm for nine categories of calcified plaques (velocity from 0 to 90mm/s) by 6-fold cross validation in the pooled test datasets, that visualizes the agreement between actual (vertical axis) and predicted category (horizontal axis). Each calcified plaque consists of three CT images. CNN = convolutional neural network

3

CNN classification accuracy and motion velocity

The accuracy of the CNN classification tended to decrease when the velocity increased. The accuracy was 85.6±4.6% for calcified plaques at rest. With increasing velocity, the accuracy was 80.1±5.2%, 80.9±4.9%, 81.8±6.0%, 75.5±4.6% and 68.2±6.5% at velocities of 10, 30, 50, 70 and 90 mm/s, respectively.

Variation of calcium scores

The Agatston scores for individual calcifications varied with velocity. The average Agatston scores of calcified plaques at rest, in motion and by CNN correction are shown in *Table 1*. The scores in motion were significantly different to those at rest (all p<0.001). Compared to the reference calcium score at rest, calcium scores for the large-size high-density lesion were overestimated, while the other lesions were underestimated (*Figure 5*). Significantly, the CNN corrected scores were similar to those at rest (all p>0.05). Despite of CNN correction, the small-size low-density plaques could not be measured, because its CT density was always under 130HU.

Table 1. Agatston scores of calcified plaques at rest, in motion and by CNN correction

Size	Density	Agatston score at rest, mean ± SD	Agatston score in motion, mean ± SD	Agatston score by CNN correction, median (1st, 3rd quartile)	P1 value	P2 value
Large	Low	35.2±6.2	22.3±13.2	38.0 (32.0, 40.5)	<0.001	0.392
	Medium	125.7±9.0	98.6±29.9	129.9 (118.8, 135.6)	<0.001	0.371
	High	171.5±3.9	304.7±95.5	171.5 (168.8, 175.0)	<0.001	1.000
Medium	Low	12.4±1.4	8.7±3.4	12.3 (12.0, 14.7)	<0.001	0.419
	Medium	50.5±1.8	27.5±17.1	50.7 (49.1, 52.5)	<0.001	0.914
	High	74.9±2.0	49.2±21.8	74.5 (73.2, 76.5)	<0.001	0.778
Small	Low	NM	NM	NM		
	Medium	10.6±0.8	7.0±3.5	10.9 (10.2, 11.7)	<0.001	0.245
	High	14.7±0.7	9.0±4.9	14.7 (14.0, 15.0)	<0.001	0.567

Note. P_1 values were for paired-samples t test, which compared Agatston scores of calcified plaques at rest and those in motion. P_2 values were for Wilcoxon signed-rank test, which compared Agatston scores at rest and those by CNN correction. CNN = convolutional neural network; SD = standard deviation; NM = not measurable, because the CT value < 130 Hounsfield units

The variation of the Agatston scores of plaques in motion and by the CNN correction is shown in *Table 2*. Compared to the Agatston score at rest, the remaining eight plaques showed an overall median score variation of 37.8% (1st, 3rd quartile: 10.5%, 68.8%) in velocity range from 10 mm/s to 90 mm/s. After corrected by

3
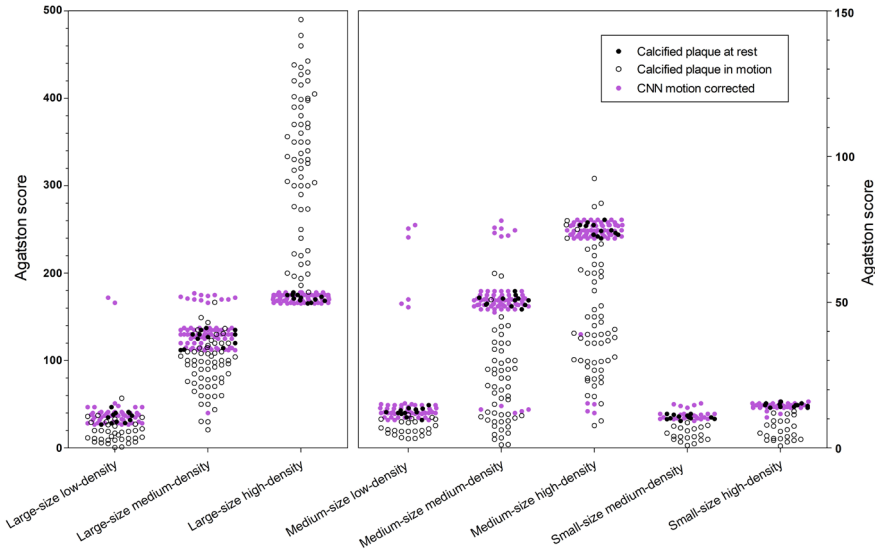
Figure 5. Grouped plots of Agatston score derived from calcified plaques at rest and in motion, and from CNN motion correction. The majority of CNN corrected scores fell into the scores at rest if the CNN classification of calcified plaque was correct, but some outliers are shown if the classification was incorrect. The score of the outlier was incorrectly considered as the score of the misclassified plaque at rest. Small-size low-density plaque is not measurable, because the CT value is under 130 Hounsfield units. CNN = convolutional neural network

Table 2. Variation between Agatston scores of calcified plaques in motion and by CNN correction, using the mean score at rest as reference

Size	Density	Variation in motion, median (1st, 3rd quartile)	Variation by CNN correction, median (1st, 3rd quartile)	P value
Large	Low	43.4% (16.3%, 68.8%)	13.9% (7.7%, 19.8%)	<0.001
	Medium	20.5% (9.0%, 36.7%)	7.9% (3.4%, 10.4%)	<0.001
	High	85.9% (21.0%, 123.0%)	2.0% (0.9%, 2.5%)	<0.001
Medium	Low	27.2% (10.4%, 53.3%)	7.3% (2.9%, 15.3%)	<0.001
	Medium	46.9% (17.7%, 77.3%)	2.8% (1.6%, 5.7%)	<0.001
	High	39.9% (8.7%, 60.0%)	2.3% (1.7%, 3.3%)	<0.001
Small	Low	NM	NM	
	Medium	32.2% (7.1%, 63.3%)	7.1% (3.9%, 0.7%)	<0.001
	High	36.8% (5.4%, 71.8%)	3.7% (1.3%, 5.7%)	<0.001
Overall		37.8% (10.5%, 68.8%)	3.7% (1.9%, 9.1%)	<0.001

Note. P values were for Wilcoxon signed-rank test, which compared variation between Agatston scores of calcified plaques in motion and those by CNN correction. CNN = convolutional neural network; NM = not measurable, because the CT value < 130 Hounsfield units

3

CNN classification, the overall median variation largely decreased to 3.7% (1.9%, 9.1%) (p<0.001). CNN correction showed a larger decrease in variation in high-density plaques than in medium- and low-density ones. The greatest impact was on the large-size high-density plaque, whose score variation declined from 85.9% (21.0%, 123.0%) to 2.0% (0.9%, 2.5%) (p<0.001).

Compared to the mass score at rest, the overall median score variation was 40.2% (20.7%, 60.5%). After corrected by CNN classification, the overall median variation largely decreased to 4.3% (1.8%, 10.1%) (p<0.001) (*Online Table 3 and 4*). For volume score, CNN classification largely decreased the score variation from 43.8% (22.9%, 67.2%) to 4.7% (2.5%, 12.5%) (p<0.001) (*Online Table 5 and 6*).

Detection of calcifications

Using a CT density threshold of 130HU in nontriggered CT, calcium scores of four large calcifications (Agatston score ≥50.5 at rest) were above zero at all velocities (*Table 3*). However, as the velocity increased, calcium scores were zero in four small calcifications (Agatston score <50.5), because the CT density dropped below the 130 HU threshold. Independent of velocity, low-density small-sized plaques' calcium scores were always zero. The overall sensitivity to detect calcifications for all nine calcified plaques was 65%. Using CNN correction, the overall sensitivity increased to 85%, because CNN successfully classified some plaques with motion artefacts (CT density <130HU) into the right category, subsequently made these plaques measurable.

DISCUSSION

In this study, a deep CNN-based correction algorithm was used to reduce variation in Agatston score from 38% to 3.7% of a dynamic phantom containing coronary calcifications. In addition, the sensitivity was improved from 65% to 85%. Although in a realistic in-vivo situation, shape, size, density and motion patterns of calcified plaques vary from patient to patient, this study simulated typical morphological and kinematical in-vivo features, which is a fundamental step towards a real clinical scenario. Therefore, a large-scale study on realistic patient data covering variable plaque at rest and in motion is potentially feasible to correct for calcium scores in motion.

Machine learning approaches were used for automatic calcium scoring and motion correction on cardiac CT .[23–29] De Vos et al. employed two ConvNets for automatic calcium scoring in chest CT, and achieved an intra-class correlation coefficient of

Table 3. Sensitivity to detect coronary calcifications in nontriggered CT with a CT density threshold of 130 HU and by CNN correction

Size	Density	CT density threshold of 130HU							CNN correction						
		Velocity, mm/s						Sensi tivity	Velocity, mm/s						Sensi tivity
		0	10	30	50	70	90		0	10	30	50	70	90	
Large	Low	■	■	■	50%	○	○	58%	■	■	■	89%	89%	83%	94%
	Medium	■	■	■	■	■	■	100%	■	■	■	■	■	97%	99%
	High	■	■	■	■	■	■	100%	■	■	■	■	■	■	100%
Medium	Low	■	■	52%	○	○	○	42%	■	■	83%	83%	81%	75%	87%
	Medium	■	■	■	■	■	■	100%	■	■	■	■	■	■	100%
	High	■	■	■	■	■	■	100%	■	■	■	■	■	94%	99%
Small	Low	○	○	○	○	○	○	0%	○	○	○	○	○	○	0%
	Medium	■	■	○	○	○	○	33%	■	■	89%	■	78%	67%	89%
	High	■	■	60%	40%	○	○	50%	■	■	89%	■	89%	97%	96%
Overall								65%							85%

Note. Sensitivity to detect calcified plaques was calculated by the percentage of non-zero scores among all the scores of an artificial calcified plaque in the velocity range (0-90mm/s). Black box indicates a sensitivity of 100%. Empty circle indicates a sensitivity of 0%. HU = Hounsfield units; CNN = convolutional neural network

0.98 between automatic and manual scoring.[23] Lossau et al. developed a coronary forward artifact model for coronary CT angiography using CNN, and proved CNNs are able to identify artifact pattern.[28] In our study the CNN correction largely decreased calcium score variation, which had not been previously reported. This large decrease was feasible because for the CNN algorithm the calcium score of the stationary plaque at rest was used, rather than the score of the plaque in motion. Hereby, the influence of motion was minimized. In this study, the overall accuracy of the CNN classification was 79%, which was the basis to achieve a very low scoring variation. Previously, Greuter et al. proposed a function to adjust calcium scores for motion and achieved a 35% decrease of scoring deviation over all velocities using a phantom.[30] This phantom study enabled a research pathway to correct calcium score in-vivo.

3

Coronary motion in our study was comparable to the velocities of coronary arteries in vivo. Several studies showed considerable velocity variations during the R-R interval, approximately 15–30 mm/s on left anterior descending (LAD), 30–70 mm/s on left circumflex (LCx) and 40–90 mm/s on right coronary artery (RCA).[7,8,31] When performing nontriggered chest CT, data are acquired over the entire cardiac cycle, and therefore it is uncertain at what coronary velocity the individual calcified plaque is scanned. Hence a velocity range of 0–90 mm/s was simulated in this study.

Coronary calcium scores are influenced by heart rate when acquired on nontriggered chest CT scans, and has also been reported in studies using ECG-triggering .[5,12] Generally motion artefacts lead to overestimation of volume and underestimation of CT density, and subsequently leads to a lower reproducibility.[30,32]

A recent phantom study showed calcium score differences up to 50% when heart rate was increased from 60 to >75 bpm.[12] Thereby influencing the individual patient risk assessment as hazards ratios for adverse cardiovascular events increase with increasing calcium score, which may even lead to alteration in treatment plans.[33]

Motion is equally important in cases with no calcium to very low coronary calcium scores, as the absence of coronary calcium is considered to be an extremely low risk of coronary events in the next ten years.[34] In the current study, motion at increasing velocity led to a more blurred image of the calcification with a peak attenuation under the threshold (<130HU), thereby leading to increasing false negatives. CNN correction largely decreased the false negative rate by classifying the blurred plaque into the correct category, and thereby more effectively identified the patients in need for further follow-up and/or treatment. Further improvement for detection of low- and medium-density plaques may strengthen the practical value.

Because aging and smoking are shared major risk factors between cardiovascular diseases and lung cancer, the candidate population for lung cancer CT screening is also at risk for cardiovascular diseases. Coronary calcifications were found in more than 70% of the participants in the NELSON lung cancer screening trial.[35] If the calcium score can accurately be derived from nontriggered chest CT examinations, there may be a great potential for additional cardiovascular risk assessment.

There are limitations to this study. First, we simulated linear movement of artificial calcified lesions in the x-y plane, in contrast to the diverse motion patterns and irregular shape of lesions in coronary arteries in vivo.[8] Further phantom studies with movement in three-dimensional space and patient studies are needed to

confirm the findings in this study. Second, the CT acquisition protocol in this study was derived from a lung cancer screening trial (NELSON). Variability potentially exists if a CT scanner with different temporal resolution is used.

CONCLUSION

In this experimental phantom study, a deep CNN architecture trained by CT images of motion artefacts showed the ability to correct coronary calcium scores from blurred images. A correction algorithm based on deep CNN can be used for a 10-fold reduction in Agatston score variations from 38% to 3.7% of moving coronary calcified plaques, and to improve the sensitivity from 65% to 85% for the detection of calcifications. Considering the extensively utilized nontriggered chest CT, this study provides a method to improve its accuracy for coronary calcium scores, which may allow for additional cardiovascular risk assessment in these already acquired images.

3

REFERENCES

1. Budoff MJ, Shaw LJ, Liu ST, et al. Long-term prognosis associated with coronary calcification: observations from a registry of 25,253 patients. J Am Coll Cardiol. 2007;49(18):1860-1870. doi:10.1016/j.jacc.2006.10.079

2. Berman DS, Arnson Y, Rozanski A. Coronary Artery Calcium Scanning: The Agatston Score and Beyond. JACC Cardiovasc Imaging. 2016;9(12):1417-1419. doi:10.1016/j.jcmg.2016.05.020

3. Greenland P, Blaha MJ, Budoff MJ, Erbel R, Watson KE. Coronary Calcium Score and Cardiovascular Risk. J Am Coll Cardiol. 2018;72(4):434-447. doi:10.1016/j.jacc.2018.05.027

4. Hecht HS, Blaha MJ, Kazerooni EA, et al. CAC-DRS: Coronary Artery Calcium Data and Reporting System. An expert consensus document of the Society of Cardiovascular Computed Tomography (SCCT). J Cardiovasc Comput Tomogr. 2018;12(3):185-191. doi:10.1016/j.jcct.2018.03.008

5. Xie X, Greuter MJW, Groen JM, et al. Can nontriggered thoracic CT be used for coronary artery calcium scoring? A phantom study. Med Phys. 2013;40(8):81915. doi:10.1118/1.4813904

6. Vonder M, Pelgrim GJ, Huijsse SEM, et al. Coronary artery calcium quantification on first, second and third generation dual source CT: A comparison study. J Cardiovasc Comput Tomogr. 2017;11(6):444-448. doi:10.1016/j.jcct.2017.09.002

7. Achenbach S, Ropers D, Holle J, Muschiol G, Daniel WG, Moshage W. In-plane coronary arterial motion velocity: measurement with electron-beam CT. Radiology. 2000;216(2):457-463. doi:10.1148/radiology.216.2.r00au19457

8. Husmann L, Leschka S, Desbiolles L, et al. Coronary Artery Motion and Cardiac Phases: Dependency on Heart Rate—Implications for CT Image Reconstruction. Radiology. 2007;245(2):567-576.

9. Xie X, Zhao Y, de Bock GH, et al. Validation and prognosis of coronary artery calcium scoring in nontriggered thoracic computed tomography: systematic review and meta-analysis. Circ Cardiovasc Imaging. 2013;6(4):514-521. doi:10.1161/CIRCIMAGING.113.000092

10. Mets OM, Vliegenthart R, Gondrie MJ, et al. Lung cancer screening CT-based prediction of cardiovascular events. JACC Cardiovasc Imaging. 2013;6(8):899-907. doi:10.1016/j.jcmg.2013.02.008

11. Hecht H, Blaha MJ, Berman DS, et al. Clinical indications for coronary artery calcium scoring in asymptomatic patients: Expert consensus statement from the Society of Cardiovascular Computed Tomography. J Cardiovasc Comput Tomogr. 2017;11(2):157-168. doi:10.1016/j.jcct.2017.02.010

12. van der Werf NR, Willemink MJ, Willems TP, Vliegenthart R, Greuter MJW, Leiner T. Influence of heart rate on coronary calcium scores: a multi-manufacturer phantom study. Int J Cardiovasc Imaging. 2017;34(6):959-966. doi:10.1007/s10554-017-1293-x

13. Ma H, Gros E, Szabo A, et al. Evaluation of motion artifact metrics for coronary CT angiography. Med Phys. 2018;45(2):687-702. doi:10.1002/mp.12720

14. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436-444. doi:10.1038/nature14539

3

15. He Y, Guo J, Ding X, et al. Convolutional neural network to predict the local recurrence of giant cell tumor of bone after curettage based on pre-surgery magnetic resonance images. Eur Radiol. 2019;29(10):5441-5451. doi:10.1007/s00330-019-06082-2

16. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. Nature. 2017;542(7639):115-118. doi:10.1038/nature21056

17. Zhao W, Yang J, Sun Y, et al. 3D Deep Learning from CT Scans Predicts Tumor Invasiveness of Subcentimeter Pulmonary Adenocarcinomas. Cancer Res. 2018;78(24):6881-6889. doi:10.1158/0008-5472.CAN-18-0696

18. Šprem J, de Vos BD, Lessmann N, de Jong PA, Viergever MA, Išgum I. Impact of automatically detected motion artifacts on coronary calcium scoring in chest computed tomography. J Med imaging (Bellingham, Wash). 2018;5(4):44007. doi:10.1117/1.JMI.5.4.044007

19. Scott DS, Arora UK, Farb A, Virmani R, Weissman NJ. Pathologic validation of a new method to quantify coronary calcific deposits in vivo using intravascular ultrasound. Am J Cardiol. 2000;85(1):37-40. doi:10.1016/s0002-9149(99)00603-7

20. Moselewski F, Ferencik M, Achenbach S, et al. Threshold-dependent variability of coronary artery calcification measurements -- implications for contrast-enhanced multi-detector row-computed tomography. Eur J Radiol. 2006;57(3):390-395. doi:10.1016/j.ejrad.2005.12.026

21. Xie X, Zhao Y, Snijder RA, et al. Sensitivity and accuracy of volumetry of pulmonary nodules on low-dose 16- and 64-row multi-detector CT: an anthropomorphic phantom study. Eur Radiol. 2013;23(1):139-147. doi:10.1007/s00330-012-2570-7

22. Blagus R, Lusa L. Joint use of over- and under-sampling techniques and cross-validation for the development and assessment of prediction models. BMC Bioinformatics. 2015;16(1):363. doi:10.1186/s12859-015-0784-9

23. de Vos BD, Wolterink JM, Leiner T, de Jong PA, Lessmann N, Isgum I. Direct Automatic Coronary Calcium Scoring in Cardiac and Chest CT. IEEE Trans Med Imaging. 2019;38(9):2127-2138. doi:10.1109/TMI.2019.2899534

24. Xie Y, Liu S, Miller A, et al. Coronary artery calcification identification and labeling in low-dose chest CT images. In: Proc.SPIE. Vol 10134. ; 2017. doi:10.1117/12.2254125

25. González G, Washko GR, Estépar RSJ. Automated Agatston score computation in large dataset of non ECG-gated chest computed tomography. Proceedings IEEE Int Symp Biomed Imaging. 2016;2016:53-57. doi:10.1109/ISBI.2016.7493209

26. Lessmann N, van Ginneken B, Zreik M, et al. Automatic Calcium Scoring in Low-Dose Chest CT Using Deep Neural Networks With Dilated Convolutions. IEEE Trans Med Imaging. 2018;37(2):615—625. doi:10.1109/tmi.2017.2769839

27. Isgum I, Prokop M, Niemeijer M, Viergever MA, van Ginneken B. Automatic coronary calcium scoring in low-dose chest computed tomography. IEEE Trans Med Imaging. 2012;31(12):2322-2334. doi:10.1109/TMI.2012.2216889

28. Lossau T, Nickisch H, Wissel T, et al. Motion artifact recognition and quantification in coronary CT angiography using convolutional neural networks. Med Image Anal. 2019;52:68-79. doi:10.1016/j.media.2018.11.003

29. Lossau Née Elss T, Nickisch H, Wissel T, et al. Motion estimation and correction in cardiac CT angiography images using convolutional neural networks. Comput Med imaging Graph Off J Comput Med Imaging Soc. 2019;76:101640. doi:10.1016/j.compmedimag.2019.06.001

30. Greuter MJW, Groen JM, Nicolai LJ, Dijkstra H, Oudkerk M. A model for quantitative correction of coronary calcium scores on multidetector, dual source, and electron beam computed tomography for influences of linear motion, calcification density, and temporal resolution: A cardiac phantom study. Med Phys. 2009;36(11):5079-5088. doi:10.1118/1.3213536

31. Contijoch F, Stayman JW, McVeigh ER. The impact of small motion on the visualization of coronary vessels and lesions in cardiac CT: A simulation study. Med Phys. 2017;44(7):3512-3524. doi:10.1002/mp.12295

32. Jacobs PCA, Isgum I, Gondrie MJA, et al. Coronary artery calcification scoring in low-dose ungated CT screening for lung cancer: interscan agreement. AJR Am J Roentgenol. 2010;194(5):1244-1249. doi:10.2214/AJR.09.3047

33. Detrano R, Guerci AD, Carr JJ, et al. Coronary calcium as a predictor of coronary events in four racial or ethnic groups. N Engl J Med. 2008;358(13):1336-1345. doi:10.1056/NEJMoa072100

34. Blaha MJ, Cainzos-Achirica M, Greenland P, et al. Role of Coronary Artery Calcium Score of Zero and Other Negative Risk Markers for Cardiovascular Disease: The Multi-Ethnic Study of Atherosclerosis (MESA). Circulation. 2016;133(9):849-858. doi:10.1161/CIRCULATIONAHA.115.018524

35. Jacobs PC, Gondrie MJA, van der Graaf Y, et al. Coronary artery calcium can predict all-cause mortality and cardiovascular events on low-dose CT screening for lung cancer. AJR Am J Roentgenol. 2012;198(3):505-511. doi:10.2214/AJR.10.5577

Online material is available via:
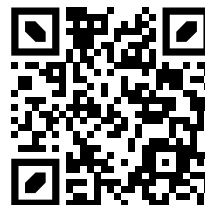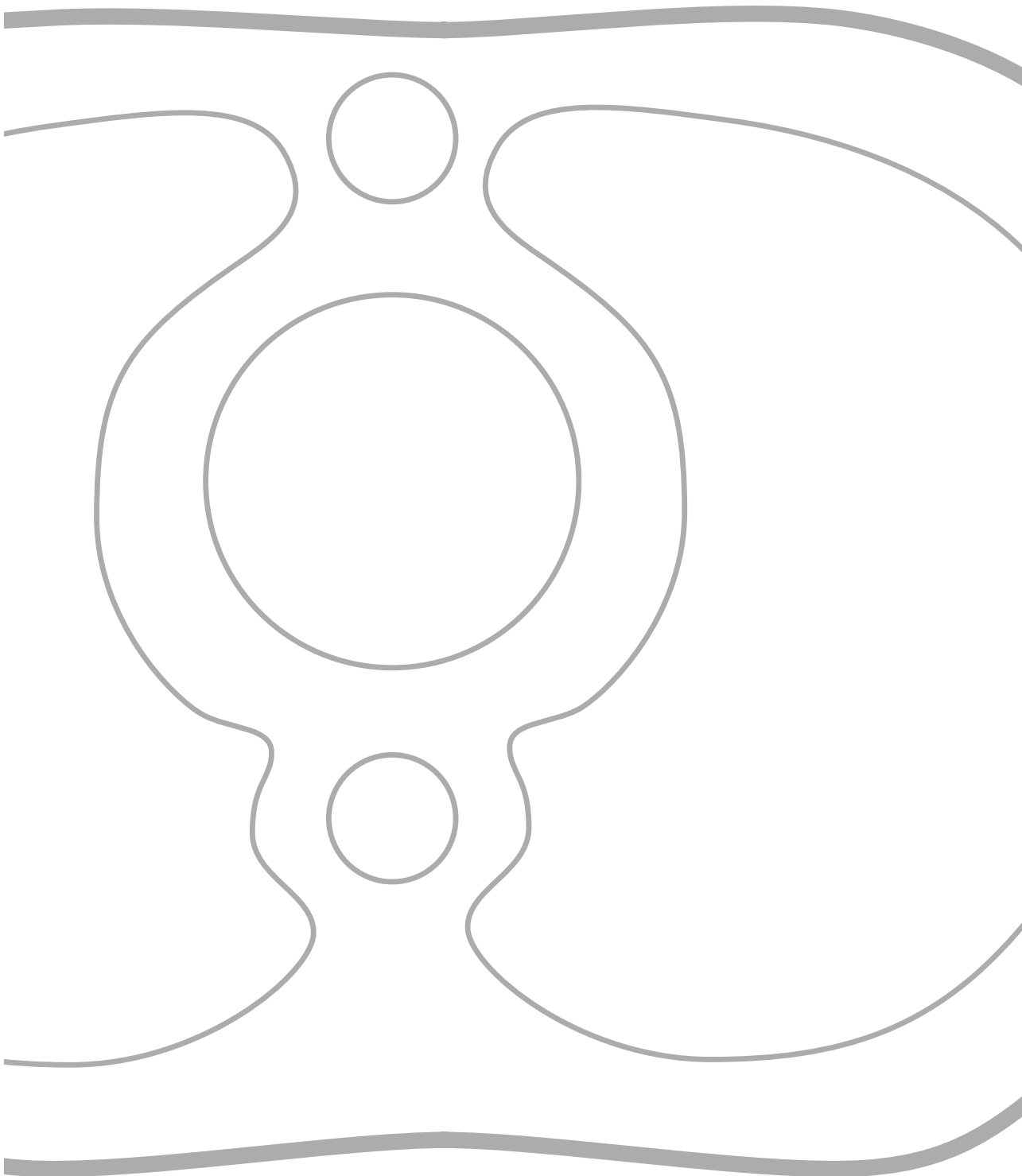
CHAPTER 4

Classification of moving coronary calcified plaques based on motion artifacts using convolutional neural networks: a robotic simulating study on influential factors

Magdalena Dobrolinska, MD

Niels R. van der Werf, MSc

Marcel J.W. Greuter, PhD

Beibei Jiang, MD

Riemer H.J.A. Slart, MD PhD

Xueqian Xie, PhD

4

ABSTRACT

Background

Motion artifacts affect the images of coronary calcified plaques. This study utilized convolutional neural networks (CNNs) to classify the motion-contaminated images of moving coronary calcified plaques and to determine the influential factors for the classification performance.

Methods

Two artificial coronary arteries containing four artificial plaques of different densities were placed on a robotic arm in an anthropomorphic thorax phantom. Each artery moved linearly at velocities ranging from 0 to 60 mm/s. CT examinations were performed with four state-of-the-art CT systems. All images were reconstructed with filtered back projection and at least three levels of iterative reconstruction. Each examination was performed at 100%, 80% and 40% radiation dose. Three deep CNN architectures were used for training the classification models. A five-fold cross-validation procedure was applied to validate the models.

Results

The accuracy of the CNN classification was 90.2±3.1%, 90.6±3.5%, and 90.1±3.2% for the artificial plaques using Inception v3, ResNet101 and DenseNet201 CNN architectures, respectively. In the multivariate analysis, higher density and increasing velocity were significantly associated with higher classification accuracy (all P<0.001). The classification accuracy in all three CNN architectures was not affected by CT system, radiation dose or image reconstruction method (all P>0.05).

Conclusion

The CNN achieved a high accuracy of 90% when classifying the motion-contaminated images into the actual category, regardless of different vendors, velocities, radiation doses, and reconstruction algorithms, which indicates the potential value of using a CNN to correct calcium scores.

INTRODUCTION

Noninvasive assessment of coronary artery disease (CAD) has gained substantial interest, due to large number of global deaths.[1,2] With the introduction of CT, the burden of coronary atherosclerosis can be expressed as a coronary artery calcium (CAC) score, generally expressed as the Agatston score, which is a strong independent predictor of coronary events in intermediate-risk asymptomatic patients.[3–5] Traditionally, the Agatston score is obtained using ECG-triggered non-contrast CT.[6]

In the US, almost 7.1 million non-ECG-triggered chest CT scans are performed each year.[7,8] Because non-ECG-triggered CT demonstrated comparable results in CAC detection to ECG-triggered CT, these scans also have the potential to assess the risk of CAD.[9] Recently, the Society of Cardiovascular CT and the Society of Thoracic Radiology recommended a CAC evaluation of every non-ECG-triggered chest CT examination as a Class I indication.[10]

Whether using ECG-triggered cardiac or non-triggered chest scans, an accompanied limitation is the presence of motion artifacts, which considerably decreases the accuracy of CAC detection and quantification.[11] To decrease the motion artifacts in ECG-triggered CT, the temporal resolution should be shorter than 10% of one cardiac cycle time.[12] Even in the relatively low motion phase of 60-70% in the R-R interval, the velocity of the coronary arteries is still approximately 10 mm/s, even at a heart rate <60 bpm.[13] However, 50% of ECG-triggered cardiac CT scans are performed with a heart rate>70 bpm, which implies a coronary velocity during the CT acquisition phase of at least 30 mm/s.[14] Notwithstanding, the influence of motion is even greater in a non-ECG-triggered chest CT, where coronary motion is up to 60 mm/s.

In a recent review, Waltz et al. concluded that the convolutional neural network (CNN) has expanded the role of automatic detection and measurement of CAC.[15] Šprem et al. proposed a CNN-based method to identify calcified plaques in-vivo that were severely affected by cardiac motion and reached an accuracy of 85.2%.[16] In a multicenter study, Eng et al. showed sensitivities of 71-94% and positive predictive values in the range of 88-100% to detect CAC on non-triggered chest CT.[17] Because motion artifacts are inevitable in the evaluation of coronary calcification, researchers have started to use CNN to alleviate motion artifacts and improve the robustness of CAC scoring. In an ex-vivo experimental study, Zhang et al. used CNN to correct coronary calcium scores and largely reduced

4

Agatston score variations from 38% to 3.7%.[18] However, the generalization ability of CNN for motion artifact recognition heavily depends on a variety of influential factors, besides coronary motion artifacts, also on other technical factors such as CT vendor, radiation dose, and reconstruction kernel.

Before applying a CNN to correct CAC scores in clinical practice in the future, we first conducted an experimental study to simulate motion artifacts of moving coronary calcified plaques using a coronary artery chest phantom, and second, established three CNN architectures to classify the motion-contaminated images of moving coronary calcified plaques, and determined the influential factors for their classification performance. The current study is a first step towards patient-specific motion artifact recognition which could be used to correct CAC scores in the future.

METHODS

Artificial coronary arteries containing cylindrical calcifications moved inside a water container, which was placed at the center of an anthropomorphic chest phantom (QRM-Chest, QRM, Moehrendorf, Germany) (*Figure 1*). Inside a shell of tissue-equivalent material, this phantom contained a spine insert and artificial lungs. To mimic an average patient, an extension ring of fat-equivalent material was placed around the phantom to increase the outer dimension to 400×300 mm (QRM Extension Ring L, QRM). Movement of the arteries inside the phantom was performed by a computer-controlled lever (QRM-Sim2D, QRM). In total, 7 velocities were assessed for the current study. Each velocity was constant during the scan phase. The used velocities ranged from 0 to 60 mm/s with an increment of 10 mm/s. The movement was in the horizontal x-direction, perpendicular to the scan direction. To acquire data during the linear motion of the artificial calcifications, the ECG-trigger of the robotic arm was used, and scans were performed at 60% of the artificial R-R interval. Two artificial coronary arteries were used, each containing two calcium hydroxyapatite (HA) calcifications of equal dimensions: diameter 5.0±0.1 mm and length 10.0±0.1 mm. One artificial coronary artery contained two calcifications with densities of 196±3 and 380±2 mgHA/cm3 (physical mass score equal 38 mg, 74 mg respectively), whereas the other contained two calcifications with densities of 408±2 and 800±2 mgHA/cm3 (physical mass score equal 80 mg, 157 mg respectively). This corresponded to one low, two medium, and one high coronary plaque burden, respectively (*Table S1*).
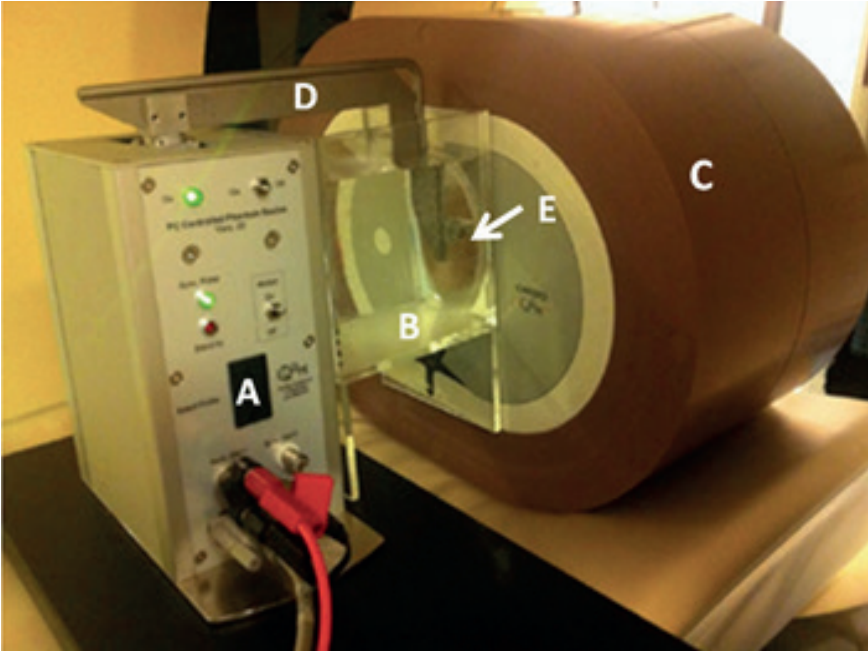
Figure 1 The moving robotic arm with a thoracic phantom. The thoracic phantom includes (**A**) a computer-controlled motion unit, (**B**) a water container, (**C**) a thoracic phantom, (**D**) a lever, and (**E**) an artificial coronary artery.

CT imaging

Thoracic CT examinations were performed using four CT systems (CT 750 HD, GE Healthcare; Brilliance iCT, Philips Healthcare; Somatom Definition Flash, Siemens Healthineers and Aquilion One, Canon Medical Systems). In the remainder of this paper, the four CT systems are denoted as CT-A to CT-D, respectively (*Table S2*). All images were reconstructed using filtered back projection (FBP) and three levels of iterative reconstruction (IR). Each examination was performed at a clinical radiation dose. Subsequently, the radiation dose was reduced by 40% and 80%.[19] Each combination of acquisition settings was repeated five times on every CT scan. In between each scan, the phantom was randomly translated by 2 mm.

Cross-validation

Since the heterogeneity of the images mainly originated from different vendors of CT, as well as different acquisition protocols, including dose levels and image reconstruction kernels, we conducted a k-fold cross-validation across each of these

variables. K-fold cross-validation is a widely used resampling procedure to evaluate machine learning models.[20] The k parameter of this process represents the number of groups into which all images are divided. In this study, 4 CT systems were used. In the 4 deep learning processes, CT images of 3 CT systems were selected as the training dataset, and the images of the remaining CT were used as the test dataset. In the same way, we performed a 3-fold cross-validation for dose level and a 4-fold cross-validation for image reconstruction kernel. The classification performance of the model is the average performance of these cross-validations.

Data preparation and image processing

The CT images were categorized into four classes, each corresponding to one calcified plaque, and they included images with different vendors, velocities, radiation doses, and reconstruction algorithms. We remapped the images using a mediastinum window setting (window width 350HU and window level 40HU) to make the images optimal for observation. After defining the calcified plaque manually in CT images, the patch images of calcified plaques were automatically cropped by an in-house developed script based on imcrop function of the Image Processing Toolbox (MATLAB R2020b, MathWorks) and resized to 299×299.

Before training the CNN, we first augmented the images using an embedded function to increase the number of training images. The data augmentation was achieved by performing geometric transformations in order to train a robust model, which was invariant to such transformations. Each image was randomly rotated from 0 to 359 degrees, zoomed on with a random aspect ratio ranging from 0.9 to 1.1, translocated from -30 to 30 pixels, and flipped vertically and/or horizontally with a probability of 0.5. In this way, each original image was augmented to 30 images, resulting in 4 (plaques) × 4 (vendors) × 7 (velocities) × 3 (doses) × 4 (kernels) × 5 (repetitions) × 30 (augmentation) =201,600 images in total.

Training algorithm and environment

Three deep CNN architectures were used, i.e., Inception v3, ResNet101, and DenseNet201 (*Tables S3-S5*), which are representative in deep learning. The network architectures consisted of 316, 347, and 709 layers, respectively. We adopted the inception and residual architectures because they increase the accuracy by using efficient and deep networks compared to the previous serial CNN architectures.[21,22] The DenseNet architecture is a logical extension that optimizes the networks by connecting each layer to every other layer in a feed-forward fashion to strengthen feature propagation.[23]

We replaced the original classifier layer consisting of 1,000 categories by adapting it to the ImageNet dataset with four categories and subsequently fine-tuned the parameters with our training dataset using a back-propagation method across all layers to optimize the networks. The mini-batch sizes were 60, 50, and 25 for these three CNN architectures, respectively, depending on the graphical memory footprints of the computer. The number of training epochs was 15 for the three models, at which point the training accuracy was close to the upper limit. The training dataset was shuffled between each epoch. We used stochastic gradient descent with momentum for the mini-batch gradient descent, which is an iterative method to minimize the result of the loss function of the CNN architectures; thus, it can be used to find suitable and optimized values of network parameters. All layers of the network were fine-tuned using the same global learning rate of 0.0003.

We performed the training and testing procedure using the Deep Learning Toolbox (MATLAB R2020b, MathWorks). The program was implemented on a workstation with a graphics processing unit (RTX 2080Ti, Nvidia).

Inference method

The CNN inference result for each image of calcified plaque consisted of a one-dimensional numeric series with a length of four, representing the matching probability of the four classifications. Since the length of each artificial calcified plaque was 10 mm, it encompassed at least 3 CT slices at a slice thickness of 3 mm. The three central slices of each calcification were selected manually by a radiologist with 20 years of experience in cardiac imaging. In the test dataset, the average matching probability of these three images was considered as the matching probability of this plaque. Among the four matching probabilities for the four calcifications, the classification of the highest matching probability was considered as the CNN's classification category.

Statistics

The ground truth of the CNN algorithms was the physical artificial calcified plaque. It was considered as true positive if the CNNs correctly classified the motion-contaminated images into the actual plaque, that had generated blurred images. We evaluated the classification accuracy and F1 score to represent CNN's classification performance. The F1 score is a weighted average of the precision and recall, where an F1 score reaches its best value at 1 and worst at 0. Precision, also called the positive predictive value, is the proportion of positive results that truly are positive. Recall is also called sensitivity.

4

The ROC curve of the CNN's classification of the calcified plaques with motion artifacts and the area under the curve (AUC) were calculated to access the CNN's classification performance. The CNN's matching probability was used as an index variable and the CNN's classification correctness was used as a reference variable. A multivariate linear regression model was used to synthesize the results of three CNN models.[24]

The association between the correctness and potentially associated factors (plaque density, CT vendor, motion velocity, dose level, and reconstruction method) was evaluated using Spearman's correlation coefficients. Because the classification accuracy was analyzed simultaneously with other variables (density, CT vendor, velocity, dose, and reconstruction algorithm), multivariate analysis was also used to access the association between the correctness and these factors.

The normally distributed outcome parameters are given as the mean values with standard deviations, and non-normally distributed parameters were given as median values with 95% confidence intervals. A p-value <0.05 was considered statistically significant. Statistical analyses were performed using a software package (MedCalc 15.8, MedCalc Software).

RESULTS

Subjective observation

The representative motion artifacts for the different CT systems, velocities, radiation doses, and reconstruction methods are shown in *Figures 2 and 3*. Increased blurring was observed as the velocity increased and radiation dose decreased. For each CT scanner, the use of IR resulted in decreased blurring. The smallest blurring was presented by CT-C.

Cross-validation

In the 4-fold cross-validation on CT system, the accuracy and F1 scores of CNN in the classification of 4 artificial plaques were high but variable (*Tables S6-S9*). The accuracies ranged from 84.8% to 95.3%, 83.8% to 95.7%, and 83.4% to 96.8% for Inception v3, ResNet101, and DenseNet201, respectively. The F1 scores ranged from 0.849 to 0.966, 0.859 to 0.969, and 0.883 to 0.969, respectively.
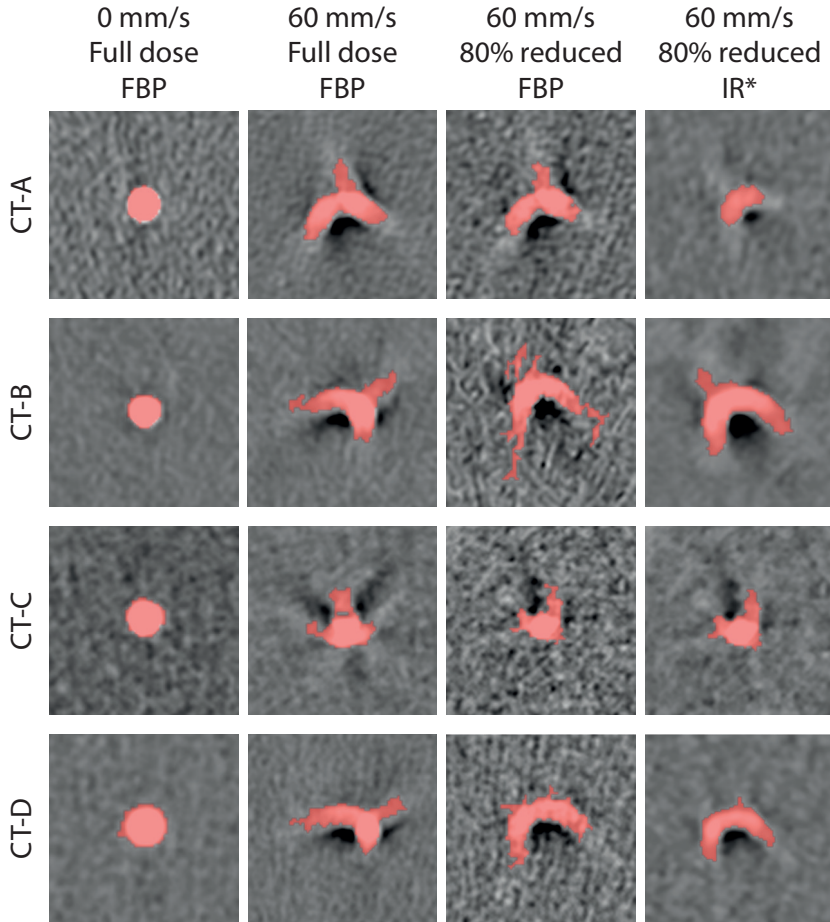
Figure 2 Representative images of the motion artifacts for all CT systems at all velocities, clinical full radiation dose and FBP. Window center was 90 HU and window width was 750 HU. * for each CT system, the highest available level of IR was used. FBP=filtered back projection;

In the 3-fold cross-validation on dose level, the accuracy and F1 scores of CNN were also high but variable. The accuracies ranged from 86.1% to 93.7%, 83.2% to 94.6%, and 85.2% to 94.4%, and the F1 scores ranged from 0.871 to 0.959, 0.881 to 0.967, 0.886 to 0.956. In the 4-fold cross-validation on reconstruction kernel, the accuracies ranged from 85.9% to 96.1%, 84.9% to 95.9%, 83.3% to 96.0%, and the F1 scores ranged from 0.893 to 0.963, 0.871 to 0.963, and 0.867 to 0.966.

4

Velocity [mm/s]

CT system

| | 0 | 10 | 20 | 30 | 40 | 50 | 60 |

CT-A, CT-B, CT-C, CT-D
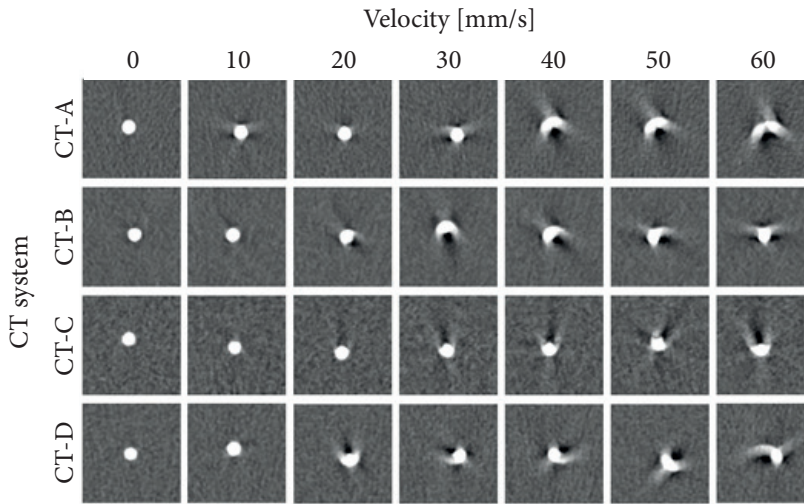
Figure 3 Representative images of the motion artifacts for all CT systems at the minimum and maximum velocity, clinical full and 80% reduced radiation dose and FBP and IR. Window center was 90 HU and window level was 750 HU. * for each CT system, the highest available level of IR was used. FBP=filtered back projection; IR= iterative reconstruction

Overall agreement between actual and predicted labels

The overall accuracy of the CNN classification for all four artificial plaques was 90.2±3.1%, 90.6±3.5%, and 90.1±3.2% for inception v3, ResNet101 and DenseNet201 CNN, respectively. The low-density plaque showed the highest accuracy of 93.3±1.6%, 92.4±2.8% and 92.7±2.8%, respectively; and F1 scores of 0.950±0.012, 0.947±0.020 and 0.945±0.019, respectively (*Table 1*). The AUCs were 0.982 (95% CI: 0.976-0.986), 0.981 (0.974-0.992) and 0.986 (0.982-0.994), respectively (*Table 2* and *Figure 4*).

Table 1 Classification accuracy and F1 scores of Inception v3, ResNet101 and DenseNet201 convolutional neural networks on calcified plaques with motion artifacts of four densities.

Plaque density	Inception v3		ResNet101		DenseNet201	
	Accuracy	*F1 score*	*Accuracy*	*F1 score*	*Accuracy*	*F1 score*
High	88.8±2.3%	0.917±0.024	90.2±2.8%	0.922±0.027	89.3±2.9%	0.919±0.023
Medium-1	88.0±3.0%	0.901±0.022	87.1±2.3%	0.896±0.020	87.7±2.9%	0.897±0.021
Medium-2	90.7±2.5%	0.939±0.024	92.9±2.9%	0.942±0.028	90.7±2.0%	0.937±0.018
Low	93.3±1.6%	0.950±0.012	92.4±2.8%	0.947±0.020	92.7±2.8%	0.945±0.019

Note. Variables are displayed as mean ± standard deviation.

Table 2 The area under receiver operating characteristic curves of convolutional neural network's classification on calcified plaques with motion artifacts.

Plaque density	Inception v3	ResNet101	DenseNet201
High	0.952 (0.939-0.964)	0.972 (0.962-0.980)	0.970 (0.960-0.978)
Medium-1	0.951 (0.939-0.962)	0.955 (0.943-0.965)	0.962 (0.951-0.972)
Medium-2	0.980 (0.970-0.989)	0.974 (0.969-0.981)	0.976 (0.970-0.982)
Low	0.982 (0.976-0.986)	0.981 (0.974-0.992)	0.986 (0.982-0.994)

Note. The data is expressed as area under the curve (95% confidence interval)

4
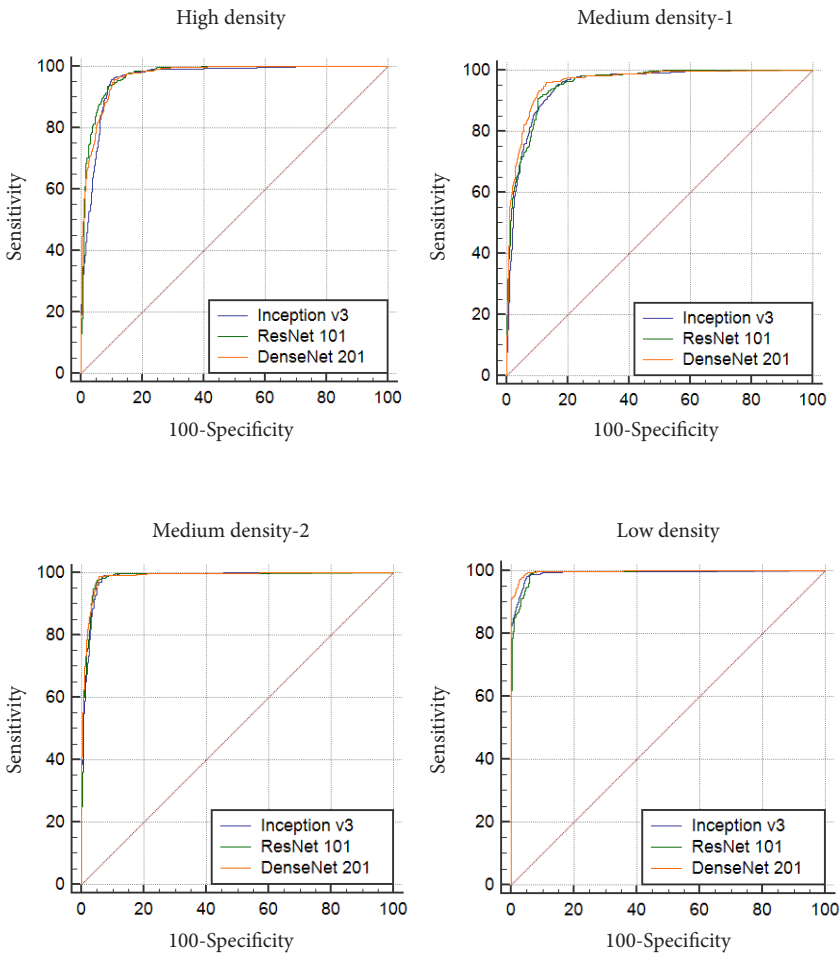
Figure 4 Receiver operating characteristic curves of convolutional neural network's classification on calcified plaques with motion artifacts.

The medium-density-1 plaque showed the lowest accuracy of 88.0±3.0%, 87.1±2.3% and 87.7±2.9%, respectively; F1 scores of 0.901±0.022, 0.896±0.020 and 0.897±0.021, respectively; and AUCs of 0.951 (0.970-0.962), 0.955 (0.943-0.981) and 0.962 (0.951-0.972), respectively. An ensemble of three CNN models by a multivariate linear regression model slightly increased the AUC to 0.990 (0.982-0.998).

With respect to different CT vendors, the accuracy was similar, and ranged from 89.8±2.7% to 91.8±0.8%, 88.0±5.3% to 92.2±2.3%, and 89.3±5.6% to 92.0±1.8% for inception v3, ResNet101 and DenseNet201 CNN, respectively (*Table 3*). Regarding the velocity, the plaques at rest showed relatively lower accuracies of 80.3±1.3%, 85.4±1.4%, and 89.2±1.2%, respectively (*Figure 5*). When the velocity increased to 60 mm/s, the accuracy increased to 93.8±1.3%, 95.1±1.6%, and 96.1±1.5%, respectively.

Table 3 Classification accuracy of Inception v3, ResNet101 and DenseNet201 convolutional neural network on calcified plaques with motion artifacts on four CT systems.

	Inception v3	ResNet101	DenseNet201
CT-A	90.2±3.1%	92.2±2.3%	92.0±1.8%
CT-B	89.8±2.7%	88.0±5.3%	89.3±5.6%
CT-C	91.0±2.8%	90.9±2.6%	90.7±2.4%
CT-D	91.8±0.8%	91.2±3.6%	91.1±2.6%

Note. Variables are displayed as mean ± standard deviation.
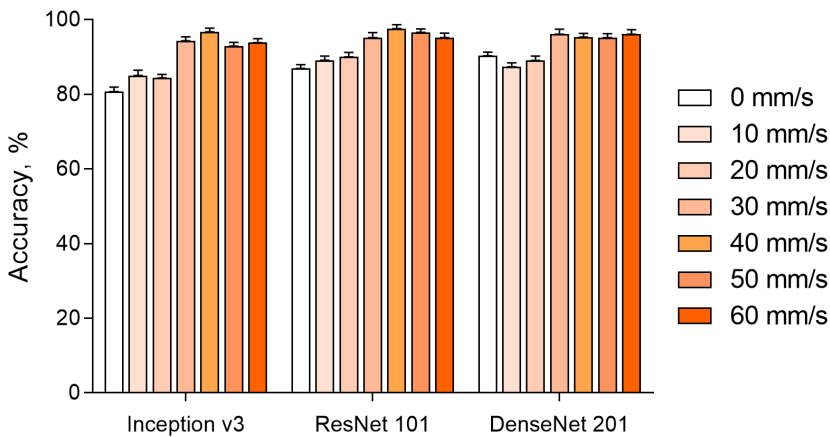
Figure 5 Classification accuracy of Inception v3, ResNet101, and DenseNet201 convolutional neural network on calcified plaques with motion artifacts in the velocity from 0 to 60 mm/s.

4

Univariate association between influencing factors and CNN's classification

Lower plaque density and increasing velocity were significantly associated with higher classification accuracy in all three CNN architectures (all P<0.001). However, other factors (CT system, radiation dose and reconstruction method) had no significant influence on the accuracy of CNN classification (all P>0.05) (*Table 4*).

Table 4 Spearman's correlation coefficients (rho) for the univariate association between influencing factors and convolutional neural network's classification on calcified plaques with motion artifacts.

	Inception v3		ResNet101		DenseNet201	
	rho (95% CI)	p-value	rho (95% CI)	p-value	rho (95% CI)	p-value
Density	0.139 (0.083, 0.195)	<0.001	0.199 (0.061, 0.337)	<0.001	0.194 (0.074, 0.314)	<0.001
CT vendor	0.031 (-0.017, 0.079)	0.249	-0.049 (-0.151, 0.053)	0.210	-0.102 (-0.192, -0.012)	0.107
Velocity	0.191 (0.105, 0.277)	<0.001	0.169 (0.029, 0.309)	<0.001	0.163 (0.024, 0.302)	<0.001
Dose	-0.028 (-0.081, 0.025)	0.518	0.028 (-0.052, 0.108)	0.477	0.046 (-0.010, 0.102)	0.239
Reconstruction	0.019 (-0.031, 0.069)	0.312	0.118 (0.030, 0.206)	0.230	0.134 (0.023, 0.245)	0.222

Note. High, medium-1, medium-2, and low-density plaque were coded as 1 to 4, respectively, four CT systems (CT-A to CT-D) as 1-4; velocities from 0 to 60 mm/s coded as 0-6; dose level 40%, 80% and full dose coded as 1-3; recon method FBP, IR1 to IR3 coded as 1 to 4.
FBP=filtered back projection; IR= iterative reconstruction

Multivariate analysis for factors associated with CNN's classification

In the multivariate analysis, a higher density and an increasing velocity were significantly associated with a higher classification accuracy for all three CNN architectures (all P<0.001). The classification accuracy for all three CNN architectures was not affected by CT system, radiation dose or reconstruction method (all P>0.05) (*Table 5*).

Table 5 Multivariate analysis for the influencing factors associated with CNN's classification on calcified plaques with motion artifacts.

	Inception v3		ResNet101		DenseNet201	
	coefficient	p-value	coefficient	p-value	coefficient	p-value
Density	0.033	<0.001	0.024	<0.001	0.319	<0.001
CT vendor	0.012	0.147	-0.025	0.091	-0.038	0.102
Velocity	0.027	<0.001	0.017	<0.001	0.015	<0.001
Dose	-0.009	0.601	-0.011	0.159	0.002	0.779
Reconstruction	0.009	0.126	0.010	0.112	0.012	0.099

Note. High, medium-1, medium-2, and low-density plaque were coded as 1 to 4, respectively, four CT systems (CT-A to CT-D) as 1-4; velocities from 0 to 60 mm/s coded as 0-6; dose level 40%, 80% and full dose coded as 1-3; recon method FBP, IR1 to IR3 coded as 1 to 4.
FBP=filtered back projection; IR= iterative reconstruction

DISCUSSION

In this experimental study, we applied three widely-used deep CNN architectures in medical image analysis to classify the CT images of calcified coronary plaques with motion artifacts into the correct category with a high accuracy, regardless of different CT vendors, velocities, radiation doses, and image reconstruction algorithms.[25–27] The overall classification accuracy of these CNNs reached 90%.

ECG-triggered CT and non-ECG-triggered CT are vulnerable to a great variety of artifacts. The most common cause of artifacts is motion.[28] Typical motion artifacts are blurring, ghosting, or windmills, which influence accurate CAC classification.[29] The severity of the motion artifacts depends not only on the heart rate and temporal resolution of the CT scanner but also on the specific coronary artery. The mean velocity of the right coronary artery is significantly higher than that of the left anterior descending and circumflex coronary artery.[30] In our study, all the CNN architectures resulted in higher classification accuracy at higher velocities. In contrast to the plaques at rest, which showed no motion artifacts, the plaques showed spatially more dispersed motion artifacts caused by increasing velocities (*Figure 2*). The magnitude of the artifact means that more image features that can be recognized and used for classification by a CNN.

Motion artifacts appear when the temporal resolution is insufficient to warrant data acquisition during the time the coronary arteries exhibit the least motion. The temporal resolution can be improved by a shorter gantry rotation time, or dedicated acquisition and reconstruction protocols.[31] It has been demonstrated

that plaque classification also depends on the density and size, where relatively small and soft calcified plaques may remain undetected.[32] Van der Werf et al. found that at an increased heart rate, the Agatston scores of low-density calcified plaques were similar to the reference scores but that the Agatston scores of medium- and high-density calcified plaques were increased by up to 50%.[11] In our study, all CNN architectures gained the highest classification accuracy for plaques with low density.

Another difficulty that a CNN encounters during plaque motion artifact classification is increased noise in low-dose CT scans. Currently, IR algorithms can be used for CT image reconstruction to reduce the noise in low-dose CT scans.[33,34] In a phantom study, increased IR levels resulted in decreased CAC scores.[19] In clinical practice for CAC scoring, a soft kernel is used in which the noise is less prominent. However, in non-contrast and non-ECG-triggered CT scans performed for other diagnostic purposes, generally, sharper kernels are used. Recently, it was found that if a CNN is trained on both soft and sharp kernels, the accuracy of the CNN in CAC scoring is similar to that of soft kernel CT scans.[35] In our study, both reconstruction methods did not influence the accuracy of all three CNN architectures. Furthermore, a lower radiation dose also did not affect the CNN's classification.

So far, for CAC classification in non-contrast and ECG-triggered CT scans, different machine-learning-based techniques have been studied to improve the diagnostic management in everyday practice. Most of these methods focused on calcium detection and coronary artery calcium scoring.[35–37] In contrast, the method presented in the current study was based on training CNNs on motion artifacts, which are one of the main limitations in accurate CAC score measurement. Motion artifact recognition methods based on motion correction algorithms have shown their value in the image improvement of coronary CT angiography (CCTA).[38] Furthermore, a CNN was used to estimate the artifact motion vectors from CCTA images and the method improved the quality and showed potential to be applied in clinical practice.[39] In our study, CNNs were solely used for motion artifact recognition from non-contrast and ECG-triggered CT scans to enhance the assessment of CAC scores in the future. Since phantoms were used for the current study, the increase of the motion artifacts as heart rate increased was known. This led to enhanced accuracy of CNNs since motion artifact features increased as the heart rate increased. Although this is of course not feasible in daily practice, our results indicated that a CNN may play an important role in clinical coronary calcium classification in the near future.

The main limitation of this study is the linear movement of artificial arteries, their relatively high diameters, and the regular shape of calcifications, which do not reflect the movement and shape of coronary arteries in vivo. Additionally, a variety of body habitus were not also included.

In this experimental study, the CNN achieved a high accuracy of 90% at classifying motion-contaminated images into the actual category, regardless of different influential technical factors. This study validated the first step towards patient-specific motion artifacts recognition which could be used to correct CAC scores in the future. Interestingly, the classification accuracy increased at higher velocities because the magnitude of the artifact had more image features that may be recognized and used for classification by a CNN, which inspires us to optimize the calcium score correction method in cases of heavily motion-contaminated images.

REFERENCES

1. Montalescot G, Sechtem U, Achenbach S, et al. 2013 ESC guidelines on the management of stable coronary artery disease: the Task Force on the management of stable coronary artery disease of the European Society of Cardiology. Eur Heart J. 2013;34(38):2949-3003. doi:10.1093/eurheartj/eht296

2. Rana JS, Khan SS, Lloyd-Jones DM, Sidney S. Changes in Mortality in Top 10 Causes of Death from 2011 to 2018. J Gen Intern Med. 2021;36(8):2517-2518. doi:10.1007/s11606-020-06070-z

3. Budoff MJ, Achenbach S, Blumenthal RS, et al. Assessment of coronary artery disease by cardiac computed tomography: A scientific statement from the American Heart Association Committee on Cardiovascular Imaging and Intervention, Council on Cardiovascular Radiology and Intervention, and Committee on C. Circulation. 2006;114(16):1761-1791. doi:10.1161/CIRCULATIONAHA.106.178458

4. Budoff MJ, Shaw LJ, Liu ST, et al. Long-term prognosis associated with coronary calcification: observations from a registry of 25,253 patients. J Am Coll Cardiol. 2007;49(18):1860-1870. doi:10.1016/j.jacc.2006.10.079

5. Yeboah J, McClelland RL, Polonsky TS, et al. Comparison of novel risk markers for improvement in cardiovascular risk assessment in intermediate-risk individuals. JAMA. 2012;308(8):788-795. doi:10.1001/jama.2012.9624

6. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. J Am Coll Cardiol. 1990;15(4):827-832. doi:10.1016/0735-1097(90)90282-T

7. Berrington de González A, Mahesh M, Kim K-P, et al. Projected cancer risks from computed tomographic scans performed in the United States in 2007. Arch Intern Med. 2009;169(22):2071-2077. doi:10.1001/archinternmed.2009.440

8. Brenner DJ, Hall EJ. Computed Tomography — An Increasing Source of Radiation Exposure. N Engl J Med. 2007;357(22):2277-2284. doi:10.1056/nejmra072149

9. Xie X, Zhao Y, de Bock GH, et al. Validation and prognosis of coronary artery calcium scoring in nontriggered thoracic computed tomography: systematic review and meta-analysis. Circ Cardiovasc Imaging. 2013;6(4):514-521. doi:10.1161/CIRCIMAGING.113.000092

10. Hecht HS, Cronin P, Blaha MJ, et al. 2016 SCCT/STR guidelines for coronary artery calcium scoring of noncontrast noncardiac chest CT scans: A report of the Society of Cardiovascular Computed Tomography and Society of Thoracic Radiology. J Cardiovasc Comput Tomogr. 2017;11(1):74-84. doi:10.1016/j.jcct.2016.11.003

11. van der Werf NR, Willemink MJ, Willems TP, Vliegenthart R, Greuter MJW, Leiner T. Influence of heart rate on coronary calcium scores: a multi-manufacturer phantom study. Int J Cardiovasc Imaging. 2017;34(6):959-966. doi:10.1007/s10554-017-1293-x

12. Steigner ML, Otero HJ, Cai T, et al. Narrowing the phase window width in prospectively ECG-gated single heart beat 320-detector row coronary CT angiography. Int J Cardiovasc Imaging. 2009;25(1):85-90. doi:10.1007/s10554-008-9347-8

13. Husmann L, Leschka S, Desbiolles L, et al. Coronary artery motion and cardiac phases: dependency on heart rate -- implications for CT image reconstruction. Radiology. 2007;245(2):567-576. doi:10.1148/radiol.2451061791

4

14. Johnson PT, Eng J, Pannu HK, Fishman EK. 64-MDCT angiography of the coronary arteries: nationwide survey of patient preparation practice. AJR Am J Roentgenol. 2008;190(3):743-747. doi:10.2214/AJR.07.2620

15. Waltz J, Kocher M, Kahn J, Dirr M, Burt JR. The Future of Concurrent Automated Coronary Artery Calcium Scoring on Screening Low-Dose Computed Tomography. Cureus. 2020;12(6):e8574. doi:10.7759/cureus.8574

16. Šprem J, de Vos BD, Lessmann N, de Jong PA, Viergever MA, Išgum I. Impact of automatically detected motion artifacts on coronary calcium scoring in chest computed tomography. J Med imaging (Bellingham, Wash). 2018;5(4):44007. doi:10.1117/1.JMI.5.4.044007

17. Eng D, Chute C, Khandwala N, et al. Automated coronary calcium scoring using deep learning with multicenter external validation. npj Digit Med. 2021;4(1):88. doi:10.1038/s41746-021-00460-1

18. Zhang Y, van der Werf NR, Jiang B, van Hamersvelt R, Greuter MJW, Xie X. Motion-corrected coronary calcium scores by a convolutional neural network: a robotic simulating study. Eur Radiol. 2020;30(2):1285-1294. doi:10.1007/s00330-019-06447-7

19. van der Werf NR, Willemink MJ, Willems TP, Greuter MJW, Leiner T. Influence of dose reduction and iterative reconstruction on CT calcium scores: a multi-manufacturer dynamic phantom study. Int J Cardiovasc Imaging. 2017;33(6):899-914. doi:10.1007/s10554-017-1061-y

20. Blagus R, Lusa L. Joint use of over- and under-sampling techniques and cross-validation for the development and assessment of prediction models. BMC Bioinformatics. 2015;16(1):363. doi:10.1186/s12859-015-0784-9

21. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the Inception Architecture for Computer Vision. Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit. 2016;2016-Decem:2818-2826. doi:10.1109/CVPR.2016.308

22. Zaeemzadeh A, Rahnavard N, Shah M. Norm-Preservation: Why Residual Networks Can Become Extremely Deep? IEEE Trans Pattern Anal Mach Intell. 2020;43(11):3980-3990. doi:10.1109/tpami.2020.2990339

23. Huang G, Liu Z, Pleiss G, Van Der Maaten L, Weinberger K. Convolutional Networks with Dense Connectivity. IEEE Trans Pattern Anal Mach Intell. Published online 2019:1-1. doi:10.1109/tpami.2019.2918284

24. Zhang YP, Heuvelmans MA, Zhang H, Oudkerk M, Zhang GX, Xie XQ. Changes in quantitative CT image features of ground-glass nodules in differentiating invasive pulmonary adenocarcinoma from benign and in situ lesions: histopathological comparisons. Clin Radiol. 2018;73(5):504.e9-504.e16. doi:10.1016/j.crad.2017.12.011

25. He Y, Guo J, Ding X, et al. Convolutional neural network to predict the local recurrence of giant cell tumor of bone after curettage based on pre-surgery magnetic resonance images. Eur Radiol. 2019;29(10):5441-5451. doi:10.1007/s00330-019-06082-2

26. Jiang B, Zhang Y, Zhang L, H de Bock G, Vliegenthart R, Xie X. Human-recognizable CT image features of subsolid lung nodules associated with diagnosis and classification by convolutional neural networks. Eur Radiol. 2021;31(10):7303-7315. doi:10.1007/s00330-021-07901-1

27. Rueckel J, Huemmer C, Fieselmann A, et al. Pneumothorax detection in chest radiographs: optimizing artificial intelligence system for accuracy and confounding bias reduction using in-image annotations in algorithm training. Eur Radiol. 2021;31(10):7888-7900. doi:10.1007/s00330-021-07833-w

28. Greuter MJW, Groen JM, Nicolai LJ, Dijkstra H, Oudkerk M. A model for quantitative correction of coronary calcium scores on multidetector, dual source, and electron beam computed tomography for influences of linear motion, calcification density, and temporal resolution: A cardiac phantom study. Med Phys. 2009;36(11):5079-5088. doi:10.1118/1.3213536

29. Padgett J, Biancardi AM, Henschke CI, Yankelevitz D, Reeves AP. Local noise estimation in low-dose chest CT images. Int J Comput Assist Radiol Surg. 2014;9(2):221-229. doi:10.1007/s11548-013-0930-7

30. Achenbach S, Ropers D, Holle J, Muschiol G, Daniel WG, Moshage W. In-plane coronary arterial motion velocity: measurement with electron-beam CT. Radiology. 2000;216(2):457-463. doi:10.1148/radiology.216.2.r00au19457

31. Aghayev A, Murphy DJ, Keraliya AR, Steigner ML. Recent developments in the use of computed tomography scanners in coronary artery imaging. Expert Rev Med Devices. 2016;13(6):545-553. doi:10.1080/17434440.2016.1184968

32. Šprem J, De Vos BD, Lessmann N, et al. Coronary calcium scoring with partial volume correction in anthropomorphic thorax phantom and screening chest CT images. PLoS One. 2018;13(12):1-20. doi:10.1371/journal.pone.0209318

33. Willemink MJ, de Jong P a, Leiner T, et al. Iterative reconstruction techniques for computed tomography Part 1: Technical principles. Eur Radiol. Published online January 2013. doi:10.1007/s00330-012-2765-y

34. Willemink MJ, Leiner T, De Jong PA, et al. Iterative reconstruction techniques for computed tomography part 2: Initial results in dose reduction and image quality. Eur Radiol. 2013;23(6):1632-1642. doi:10.1007/s00330-012-2764-z

35. Lessmann N, van Ginneken B, Zreik M, et al. Automatic Calcium Scoring in Low-Dose Chest CT Using Deep Neural Networks With Dilated Convolutions. IEEE Trans Med Imaging. 2018;37(2):615—625. doi:10.1109/tmi.2017.2769839

36. Isgum I, Prokop M, Niemeijer M, Viergever MA, van Ginneken B. Automatic coronary calcium scoring in low-dose chest computed tomography. IEEE Trans Med Imaging. 2012;31(12):2322-2334. doi:10.1109/TMI.2012.2216889

37. Brunner G, Chittajallu DR, Kurkure U, Kakadiaris IA. Toward the automatic detection of coronary artery calcification in non-contrast computed tomography data. Int J Cardiovasc Imaging. 2010;26(7):829-838. doi:10.1007/s10554-010-9608-1

38. Rohkohl C, Bruder H, Stierstorfer K, Flohr T. Improving best-phase image quality in cardiac CT by motion correction with MAM optimization. Med Phys. 2013;40(3):31901. doi:10.1118/1.4789486

39. Lossau Née Elss T, Nickisch H, Wissel T, et al. Motion estimation and correction in cardiac CT angiography images using convolutional neural networks. Comput Med imaging Graph Off J Comput Med Imaging Soc. 2019;76:101640. doi:10.1016/j.compmedimag.2019.06.001

Online material is available via:

PART II

Acquisition and reconstruction parameters
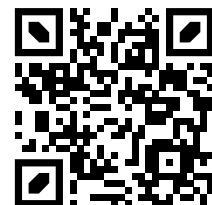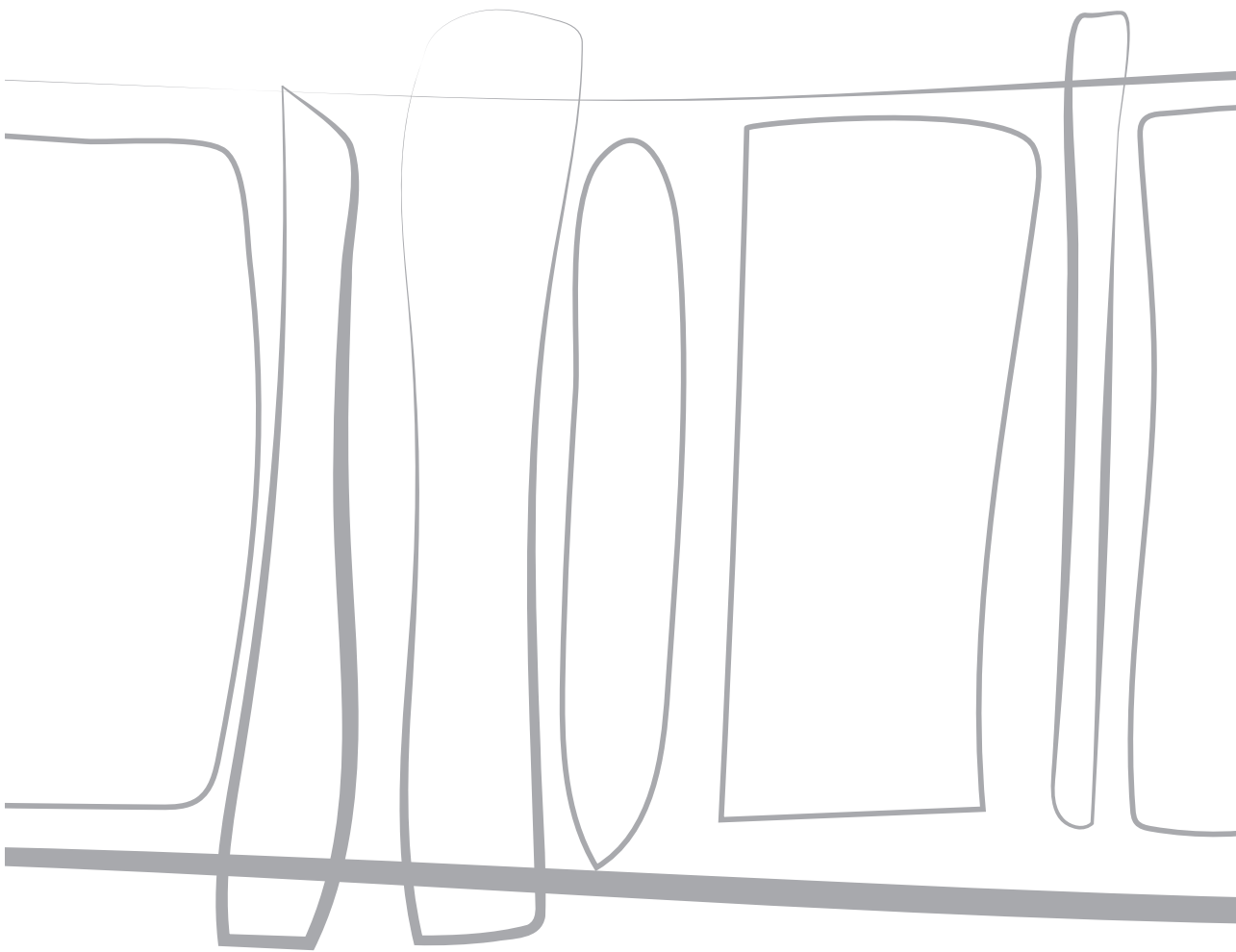
CHAPTER 5

The impact of dose reduction on the quantification of coronary artery calcifications and risk categorization: a systematic review

Marleen Vonder, MSc

Niels R. van der Werf, MSc

Tim Leiner, MD PhD

Marcel J. W. Greuter, PhD

Dominik Fleischmann, MD

Rozemarijn Vliegenthart, MD PhD

Matthijs Oudkerk, MD PhD

Martin J. Willemink, MD PhD

5

ABSTRACT

Multiple dose reduction techniques have been introduced for coronary artery calcium (CAC) computed tomography (CT), but few have emerged into clinical practice while an increasing number of patients undergo CAC scanning. We sought to determine to what extend the radiation dose in CAC CT can be safely reduced without a significant impact on cardiovascular disease (CVD) risk stratification. A systematic database-review of articles published from 2002 until February 2018 was performed in Pubmed, WebOfScience, and Embase. Eligible studies reported radiation dose reduction for CAC CT, calcium scores and/or risk stratification for phantom or patient studies. Twenty-eight studies were included, under which 17 patient studies, 10 phantom/ex-vivo studies, and 1 study evaluated both phantom and patients. Dose was reduced with tube voltage reduction and tube current reduction with and without iterative reconstruction (IR), and tin-filter spectral shaping. The different dose reduction techniques resulted in varying final radiation doses and had varying impact on CAC scores and CVD risk stratification. In 78% of the studies the radiation dose was reduced by ≥50% ranging from (CTDIvol) 0.6 to 5.5 mGy, leading to reclassification rates ranging between 3% and 21%, depending on the acquisition technique. Specific dose reduced protocols, including either tube current reduction and IR or spectral shaping with tin filtration, that showed low reclassification rates may potentially be used in CAC scanning and in future population-based screening for CVD risk stratification.

INTRODUCTION

The amount of coronary artery calcification (CAC) expressed in Agatston scores has shown to be strongly associated with risk of cardiovascular disease (CVD).[1] CAC assessment with computed tomography (CT) has substantially gained interest, resulting in increased numbers of CAC CT examinations. Ongoing and future research will evaluate the feasibility of population based screening for CVD by determining the amount of CAC on CT images.[2,3] If positive, millions of people worldwide will be eligible for screening, leading to an even further increase of individuals exposed to ionizing radiation. Moreover, repetitive screening or follow-up scans might be required, adding to the cumulative radiation dose.[4]

Therefore, continual efforts have been made to reduce the radiation dose in cardiac CT, resulting in the introduction of multiple dose reduction techniques. While radiation exposure has been dramatically reduced for coronary CT angiography in the last decade,[5–11] this has not been the case for CAC CT. In fact, clinically used acquisition protocols are nowadays still similar to the methods used in the 1990s on electron beam tomography.[12] The impact of the available dose reduction techniques were examined in multiple small-sized phantom and/or patient studies on a variety of CT scanners from different vendors. Although many studies evaluated these techniques, there is no clear overview and guidelines regarding their impact and there is only limited implementation of these techniques into clinical practice for CAC imaging.[13]

The aim of the current study was therefore to systematically review the available dose reduction techniques for CAC CT and to determine to what extend the radiation dose can be safely reduced without significantly impacting the CAC score and/or CVD risk stratification.

METHODS

Search strategy

A systematic literature search was performed in February 2018 for studies assessing dose reduction in CAC CT using the Pubmed, Embase and Web of Science databases. The following search strategy was used in Pubmed: ((((((coronar*)) AND (calcium OR calcification*)) AND (radiation OR dose) AND (reduc* OR low*)). Additionally, Embase and Web of Science were searched using adjusted search strategy to fit the search matrix of the database source.

Inclusion and exclusion criteria

Inclusion criteria were published studies less than 15 years old; single or multicenter; either included phantom, ex vivo and/or patient data; included non-contrast electrocardiography (ECG) triggered cardiac CT; reported quantification of radiation dose reduction, CAC scores (e.g. Agatston score, volume score, mass score), and/or CVD risk stratification. Exclusion criteria were non-English written full text articles; abstracts without full text; editorials, reviews, case reports, letters and guidelines. Studies were excluded that did not report the outcome of interest or if the outcome of interest could not be calculated from the results. We also excluded studies with protocols for which the primary indication was not CAC quantification (e.g. lung CT scans and CT angiography).

Study selection and data extraction

Studies for the systematic review were selected using the PRISMA flow diagram, see *Figure 1*.[14] The screening of title and abstract of each paper was independently performed by two reviewers (MV, NvW). Subsequently, both reviewers independently evaluated the full-text of each article for eligibility based on the in- and exclusion criteria. In case of disagreement, eligibility of the article was discussed between the two reviewers to obtain consensus.
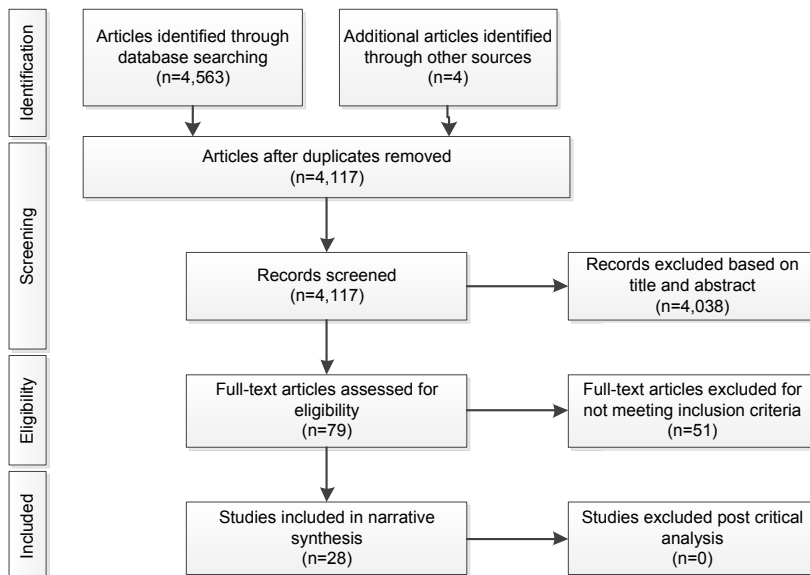
Figure 1 Flowchart of the systematic review of dose reduction in CAC scanning.

Study characteristics and data extraction of selected articles was performed independently by two authors (MV, NvW) according to a predefined protocol. The following study characteristics were collected: author, year of publication, study type, radiation dose reduction technique, scanner type and vendor, acquisition and reconstruction and radiation dose parameters, CAC scores, and percentage of dose reduction, number of included patients, and impact on CVD risk stratification. The final retrieved data were reviewed by one author (MV).

Analysis of data

Data were grouped per radiation dose technique, and per IR algorithm that was applied. The key parameter setting leading to the radiation dose reduction was extracted for the full dose and low dose protocols. Remaining acquisition and reconstruction parameters were logged. If a study investigated multiple low dose protocols, only the results of the protocols leading to no significant different Agatston scores or showing high agreement for risk categorization with the full dose protocol were included in the tables. The impact on Agatston scores was extracted. If available, volume and mass score were also extracted. Percentage differences between the radiation dose for the full and reduced dose scans and impact on CVD risk stratification were extracted or calculated.

RESULTS

Characteristics of included studies

In total 28 studies were included, of which 17 were patient studies, 10 were phantom/ex vivo studies, and 1 study included both phantom and patients. The used dose reduction techniques were tube voltage reduction, tube voltage reduction with iterative reconstruction (IR), tube current reduction, tube current reduction with IR, and spectral shaping with tin-filter. All studies used multi-detector or dual-source CT and used either retrospectively or prospectively ECG-gated acquisition in patients. All studies used a tube voltage of 120 kVp (except for the study by Mahnken et al.[15]), either a fixed or adaptive tube current and FBP as the reference full dose protocol. An overview of the $CTDI_{vol}$ of the full and reduced dose protocols per dose reduction technique is shown in *Figure 2*.

5
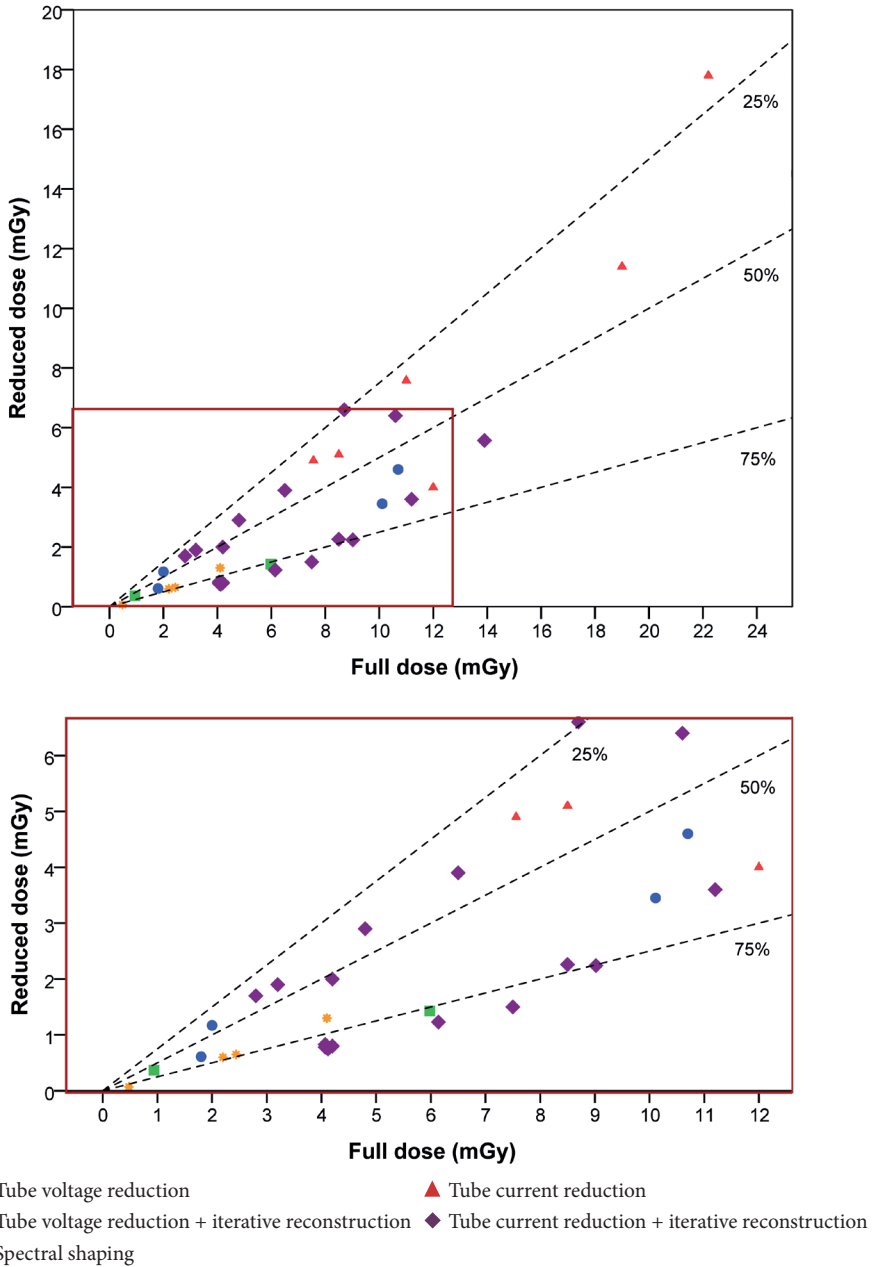
Figure 2 Radiation dose of the full and reduced dose protocols grouped per dose reduction technique, with reference lines representing 25%, 50% and 75% dose reduction.

Tube voltage reduction

In total, four studies examined the impact of tube voltage reduction on CAC quantification, see *Table 1* and *Figure 3*. A lower (peak) tube voltage results in an X-ray spectrum with overall lower-energy photons, resulting in a lower radiation dose. However, this results in increased image noise, which may affect CAC quantification. Besides increased image noise, a reduction in tube voltage also results in increased attenuation of most materials, including calcium.[16] Tube voltage reduction may therefore require adaptation of the Hounsfield unit (HU) threshold for calcium measurements.[17] At 120 kVp, the threshold is set at 130 HU and in general a higher threshold is used for lower tube voltages. Studies by Thomas et al. and Marwan et al. showed that despite an adapted HU threshold for 80 and 100 kVp (187 and 147 HU, respectively), reduced tube voltage from 120 kVp to 100 or 80 kVp led to an overestimation of the Agatston, volume and mass score in a phantom and 71 patients.[18,19] The number of patients (n=69) with zero-scores was similar, however 5 out of 71 patients with a positive-score (7%) were reclassified to higher risk categories.[19] Contrary, another study by Jakobs et al. showed that the mass score did not differ between 120 and 80 kVp protocols in 34 patients with a positive score.[20] Whereas in a patient study (n=103) of Gräni et al., the Agatston score was underestimated for 70 and 80 kVp protocols with adapted HU threshold compared to a 120 kVp protocol. Subsequently, these patients were mostly reclassified into lower risk categories (6.1% reclassifications for 70 kVp and 2.8% for 80 kVp, respectively).[21] Radiation dose ranged from 1.8 to 10.7 mGy for 120 kVp and ranged from 0.6 to 4.6 mGy for tube voltage reduced protocols.[18–21] These dose reductions resulted in reclassification rates ranging from 2.8% to 7%.

Tube voltage reduction and iterative reconstruction

The issue of increased image noise at lower tube voltages could be solved by combining a lower tube voltage with IR.[22] IR is a noise-reducing technique that allows for radiation dose reduction while maintaining low image noise and diagnostic image quality.[23] Most IR algorithms have different noise reduction levels, in general lower IR levels correspond to higher image noise levels. In total, 2 phantom studies were included that examined the impact of tube voltage reduction in combination with IR, see *Table 2* and *Figure 3*. These studies conducted by Blobel et al.[24] and Vonder et al.[25] showed that tube voltage reductions to 80, 90, or 100 kVp did not significantly affect Agatston and CAC volume scores with the use of IR.[24,25] However, risk reclassification was not evaluated since both studies

5

Table 1 – Characteristics of included studies using tube voltage reduction

Study	Phantom/ patients	Reference full dose protocol			
		Tube voltage	HU Threshold	CTDI$_{vol}$/ Eff dose	CAC score
Jakobs et al.[20] (2003)	Patients (n=46)	120 kVp	130 HU	10.1 mGy	Mass: 86 mg
Thomas et al.[18] (2006)	Phantom	120 kVp	130 HU	10.7 mGy	Agatston: 649.0±18.1*
Marwan et al.[19] (2013)	Patients (n=71)	120 kVp	130 HU	2.0 mGy 0.3 mSv	Agatston: 105±245*
Gräni et al.[21] (2018)	Patients (n=103)	120 kVp	130 HU	1.8 mGy 0.6 mSv	Agatston: 212 (25-901)**

Study	Phantom/ patients	Reduced dose protocol			
		Tube voltage	HU Threshold	CTDI$_{vol}$/ Eff dose	CAC score
Jakobs et al.[20] (2003)	Patients (n=46)	80 kVp	130 HU	3.5 mGy	Mass: 84 mg
Thomas et al.[18] (2006)	Phantom	80 kVp	187 HU	4.6 mGy	Agatston: 671.3±19.5*
Marwan et al.[19] (2013)	Patients (n=71)	100 kVp	147 HU	1.2 mGy 0.2 mSv	Agatston: 116±261*
Gräni et al.[21] (2018)	Patients (n=103)	70 kVp 80 kVp	Adapted HU threshold	0.6 mGy 0.1, 0.2 mSv	Agatston mean difference: 80 kVp=-31(-5.2%) 70 kVp=-103(-18.4)

Study	Phantom/ patients	Impact		
		CAC score	Risk classification	Dose reduction
Jakobs et al.[20] (2003)	Patients (n=46)	Mass: similar Agatston: n.s.	n.s.	65%
Thomas et al.[18] (2006)	Phantom	Agatston: Significant increase	n.s.	57%
Marwan et al.[19] (2013)	Patients (n=71)	Agatston: Significant increase	7% reclassified	33%
Gräni et al.[21] (2018)	Patients (n=103)	Agatston: Significant decrease	6.1% and 2.8% reclassified for 70 and 80 kVp	Up to 80%
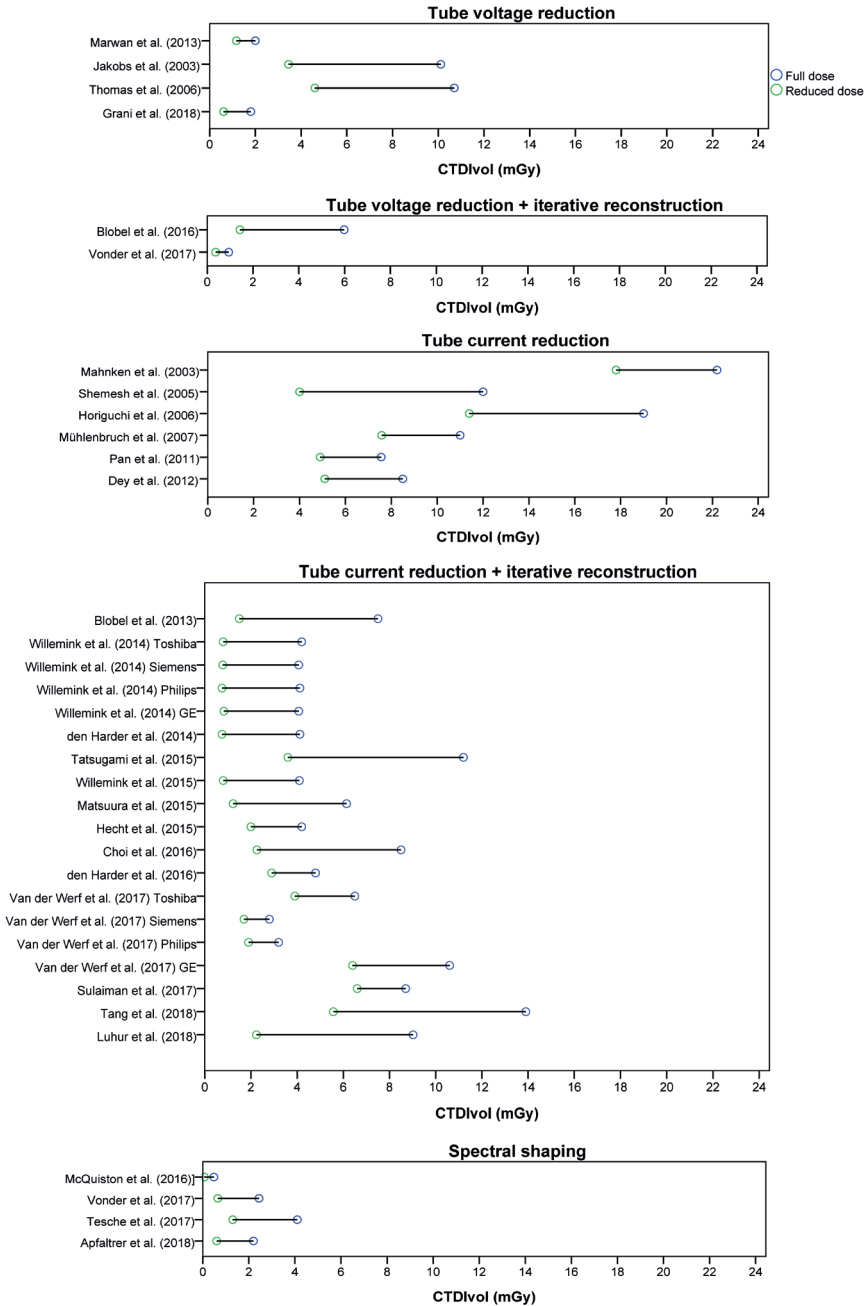
Figure 3 Mean radiation dose for the full and reduced dose protocol per study grouped per dose reduction technique.

5

Table 2 – Characteristics of included studies using tube voltage reduction and iterative reconstruction

Study	Phantom/ patients	Reference full dose protocol				
		Tube voltage	Recon.	HU threshold	CTDI$_{vol}$/ Eff dose	CAC score
Blobel et al.[24] (2016)	Phantom	120 kVp	FBP	130 HU	6.0 mGy	Agatston: 698.5* Volume: 586.8 mm³*
Vonder et al.[25] (2017)	Phantom	120 kVp	FBP	130 HU	0.9 mGy	Agatston: 33.6 (30.4-38.4)**

Study	Phantom/ patients	Reduced dose protocol				
		Tube voltage	Recon.	HU threshold	CTDI$_{vol}$/ Eff dose	CAC score
Blobel et al.[24] (2016)	Phantom	100 kVp 80 kVp	AIDR-3D (Thosiba)	Adapted HU threshold	1.6 mGy 1.5 mGy	Agatston: (100, 80 kVp) 697.6, 698.5* Volume: 583.6, 585.6 mm³*
Vonder et al.[25] (2017)	Phantom	100 kVp 90 kVp	ADMIRE level 1, 3 (Siemens)	130 HU	0.4 mGy	Agatston: n.s.

Study	Phantom/ patients	Impact		
		CAC score	Risk classification	Dose reduction
Blobel et al.[24] (2016)	Phantom	Similar Agatston and volume scores	n.s.	76%
Vonder et al.[25] (2017)	Phantom	Similar Agatston scores	n.s.	60%

*mean; ** median (IQR); Recon.: Reconstruction algorithm; FBP: filtered back projection; CTDIvol: Computed Tomography Dose Index volume; Eff dose: effective radiation dose; AIDR-3D: Adaptive iterative dose reduction 3D; ADMIRE: Advanced model based iterative reconstruction; n.s.: not specified

concerned phantom studies. Of note, these studies used different IR algorithms on different CT systems. Besides, in one study the HU threshold was adapted [24] and in the other study the variability of the scores was higher for the tube voltage reduced protocol with IR compared to the reference 120 kVp protocol.[25] Radiation dose ranged from 0.9 to 6.0 mGy for 120 kVp and ranged from 0.4 to 1.4 mGy for tube voltage reduced protocols.[24,25]

Tube current reduction

Another approach to reduce the radiation dose in CT imaging is to lower the current of the X-ray tube. By reducing the tube current, less X-ray photons will be emitted. Therefore, similar to tube voltage reduction, a lower tube current will result in increased image noise. However, since the energy of the X-ray photons is not affected, a reduced tube current will not affect attenuation of calcium and thus, there is no need for adjusted calcium quantification thresholds. In four studies with a total of 253 patients, the tube current was optimized and/or a reduced (fixed) tube current was used, see *Table 3* and *Figure 3*. Different optimization techniques for tube current reduction were used: body mass index (BMI), body-weight, or attenuation based. For all studies, the image noise increased for the lower tube current protocols, but overall image quality was higher and more constant across different patients' body sizes.[15,26–28] Notably, the studies mention different maximum noise levels ranging from 18 to 30 HU for the dose reduced protocols for being regarded as an acceptable noise level. Studies reported no significant differences or a high agreement for Agatston scores for the optimized and reduced protocols compared to the full dose protocol. However, none of these studies reported on the effect of tube current optimization and reduction on risk stratification. Radiation dose of the full dose protocols ranged from 8.5 to 22.2 mGy and reduced radiation doses ranged from 4.0 to 17.8 mGy.

Tube current reduction and iterative reconstruction

Similar to tube voltage reduction, the increased noise issues with lower tube current can potentially be solved by applying IR.[29] Since only noise is affected by tube current reductions and mean HU-values are not affected, IR has the potential to allow for radiation dose reduction without affecting CAC scores. The majority of studies (n=13) included in this systematic review examined the impact of tube current reduction combined with IR, see *Table 4* and *Figure 3*. IR algorithms are vendor and scanner type specific. Therefore, the results of the studies are grouped per IR algorithm that was applied.

iDose4 & Iterative Model based Reconstuction (IMR)

The fourth version of iDose (iDose4, Philips Healthcare, Best, The Netherlands) is based on an algorithm iterating first in the raw data domain and is based on maximum likelihood denoising algorithm based on Poisson statistics. Next, iterations in the image domain are performed to reduce uncorrelated noise while

Table 3 – Characteristics of included studies with tube current reduction

Study	Phantom/ patients	Reference full dose protocol			
		Tube current	ECG gating	CTDI$_{vol}$/ Eff dose	CAC score
Mahnken et al.[15] (2003)	Two Patient groups (n=2*50)	Fixed	Retrospective	22.2 mGy	Agatston: 588.6±762.5*
Shemesh et al.[26] (2005)	Patients (n=51)	165 mAs	Prospective	12.0 mGy	Agatston: 123±223* Mass: 23±43 mg*
Horiguchi et al.[27] (2006)	Patients (n=86)	100 mA	Retrospective	19 mGy 3.2 mSv	Agatston: 479±778*
Dey et al.[28] (2012)	Patients (n= 66)	150 mAs	Prospective	8.5 mGy 1.7 mSv	Agatston: 236±581 Volume: 189±460mm³

Study	Phantom/ patients	Reduced dose protocol			
		Tube current	ECG gating	CTDI$_{vol}$/ Eff dose	CAC score
Mahnken et al.[15] (2003)	Two Patient groups (n=2*50)	Body weight adapted	Retrospective	17.8 mGy	Agatston: 496.0±917.0*
Shemesh et al.[26] (2005)	Patients (n=51)	55 mAs	Prospective	4.0 mGy	Agatston: 126±225* Mass: 24±44 mg*
Horiguchi et al.[27] (2006)	Patients (n=86)	Body weight adapated (60 ±11 mA)	Retrospective	11.4 mGy	Agatston: 463±641*
Dey et al.[28] (2012)	Patients (n= 66)	BMI based tube current selection (120 mAs or 80 mAs)	Prospective	5.1 mGy 1.0 mSv	Agatston: 234±586 Volume: 184±455mm³

Study	Phantom/ patients	Impact		
		CAC score	Risk classification	Dose reduction
Mahnken et al.[15] (2003)	Two Patient groups (n=2*50)	Similar Agatston scores	n.s.	11.6% for men 24.8% for women
Shemesh et al.[26] (2005)	Patients (n=51)	Similar Agatston and mass scores	n.s.	66%
Horiguchi et al.[27] (2006)	Patients (n=86)	High correlation for Agatston score	n.s.	40%
Dey et al.[28] (2012)	Patients (n= 66)	Similar Agatston and volume scores	n.s.	40%

* mean ±SD; ECG: Electrocardiographic; CTDIvol: Computed Tomography Dose Index volume, Eff. dose: effective radiation dose; BMI: body mass index; n.s.: not specified;

Table 4 – Characteristics of included studies with tube current reduction and iterative reconstruction

Study	Phantom/ patients	Reference full dose protocol				
		Tube current	Recon.	Tube voltage	CTDI$_{vol}$/ Eff dose	CAC score
den Harder et al. [33] (2014)	Ex vivo hearts (n=15)	Routine dose	FBP	120 kVp	4.1 mGy	Agatston: 397 (212-1,413)**
den Harder et al. [36] (2016)	Patients (n=28)	50 mAs (<80 kg) 60 mAs (≥80 kg)	FBP	120 kVp	4.8 mGy 0.9 mSv	Agatston: 28.0 (2.1–193.0)**
Willemink et al. [34] (2015)	Patients (n=30)	60 mAs	FBP	120 kVp	4.1-4.9 mGy 0.7-0.9 mSv	Agatston: 26.1 (5.2-192.2)**
Matsuura et al. [35] (2015)	Patients (n=77)	80 mAs	FBP	120 kVp	1.20 mSv	Agatston: 390.7±n.s*
Hecht et al. [37] (2015)	Patients (n=102)	48.5±17.8 mAs	FBP	120 kVp	4.2±1.7 mGy 0.76±0.34 mSv	Agatston: 248.4±497.1*
Blobel et al. [39] (2013)	Phantom§	580 mA	FBP	120 kVp	7.5-14.5 mGy	Agatston: 697.8±7.7*
Tatsugami et al. [40] (2015)	Patients (n=54)	315 mA	FBP	120 kVp	11.2 mGy 2.2 mSv	Agatston: 361.6±n.s.*
Choi et al. [43] (2016)	Patients (n=200)	390 mA	FBP	120 kVp	8.5 mGy 1.4 mSv	n.s.
Tang et al. [41] (2018)	Static phantom§	150-500 mA	FBP	120 kVp	4.1-16.1 mGy	Agatston: 754-765 (small) 716-728 (medium)
	Patient (n=24)	300 mA			13.9 ± 1.2 mGy 2.1 ± 0.3 mSv	Agatston: 258(139-896)**
Luhur et al. [42] (2018)	Patient (n=163)	Automatic exposure control: Mean: 341.7 (SD 147.5) mA	FBP	120 kVp	9.02 (SD 3.98) mGy	Agatston: 184.8±346.4*
Sulaiman [45] (2017)	Patients (n=100)	632 mAs	FBP	120 kVp	8.7±1.4 mGy 2.3±0.4 mSv	Agatston: 138.2±360.6*
Willemink [46] (2014)	Ex vivo hearts (n=15)	259 mA	FBP	120 kVp	4.1 mGy	Agatston: 353.4(n.s.)**
		160 mA			4.2 mGy	409.5(n.s.)**
		220 mA			4.1 mGy	469.0(n.s.)**
		252 mA			4.1 mGy	332.1(n.s.)**
Van der Werf [47] (2017)	Dynamic phantom	500 mA	FBP	120 kVp	10.6 mGy	n.s.
		185 mA			3.2 mGy	n.s.
		285 mA			2.8 mGy	n.s.
		230 mA			6.5 mGy	n.s.

† Scores of the protocol with highest dose reduction and highest level of IR is shown; * mean; ** median (IQR); *** range of medians for the different protocols; § small and medium size thorax
Recon.: Reconstruction algorithm; FBP: filtered back projection; CTDI$_{vol}$: Computed Tomography Dose Index volume; Eff dose: effective radiation dose; n.s.: not specified

5

Study	Phantom/ patients	Reduced dose protocol		
		Tube current	Reconstrucion algorithm	Tube voltage
den Harder et al. [33] (2014)	Ex vivo hearts (n=15)	27%,55%, 82% reduced dose protocol	iDose4 level 1, 4 (Philips)	120 kVp
den Harder et al. [36] (2016)	Patients (n=28)		Prototype IMR level 1,2,3 (Philips)	
Willemink et al. [34](2015)	Patients (n=30)	20 mAs (<80 kg) 2	iDose4 level 4 (Philips)	120 kVp
Matsuura et al. [35] (2015)	Patients (n=77)	4 mAs (≥80 kg) (60% reduced)	IMR level 1,2,3 (Philips)	
Hecht et al. [37] (2015)	Patients (n=102)	12 mAs	iDose4 level 7 (Philips)	120 kVp
Blobel et al. [39] (2013)	Phantom[§]	16 mAs	iDose4 level 7 (Philips)	120 kVp
Tatsugami et al. [40] (2015)	Patients (n=54)	24.5±8.8 mAs	iDose4 level 3 (Philips)	120 kVp
Choi et al. [43] (2016)	Patients (n=200)	10-580 mA in 21 steps	AIDR-3D level standard (Toshiba)	120 kVp
Tang et al. [41] (2018)	Static phantom[§]	104.6 mA	AIDR-3D level standard (Toshiba)	120 kVp
	Patient (n=24)	120 mA	AIDR-3D level standard (Toshiba)	120 kVp
Luhur et al. [42] (2018)	Patient (n=163)	Automatic exposure control: Target image noise: 16-24 HU	AIDR-3D level standard (Toshiba)	120 kVp
Sulaiman [45] (2017)	Patients (n=100)	Automatic exposure control: Target image noise: 20 HU		
Willemink [46] (2014)	Ex vivo hearts (n=15)	91.1 (SD 40.4) mA	AIDR-3D level standard (Toshiba)	120 kVp
		481 mAs	ASIR-V 50% (GE)	120 kVp
		189, 118, 47 mA	FBP, iDose4 level 1 and 6 (Philips)	120 kVp
		120, 70, 30 mA	FBP, AIDR-3D level mild and strong (Toshiba)	
Van der Werf [47] (2017)	Dynamic phantom	160, 105, 45 mA	FBP, ASIR 20% and 80% (GE)	
		184, 116, 48 mA	FBP, SAFIRE level 1 and 5 (Siemens)	
		300 mA	FBP,	120 kVp
		100 mA	ASIR-V Level 20%, 60%, 100% (GE)	
		111mA	FBP,	
		37 mA	iDose4 Level 1, 5 and 7 (Philips)	
		171 mA	FBP,	
		37 mA	SAFIRE level 1,3 and 5 (Siemens)	
		138 mA	FBP,	
		46 mA	AIDR-3D level: Weak, standard, strong (Toshiba)	

† Scores of the protocol with highest dose reduction and highest level of IR is shown; * mean; ** median (IQR); *** range of medians for the different protocols; § small and medium size thorax

Reduced dose protocol		Impact		
CTDI$_{vol}$/ Eff dose	CAC score†	CAC score	Risk classification	Dose reduction
3.0 mGy 1.9 mGy 0.8 mGy	Agatston: 385 (211-1,428)**	Similar Agatston scores	n.s.	27-82%
	Agatston: 377 (198-1,403)**	Lower Agatston scores	n.s.	
1.9 mGy 0.4 mSv	Agatston: 26.9 (0.0–230.1)**	Similar Agatston scores	18% reclassification	60%
	Agatston: 2 1.4 (0.0–197.0)**	Similar Agatston scores	21% reclassification	
0.8-1.0 mGy 0.2-0.2 mSv	Agatston: 22.9 (5.9-195.5)**	Lower Agatston scores	15% reclassification	80%
0.24 mSv	Agatston: 377.7±n.s*	Lower Agatston scores	n.s.	80%
2.0±0.7 mGy 0.37±0.16 mSv	Agatston: 237.9±489.5*	Lower Agatston scores	8% reclassification	47%
1.5-2.6 mGy	Agatston: 678.8±14.3 * (s) 643.9±13.4 * (m)	Lower Agatston scores	n.s.	82%
3.6 mGy 0.7 mSv	Agatston: 356.8±n.s.*	Lower Agatston scores	n.s.	67%
2.3 mGy 0.37 mSv	n.s.	n.s.	11% reclassification	74%
0.7-5.5 mGy	Agatston: 764-795 (s) 708-722 (m)	Higher Agatston score for for small phantom. Similar Agatston score for medium sized phantom	n.s.	66-83%
5.7 ± 2.2 mGy 0.9 ± 0.4 mSv	Agatston: 226(138-993)**	Similar Agatston scores	8% reclassification	57.8%
2.2 (SD 1.0) mGy	Agatston: 185.3±351.3*	Similar Agatston scores	5% reclassification	75%
6.6±2.9 mGy 1.8±0.8 mSv	Agatston: 137.3±356.4*	Similar Agatston scores	3% reclassification	25%
0.8 mGy	Agatston: 354.3-359.6***	Trend towards lower Agatston scores for IR	n.s.	80%
0.8 mGy	381.0-408.0***			
0.8 mGy	438.5-466.6***			
0.8 mGy	318.5-327.2***			
6.4 mGy	n.s.	Decrease in Agatston score for increasing levels of IR	n.s.	40%
2.1 mGy				80%
1.9 mGy	n.s.			40%
0.6 mGy				80%
1.7 mGy	n.s.			40%
0.6 mGy				80%
3.9 mGy	n.s.			40%
1.3 mGy				80%

Recon.: Reconstruction algorithm; FBP: filtered back projection; CTDI$_{vol}$: Computed Tomography Dose Index volume; Eff dose: effective radiation dose; n.s.: not specified; (s) = small; (m) = medium

preserving underlying edges associated with true anatomy.[30] iDose4 can be set to seven noise-reducing levels, higher levels result in lower noise. Besides accounting for noise behavior in an image, the advanced model based IR (IMR, Philips Healthcare, Best, The Netherlands) also accounts for data and image statistics and detailed CT system geometry in the iterative process.[31,32] In total 5 studies determined the impact of iDose4 including 237 patients and 15 ex-vivo hearts, and two studies determined the impact of (prototype) IMR in 28 patients and 15 ex-vivo hearts.

Three studies by den Harder et al., Willemink et al., and Matsuura et al. showed that a high level of iDose4 (level 7) in combination with up to 80% dose reduction results in lower Agatston scores.[33–35] Reclassification for high level iDose4 occurred in 15% of the cases with 2 out of 30 patients going from a positive score with full dose FBP to a 0-score for iDose4 level 7 for one study[34], while in another study none of the patients (n=77) were reclassified from a positive to 0-score.[35] Besides, the former study by Willemink et al. showed that lower levels of iDose4 and higher dose did not lead to less reclassification.[34] This is confirmed in another patient study (n=28) by den Harder et al.: iDose4 level 4 with tube current reduction of 60% led to reclassification in 18% of the cases, compared to the full dose FBP protocol.[36] Contrary, a larger patient study (n=102) by Hecht et al. showed that a lower level of iDose4 (level 3) with tube current reduction results in reclassification of 8% of the cases.[37] In this study the difference in mean Agatston score between full dose and low dose was 17.4±25.8 which was smaller than the variability for repetitive scanning with the same mAs.[37] None of the patients were reclassified from a positive to 0-score or from very high score (400+) to a lower risk category.[37] Radiation dose of the full dose protocols ranged from 4.1 to 4.9 mGy and was reduced to 0.8 to 2.0 mGy with the use of iDose4 with reclassification of 8% to 18%.

The combination of tube current reduction with IMR levels 1-3 led to lower Agatston scores in 15 ex vivo hearts [33] but to similar results at 60% of the full dose in 28 patients.[36] Nonetheless, reducing the dose further led to lower Agatston scores at all IMR levels and reclassification in 21% of the cases.[36]

Adaptive Iterative Dose Reduction 3D (AIDR-3D)

The Adaptive Iterative Dose Reduction 3D (AIDR-3D, Toshiba Medical Systems, Otawara, Japan) is based on iteratively adaptive filtering in the image domain and noise reduction in the raw data domain, by taking scanner model, statistical noise model and anatomical model into account.[38]

The impact of AIDR-3Dand tube current reduction was examined in five studies, including two phantoms and 441 patients.[39–43] These studies showed varying results for the impact on the Agatston score of AIDR-3D (level: standard) and tube current reduction. In a phantom study by Blobel et al. the mean Agatston score was approximately 5% lower for reduced dose AIDR-3D protocol compared to full dose FBP protocol.[39] These differences were significant, however reclassification of individuals was not evaluated. Likewise, in a patient study (n=54) by Tatsugami et al. the mean Agatston score was 7% lower for AIDR-3D compared to FBP and per individual the mean difference was 15.9%, with no false negatives.[40] Another patient study (n=24) by Tang et al. reported similar mean Agatston scores for AIDR-3D and full dose protocol, but the median Agatston score was reduced with 12% and two patients (8%) were reclassified.[41] Contrary, in a larger patient study (n=163) by Luhur et al. the mean Agatston score difference between full and reduced dose protocol showed no systematic deviation.[42] Besides, for the reduced AIDR-3D protocol, no false positive and false negative scores were seen, but 5% of the patients were reclassified.[42] In a large patient study (n=200) by Choi et al. the reproducibility and reclassification were assessed for low dose AIDR-3D compared to full dose FBP.[43] In total, 11% of the patients were reclassified (κ=0.86), and of these: 8 patients were reclassified from a 0-score to a positive score or vice versa. The rescan agreement for risk categorization was κ=0.87 (95% CI:0.83-0.93) and κ=0.91 (95% CI:0.86-0.95) for low dose AIDR-3D and full dose FBP, respectively.[43] Radiation doses of the full dose protocols ranged from 4.1 to 16.1 mGy and reduced to radiation doses ranging from 0.7 to 5.7 mGy with a reclassification of 5% to 11% when AIDR-3D was used.

Adaptive Statistical IR (ASIR-V)

The Adaptive Statistical IR-V (ASIR-V, GE Healthcare, Chicago, Illinois, USA) is based on an algorithm focusing mainly on the modeling of system noise statistics, objects and physics and less focused on the modeling of system optics.[44] Unlike the former mentioned IR algorithms, ASIR-V entails forward and backward projection between the raw data and image data domain. In one patient study (n=100) by Sulaiman et al. a low dose protocol was used including Adaptive Statistical IR-V and a low tube current and was compared with a full dose FBP protocol.[45] The Agatston score and mass score were not significantly different for low and full dose, respectively. Two patients were reclassified to a lower and one patient to a higher risk category for the low dose protocol, but no patients were reclassified from positive to 0-score or vice versa. The mean radiation dose of the full dose protocols was 8.7±1.4 mGy and reduced to 6.6±2.9 mGy, corresponding to a reclassification of 3% of the patients with the use of ASIR-V.[45]

Comparison of various IR algorithms

In a study by Willemink et al. the impact of tube current reduction with FBP and IR (low and high level) was determined on four scanners from different vendors in 15 ex-vivo hearts.[46] The Agatston score was similar for all dose-reduced protocols with FBP compared to the full dose protocol with FBP. Using IR with reduced-dose protocols, resulted in a trend towards lower Agatston scores, with significant differences for one vendor compared to the full dose FBP protocol. Similar results were found for the other vendors, but IR-reduced dose protocols did not improve the reproducibility compared to FBP-reduced-dose protocols. No hearts were reclassified for the reduced dose protocols of Philips and Siemens CT systems, whereas maximum of two hearts (13%) were reclassified to a lower risk category for IR-reduced-dose protocols for Toshiba and GE CT systems. Radiation dose was reduced with 80% by lowering the tube current and using FBP.[46]

In a dynamic phantom study by Van der Werf et al. on four CT systems from different vendors a dose reduction of 40% resulted in not significantly different Agatston scores when FBP (Philips, Siemens, Toshiba) or IR (GE) was applied.[47] For tube current reductions of 80%, Agatston scores were not significantly different for Philips in combination with iDose levels 5 and 7. For mass scores, similar results were found. For one vendor (Siemens) 80% dose reduction in combination with IR also led to not significantly different mass scores.

Spectral shaping with tin filtration

In the past two years, four studies evaluated the impact of spectral shaping by tin-filtration on CAC scoring in 140 patients and two phantoms, see *Table 5* and *Figure 3*. In spectral shaping CT, a tin-filter is used at the X-ray tube. Low energy photons that contribute little to the final image are filtered out of the x-ray beam by this tin-filter, thereby reducing the radiation dose received by the patient. Studies report dose reductions of 62 to 85% with the use of a tin-filter (Sn) at 100 kVp compared to reference full dose acquisitions with 120 kVp.[48–51]

Phantoms studies have shown that CAC scoring is feasible when using tin-filtration.[48,51] In a static phantom study by McQuiston et al., tin-filtration at Sn100 kVp acquisition yielded similar Agatston scores as the 120 kVp sequential acquisition at FBP projection.[51] However, risk reclassification was not evaluated. In another static and dynamic phantom study by Vonder et al., high-pitch spiral Sn100 kVp acquisitions led to lower Agatston scores compared to the 120 kVp

Table 5 – Characteristics of included studies with spectral shaping

Study	Phantom/ patients	Reference full dose protocol			
		Tube voltage (kVp)	HU threshold	CTDI$_{vol}$/ Eff dose	CAC score
McQuiston et al. [51] (2016)	Static phantom	120	130 HU	0.48 mGy	Agatston: 686-688
Vonder et al. [48] (2017)	Static and dynamic phantom	120	130 HU	2.44 mGy	Agatston: 638.3(625.7-653.2)**
Tesche et al. [49] (2017)	Patient (n=70)	120	130 HU	4.1 mGy 0.82±0.32 mSv	Agatston: 41.2(2.1- 180.2)**
Apfaltrer et al. [50] (2018)	Patient (n=70)	120	130 HU	2.2 ± 0.7 mGy 0.57±0.2 mSv	Agatston: 41.7(0.7-207.2)**

Study	Phantom/ patients	Reduced dose protocol			
		Tube voltage (kVp)	HU threshold	CTDI$_{vol}$/ Eff dose	CAC score
McQuiston et al. [51] (2016)	Static phantom	Sn100	130 HU	0.07 mGy	Agatston: 639-672
Vonder et al. [48] (2017)	Static and dynamic phantom	Sn100	130HU	0.65 mGy	Agatston: 600.8(593.7-621.3)**
Tesche et al. [49] (2017)	Patient (n=70)		117 HU		657.4 (651.1-675.2)**
Apfaltrer et al. [50] (2018)	Patient (n=70)	Sn100	130 HU	1.3 mGy 0.19±0.05 mSv	Agatston: 38.2(1.4-156.9)**
		Sn100	130 HU	0.6 ± 0.3 mGy 0.13±0.07 mSv	Agatston: 34.9(0.7–197.1)**

Study	Phantom/ patients	Impact		
		CAC score	Risk classification	Dose reduction
McQuiston et al. [51] (2016)	Static phantom	Similar Agatston score	n.s.	85%
Vonder et al. [48] (2017)	Static and dynamic phantom	Lower Agatston scores	n.s.	62%
Tesche et al. [49] (2017)	Patient (n=70)	Similar Agatston scores	n.s.	
Apfaltrer et al. [50] (2018)	Patient (n=70)	Lower Agatston scores	Excellent agreement (κ=0.98), 3% reclassification	75%
		Lower Agatston scores	Excellent agreement (κ=0.98), 4% reclassification	78%

** median (IQR); CTDI$_{vol}$: Computed Tomography Dose Index volume; Eff dose: effective radiation dose; Sn100: Tin-filter with 100 kVp; n.s.: not specified;

high-pitch spiral acquisition.[48] However, similar Agatston scores were achieved by using an adapted 117 HU threshold, while reducing the dose with 62% for Sn100 kVp compared to 120 kVp. In this study risk reclassification was also not assessed.

Two patient studies (n=70) by Tesche et al. and Apfaltrer et al. also showed an underestimation of the Agatston score with sequential and high-pitch spiral Sn100 kVp acquisitions compared to sequential and high-pitch spiral 120 kVp acquisitions.[49,50] In 3% and 4% of the cases, patients (n=2, n=3) were reclassified into a lower risk category for Sn100 kVp compared to the 120 kVp acquisitions, for respectively sequential and high-pitch spiral acquisitions.[49,50] Nevertheless, in none of the patients false negative or false positive scores were encountered for Sn100 kVp.[49,50] The radiation dose was reduced by 75% and 78% to $CTDI_{vol}$ of 1.3±1.7 mGy and 0.6±0.3 mGy for Sn100 kVp sequential and high-pitch spiral acquisitions and led to reclassification of 3% and 4%, respectively.[49,50]

DISCUSSION

In this study, dose reduction techniques in CAC imaging were systematically reviewed for their impact on Agatston score and CVD risk stratification. In 78% of the studies, the radiation dose was reduced ≥50%, with $CTDI_{vol}$ ranging from 0.6 to 5.5 mGy. However, the different dose reduction techniques had varying impact on Agatston scores and CVD risk stratification.

CVD risk stratification is based on the Agatston score. Agatston score risk categories are typically defined as very low, moderate, high, and very high risk for scores of 0, 1-99, 100-399 and ≥400, respectively. Under- or overestimation of the Agatston score could lead to reclassification of an individual and might lead to inadequate prevention treatment in future CVD CT screening trials. Besides, progression analysis of CAC could become relevant, and no or limited under- or overestimation of the Agatston score would then be required. However, there is no consensus about the clinical acceptable range of difference in Agatston score and acceptable percentage of reclassification. In general, inter scan reproducibility of the full dose protocol is regarded as the acceptable level of variance by most of the studies included in this review, however the level of variance of a full dose protocol can vary among the different protocols on different scanners.[52] Besides, more studies with large sample sizes (n>200 patients) are needed to reliably show the variance and impact of a specific dose-reduction technique.

Impact on Agatston score

This review showed that tube voltage reduction led to high agreement for Agatston score but with a systematic under- or overestimation of the Agatston score even if the HU-threshold was adapted. Tube voltage reduction resulted in radiation dose protocols of 0.6-1.2 mGy. Contrary to only tube voltage reduction, tube voltage reduction with IR showed similar results for Agatston score at a radiation dose of 0.4-1.6 mGy. Likewise, studies evaluating tube current-optimized and tube current-reduced protocols, reported similar Agatston scores or a high agreement of Agatston scores. Further reduction of the tube current resulted in excessive noise levels. Although guidelines suggest to keep the noise level below 23 HU,[53] various maximum noise levels were used in the included studies of this review. To not exceed the maximum noise level, IR can be applied which allowed for a large decrease in tube current, while maintaining similar noise levels at reduced radiation doses. However, high IR levels showed an underestimation of the Agatston score, whereas low and moderate levels showed similar Agatston scores at a radiation dose of 0.8-3.9 mGy. Spectral shaping with tin-filtration resulted in substantial dose reductions in CAC imaging (total dose: 0.6-1.3 mGy), but resulted in underestimated Agatston scores. One study showed that adaptation of the HU threshold resulted in similar Agatston scores for a tin-filter protocol compared to the full dose protocol.

Impact on risk stratification

Only twelve studies (42%) in this review reported results regarding risk stratification. Although two studies reported significant differences for Agatston score for tube voltage reduction, reclassification of patients into lower risk categories was limited to 2% to 7%. Remarkably, no studies reported reclassification percentages of tube voltage reduction with IR or of tube current reduction alone. Contrary, eight studies reported reclassification rates for tube current reduction with IR. Lower levels of IR tended to result in less reclassification, but a wide range of reclassification was reported: 3% up to 21% across all types of IR. One patient study (n=28) reported similar Agatston scores, but showed high reclassification rates (18% and 21%) for iDose4 level 4 and IMR. Contrary, another patient study (n=102) used iDose4 level 3, and showed a moderate reclassification rate (8%). This difference could be caused by the limited number of included patients in the former study, or due to the fact that the Agatston score distribution of both populations is considerably different. Spectral shaping with tin-filter led to low reclassification rates (3% and

4%). Nevertheless, so far only limited patient studies have been performed and tin-filter is only available on latest generation dual-source CT. Future research is needed with larger sample sizes and evaluation of reclassification rates.

Limitations

This systematic review study has some limitations. First, different maximum noise levels were used in the included studies. Therefore, a higher decrease in tube current was allowed in some studies. This may explain the wide range of reduced radiation doses reported by the different studies. Second, there is no consensus about the maximum acceptable Agatston difference or reclassification, therefore studies used different ways of reporting and interpreting the clinical significance of their results. For instance, some studies report no significant difference in mean Agatston scores, while reporting median Agatston scores would be more appropriate. Besides, studies reported high kappa values for dose-reduction techniques and the full dose protocol, and therefore concluded that there is a high agreement for risk classification. However, it remains unclear whether only a high agreement ($\kappa \geq 0.8$) for risk classification is sufficient to allow for a wide-spread implementation of a dose-reduction technique for screening, because also at high kappa values a considerable percentage of individuals may still be reclassified and the clinical impact of that is unknown. Third, there were substantial differences among studies for reported full dose and reduced dose protocols. All included studies (n=5) published in 2006 or earlier reported a full dose of >10.0 mGy and reduced dose of >3.5 mGy. Whereas the majority (n=15) of the published studies in 2012 and later reported full doses ranging from 0.9 to 4.8 mGy and reduced doses of 0.1 to 3.0 mGy. Advances in CT technology over time have thus resulted in dose reductions in the last 15 years. Finally, only one phantom study in this review investigated the impact of combining different dose reduction techniques on one CT system. Potentially, combining different dose reduction techniques could lead to an even further dose reduction in patients undergoing CAC imaging, than reported so far for the separate techniques.

CONCLUSION

Radiation dose reduction techniques allowed for radiation dose reductions of 50% or more in 78% of CAC CT studies. However, risk reclassification was influenced in 3% (dose reduction of 75%) up to 21% (dose reduction of 60%) of individuals, depending on the acquisition technique. Specific dose reduced protocols, including

either tube current reduction and IR or spectral shaping with tin filtration, that showed low reclassification rates may potentially be used in CAC scanning and in future population-based screening for CVD risk stratification. Tube current reduction with IR is most intensively investigated in current literature as a method for dose reduction in CAC imaging. Contrary to tin-filter, tube current reduction is applicable on all type of CT scanners with limited impact on risk stratification. Future research in dose reduction techniques of CAC imaging should focus on larger patient studies evaluating CVD risk reclassification rates, Agatston score distribution and the reproducibility of the dose reduced protocol.

5

5

REFERENCES

1. Mahabadi AA, Mohlenkamp S, Lehmann N, et al. CAC Score Improves Coronary and CV Risk Assessment Above Statin Indication by ESC and AHA/ACC Primary Prevention Guidelines. JACC Cardiovasc Imaging. 2017;10(2):143-153. doi:10.1016/j.jcmg.2016.03.022

2. Vonder M, van der Aalst CM, Vliegenthart R, et al. Coronary Artery Calcium Imaging in the ROBINSCA Trial. Acad Radiol. 2018;25(1):118-128. doi:10.1016/j.acra.2017.07.010

3. Hecht H, Blaha MJ, Berman DS, et al. Clinical indications for coronary artery calcium scoring in asymptomatic patients: Expert consensus statement from the Society of Cardiovascular Computed Tomography. J Cardiovasc Comput Tomogr. 2017;11(2):157-168. doi:10.1016/j.jcct.2017.02.010

4. Benjamin EJ, Blaha MJ, Chiuve SE, et al. Heart Disease and Stroke Statistics'2017 Update: A Report from the American Heart Association. Vol 135.; 2017. doi:10.1161/CIR.0000000000000485

5. Hausleiter J, Martinoff S, Hadamitzky M, et al. Image Quality and Radiation Exposure With a Low Tube Voltage Protocol for Coronary CT Angiography. JACC Cardiovasc Imaging. 2010;3(11):1113-1123. doi:10.1016/j.jcmg.2010.08.016

6. Øvrehus KA, Marwan M, Bøtker HE, Achenbach S, Nørgaard BL. Reproducibility of coronary plaque detection and characterization using low radiation dose coronary computed tomographic angiography in patients with intermediate likelihood of coronary artery disease (ReSCAN study). Int J Cardiovasc Imaging. 2012;28(4):889-899. doi:10.1007/s10554-011-9895-1

7. Hell MM, Bittner D, Schuhbaeck A, et al. Prospectively ECG-triggered high-pitch coronary angiography with third-generation dual-source CT at 70 kVp tube voltage: Feasibility, image quality, radiation dose, and effect of iterative reconstruction. J Cardiovasc Comput Tomogr. 2014;8(6):418-425. doi:10.1016/j.jcct.2014.09.003

8. Chinnaiyan KM, Bilolikar AN, Walsh E, et al. CT dose reduction using prospectively triggered or fast-pitch spiral technique employed in cardiothoracic imaging (the CT dose study). J Cardiovasc Comput Tomogr. 2014;8(3):205-214. doi:10.1016/j.jcct.2014.04.001

9. Deseive S, Pugliese F, Meave A, et al. Image quality and radiation dose of a prospectively electrocardiography-triggered high-pitch data acquisition strategy for coronary CT angiography: The multicenter, randomized PROTECTION IV study. J Cardiovasc Comput Tomogr. 2015;9(4):278-285. doi:10.1016/j.jcct.2015.03.001

10. Deseive S, Chen MY, Korosoglou G, et al. Prospective Randomized Trial on Radiation Dose Estimates of CT Angiography Applying Iterative Image Reconstruction. JACC Cardiovasc Imaging. 2015;8(8):888-896. doi:10.1016/j.jcmg.2015.02.024

11. Schuhbaeck A, Achenbach S, Layritz C, et al. Image quality of ultra-low radiation exposure coronary CT angiography with an effective dose <0.1 mSv using high-pitch spiral acquisition and raw data-based iterative reconstruction. Eur Radiol. 2013;23(3):597-606. doi:10.1007/s00330-012-2656-2

12. Blaha MJ, Mortensen MB, Kianoush S, Tota-Maharaj R, Cainzos-Achirica M. Coronary Artery Calcium Scoring: Is It Time for a Change in Methodology? JACC Cardiovasc Imaging. Published online 2017. doi:10.1016/j.jcmg.2017.05.007

13. Vonder M, van der Aalst CM, Vliegenthart R, et al. Coronary artery calcium imaging in the ROBINSCA trial: Rationale, Design and Technical Background. Acad Radiol. 2017;to be publ.

14. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med. 2009;6(7):e1000097. doi:10.1371/journal.pmed.1000097

15. Mahnken AH, Wildberger JE, Simon J, et al. Detection of coronary calcifications: Feasibility of dose reduction with a body weight-adapted examination protocol. Am J Roentgenol. 2003;181(2):533-538. doi:10.2214/ajr.181.2.1810533

16. van Hamersvelt RW, Eijsvoogel NG, Mihl C, et al. Contrast agent concentration optimization in CTA using low tube voltage and dual-energy CT in multiple vendors: a phantom study. Int J Cardiovasc Imaging. 2018;0(0):1-11. doi:10.1007/s10554-018-1329-x

17. Groen JM, Kofoed KF, Zacho M, Vliegenthart R, Willems TP, Greuter MJW. Calcium score of small coronary calcifications on multidetector computed tomography: Results from a static phantom study. Eur J Radiol. 2013;82(2):e58-e63. doi:10.1016/j.ejrad.2012.09.018

18. Thomas CK, Mühlenbruch G, Wildberger JE, et al. Coronary artery calcium scoring with multislice computed tomography: in vitro assessment of a low tube voltage protocol. Invest Radiol. 2006;41(9):668-673. doi:10.1097/01.rli.0000233324.09603.dd

19. Marwan M, Mettin C, Pflederer T, et al. Very low-dose coronary artery calcium scanning with high-pitch spiral acquisition mode: Comparison between 120-kV and 100-kV tube voltage protocols. J Cardiovasc Comput Tomogr. 2013;7(1):32-38. doi:10.1016/j.jcct.2012.11.004

20. Jakobs TF, Wintersperger BJ, Herzog P, et al. Ultra-low-dose coronary artery calcium screening using multislice CT with retrospective ECG gating. Eur Radiol. 2003;13(8):1923-1930. doi:10.1007/s00330-003-1895-7

21. Gräni C, Vontobel J, Benz DC, et al. Ultra-low-dose coronary artery calcium scoring using novel scoring thresholds for low tube voltage protocols—a pilot study. Eur Hear J - Cardiovasc Imaging. 2018;(February):1-10. doi:10.1093/ehjci/jey019

22. Naoum C, Blanke P, Leipsic J. Iterative reconstruction in cardiac CT. J Cardiovasc Comput Tomogr. 2015;9(4):255-263. doi:10.1016/j.jcct.2015.04.004

23. Willemink MJ, de Jong P a, Leiner T, et al. Iterative reconstruction techniques for computed tomography Part 1: Technical principles. Eur Radiol. Published online January 2013. doi:10.1007/s00330-012-2765-y

24. Blobel J, Mews J, Goatman KA, Schuijf JD, Overlaet W. Calibration of coronary calcium scores determined using iterative image reconstruction (AIDR 3D) at 120, 100, and 80 kVp. Med Phys. 2016;43(4):1921-1932. doi:10.1118/1.4942484

25. Vonder M, Pelgrim GJ, Meyer M, Henzler T, Oudkerk M, Vliegenthart R. Dose reduction techniques in coronary calcium scoring: The effect of iterative reconstruction combined with low tube voltage on calcium scores in a thoracic phantom. Eur J Radiol. 2017;93(January):229-235. doi:10.1016/j.ejrad.2017.06.001

26. Shemesh J, Evron R, Koren-Morag N, et al. Coronary Artery Calcium Measurement with Multi–Detector Row CT and Low Radiation Dose: Comparison between 55 and 165 mAs. Radiology. 2005;236(3):810-814. doi:10.1148/radiol.2363040039

27. Horiguchi J, Yamamoto H, Hirai N, et al. Variability of Repeated Coronary Artery Calcium Measurements on Low-Dose ECG-Gated 16-MDCT. Am J Roentgenol. 2006;187(1):W1-W6. doi:10.2214/AJR.05.0052

5

28. Dey D, Nakazato R, Pimentel R, et al. Low radiation coronary calcium scoring by dual-source CT with tube current optimization based on patient body size. J Cardiovasc Comput Tomogr. 2012;6(2):113-120. doi:10.1016/j.jcct.2011.12.008

29. Baron KB, Choi AD, Chen MY. Low Radiation Dose Calcium Scoring: Evidence and Techniques. Curr Cardiovasc Imaging Rep. 2016;9(4):1-8. doi:10.1007/s12410-016-9373-1

30. Noël PB, Fingerle AA, Renger B, Münzel D, Rummeny EJ, Dobritz M. Initial Performance Characterization of a Clinical Noise–Suppressing Reconstruction Algorithm for MDCT. Am J Roentgenol. 2011;197(6):1404-1409. doi:10.2214/AJR.11.6907

31. Morton T. Basic IMR testing , considerations and image quality trends. Philips computed tomograhpy.

32. Mehta D, Thompson R, Morton T, Dhanantwari A, Shefer E. Iterative model reconstruction: simultaneously lowered computed tomography radiation dose and improved image quality. Med Phys Int. 2013;1(2):147-155.

33. den Harder AM, Willemink MJ, Bleys RL a W, et al. Dose reduction for coronary calcium scoring with hybrid and model-based iterative reconstruction: an ex vivo study. Int J Cardiovasc Imaging. 2014;30(6):1125-1133. doi:10.1007/s10554-014-0434-8

34. Willemink MJ, den Harder AM, Foppen W, et al. Finding the optimal dose reduction and iterative reconstruction level for coronary calcium scoring. J Cardiovasc Comput Tomogr. 2016;10(1):69-75. doi:10.1016/j.jcct.2015.08.004

35. Matsuura N, Urashima M, Fukumoto W, et al. Radiation Dose Reduction at Coronary Artery Calcium Scoring by Using a Low Tube Current Technique and Hybrid Iterative Reconstruction. J Comput Assist Tomogr. 2014;00(00):1-6. doi:10.1097/RCT.0000000000000168

36. den Harder AM, Wolterink JM, Willemink MJ, et al. Submillisievert coronary calcium quantification using model-based iterative reconstruction: A within-patient analysis. Eur J Radiol. 2016;85(11):2152-2159. doi:10.1016/j.ejrad.2016.09.028

37. Hecht HS, De Siqueira MEM, Cham M, et al. Low- vs. Standard-dose coronary artery calcium scanning. Eur Heart J Cardiovasc Imaging. 2015;16(4):358-363. doi:10.1093/ehjci/jeu218

38. Angel E. AIDR 3D Iterative Reconstruction: Integrated, Automated and Adaptive Dose reduction. Toshiba - White paper.

39. Blobel J, Mews J, Schuijf JD, Overlaet W, Dipl-Ing Desiree Schuijf J, Overlaet W. Determining the Radiation Dose Reduction Potential for Coronary Calcium Scanning With Computed Tomography the Adaptive Iterative Dose Reduction Algorithm for Image Reconstruction. Invest Radiol. 2013;48(12):857-862.

40. Tatsugami F, Higaki T, Fukumoto W, et al. Radiation dose reduction for coronary artery calcium scoring at 320-detector CT with adaptive iterative dose reduction 3D. Int J Cardiovasc Imaging. 2015;31(5):1045-1052. doi:10.1007/s10554-015-0637-7

41. Tang Y-C, Liu Y-C, Hsu M-Y, Tsai H-Y, Chen C-M. Adaptive Iterative Dose Reduction 3D Integrated with Automatic Tube Current Modulation for CT Coronary Artery Calcium Quantification: Comparison to Traditional Filtered Back Projection in an Anthropomorphic Phantom and Patients. Acad Radiol. Published online January 2018. doi:10.1016/j.acra.2017.12.018

42. Luhur R, Schuijf JD, Hamm B, et al. Accuracy of coronary artery calcium scoring with tube current reduction by 75%, using an adaptive iterative reconstruction algorithm. Br J Radiol. 2018;(September 2017):20170678. doi:10.1259/bjr.20170678

43. Choi AD, Leifer ES, Yu J, et al. Prospective evaluation of the influence of iterative reconstruction on the reproducibility of coronary calcium quantification in reduced radiation dose 320 detector row CT. J Cardiovasc Comput Tomogr. 2016;10(5):359-363. doi:10.1016/j.jcct.2016.07.016

44. Yue J. Fan M, Melknyk R. Benefits of ASiR-V Reconstruction for Reducing Patient Radiation Dose and Preserving Diagnostic Quality in CT Exams. GE Healthcare - White paper.

45. Sulaiman N, Soon J, Park J kwan, et al. Comparison of low-dose coronary artery calcium scoring using low tube current technique and hybrid iterative reconstruction vs. filtered back projection. Clin Imaging. 2017;43(November 2015):19-23. doi:10.1016/j.clinimag.2017.01.017

46. Willemink MJ, Takx RAP, De Jong PA, et al. The impact of CT radiation dose reduction and iterative reconstruction algorithms from four different vendors on coronary calcium scoring. Eur Radiol. 2014;24(9):2201-2212. doi:10.1007/s00330-014-3217-7

47. van der Werf NR, Willemink MJ, Willems TP, Greuter MJW, Leiner T. Influence of dose reduction and iterative reconstruction on CT calcium scores: a multi-manufacturer dynamic phantom study. Int J Cardiovasc Imaging. 2017;33(6):899-914. doi:10.1007/s10554-017-1061-y

48. Vonder M, Pelgrim GJ, Huijsse SEM, et al. Feasibility of spectral shaping for detection and quantification of coronary calcifications in ultra-low dose CT. Eur Radiol. 2017;27(5):2047-2054. doi:10.1007/s00330-016-4507-z

49. Tesche C, De Cecco C, Vliegenthart R, et al. Accuracy and Radiation Dose Reduction Using Low-Voltage Computed Tomography Coronary Artery Calcium Scoring With Tin Filtration. Am J Cardiol. 2017;119(4):675-680. doi:10.1016/j.amjcard.2016.10.051

50. Apfaltrer G, Albrecht MH, Schoepf UJ, et al. High-pitch low-voltage CT coronary artery calcium scoring with tin filtration: accuracy and radiation dose reduction. Eur Radiol. Published online February 2018. doi:10.1007/s00330-017-5249-2

51. McQuiston AD, Muscogiuri G, Schoepf UJ, et al. Approaches to ultra-low radiation dose coronary artery calcium scoring based on 3rd generation dual-source CT: A phantom study. Eur J Radiol. 2016;85(1):39-47. doi:10.1016/j.ejrad.2015.10.023

52. Willemink MJ, Vliegenthart R, Takx RAP, et al. Coronary Artery Calcification Scoring with State-of-the-Art CT Scanners from Different Vendors Has Substantial Effect on Risk Classification. Radiology. 2014;273(3):695-702. doi:10.1148/radiol.14140066

53. Voros S, Rivera JJ, Berman DS, et al. Guideline for minimizing radiation exposure during acquisition of coronary artery calcium scans with the use of multidetector computed tomography: A report by the Society for Atherosclerosis Imaging and Prevention Tomographic Imaging and Prevention Counci. J Cardiovasc Comput Tomogr. 2011;5(2):75-83. doi:10.1016/j.jcct.2011.01.003
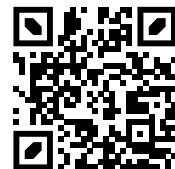
Online material is available via:
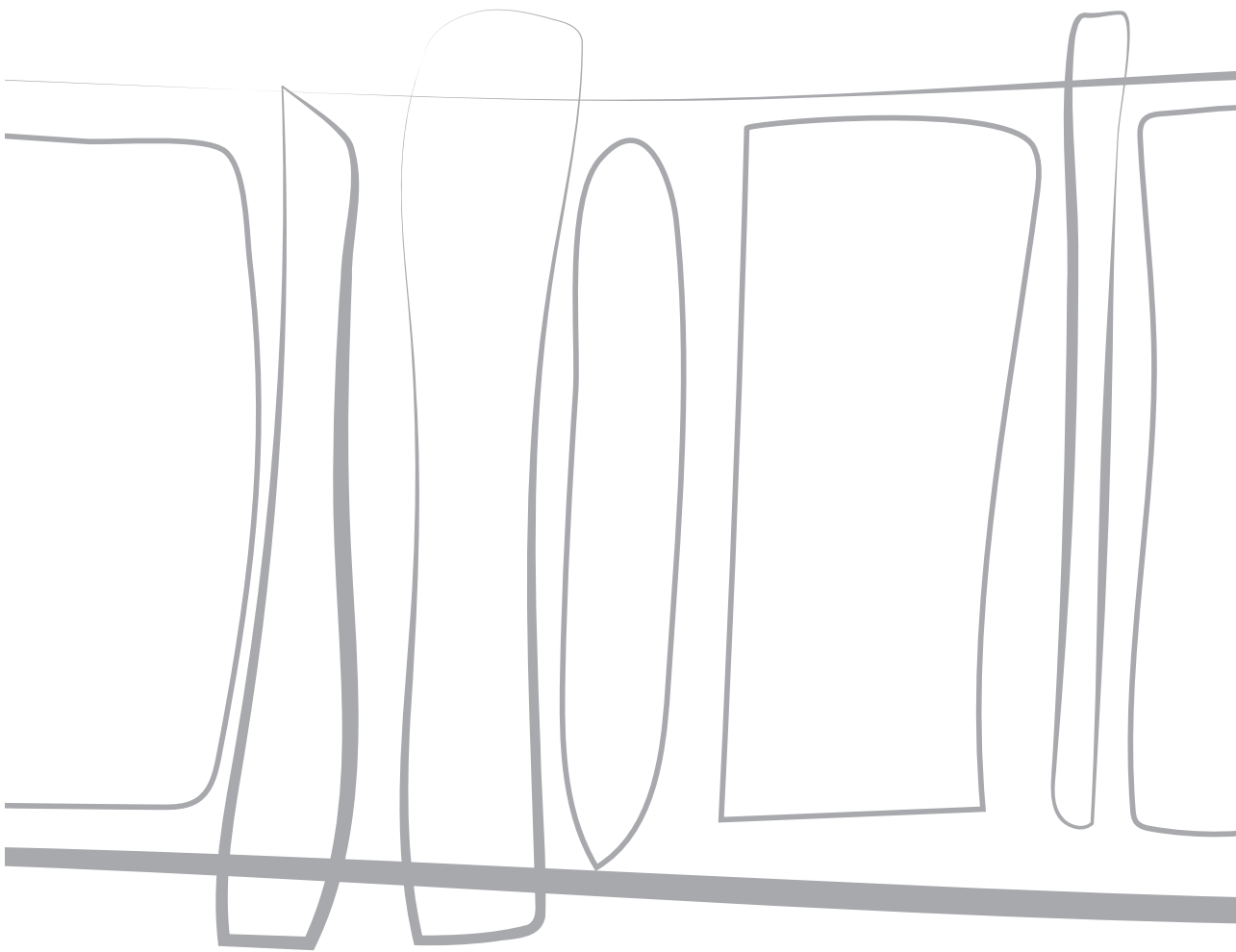
CHAPTER 6

Coronary artery calcium: a technical argument for a new scoring method

Martin J. Willemink, MD PhD

Niels R. van der Werf, MSc

Koen Nieman, MD PhD

Marcel J.W. Greuter, PhD

Lynne Koweek, MD

Dominik Fleischmann, MD

6

ABSTRACT

Coronary artery calcium (CAC) is a strong predictor for future cardiovascular events. Traditionally CAC is quantified with the Agatston score, which was developed in the 1990s using electron beam tomography (EBT). While EBT technology has been replaced by modern multiple-detector row CT scanners, CAC quantification has not changed although current literature indicates that the traditional Agatston score lacks reproducibility resulting in subjects being incorrectly classified. The traditional Agatston score is affected by newer CT technology including more advanced image acquisition and reconstruction techniques. Moreover, recent studies have shown that low-density CAC is associated with increased cardiovascular risk and assessment of CAC distribution may be a beneficial prognostic value. CAC quantification should be optimized for modern CT scanners. In this paper we describe how to optimize CAC quantification using a lower radiation dose and isovolumetric data set that can be applied to contemporary CT scanners and acquisition and reconstruction protocols. Optimized CAC images should be acquired at a low tube voltage allowing for detection of small calcifications with lower densities and reconstructed with thin slices and iterative reconstruction allowing for more precise 3-dimensional CAC assessment to improve reproducibility. A low-dose optimized acquisition and reconstruction method combined with advanced quantification will result in improved reproducibility and allow for evaluation of CAC shape, density and distribution along the coronary artery tree at a lower radiation dose exposure.

INTRODUCTION

Coronary artery disease causes approximately one in every six deaths in the United States, making it the leading cause of death in the US and in other developed countries.[1] Cardiac risk assessment and primary prevention are thus very important. The amount of coronary artery calcium (CAC) is a strong predictor for future cardiovascular events.[2] Clinically, CAC scoring is used to reclassify individuals at intermediate risk into more appropriate risk categories.[3] For example, absence of CAC indicates a subject can be reclassified to low risk, and large amounts of calcium indicate high risk for future cardiac events. Traditionally CAC is quantified using the Agatston score, which can be easily quantified and interpreted within a couple of minutes.[4] While it showed an independent prognostic value in large populations resulting in robust reference score tables, this time-tested method has important limitations.[5] Optimized image acquisition and reconstruction combined with improved quantification can be more robust than traditional Agatston scoring,[6,7] with better reproducibility due to thin-slice reconstructions,[8,9] iterative reconstruction,[10] and improved quantification of low-density calcifications enabled by low tube-voltage acquisitions.[11] This may also result in a reduced radiation dose. These optimizations will most likely improve quantification of zero scores, low positive, or intermediate scores. Moreover, it will allow for evaluation of distribution and shape of calcifications, which may allow refining risk estimates.[12] Machine learning algorithms may be helpful for these evaluations.

IMAGE ACQUISITION

Electron beam tomography

The Agatston score was first published in the early 1990s.[4] At that time the gantry rotation times of mechanically rotating 3rd generation computed tomography (CT) scanners were not fast enough to adequately image a rapidly moving organ, such as the heart. Agatston and colleagues used an electron beam tomography (EBT) system. This scanner had no rotating mechanical elements and with exposure times as short as 50 ms allowed for motionless imaging of the coronary arteries. EBT had limitations pertaining to tube power, slice thickness, and spatial resolution. The X-ray source could only be operated at a single energy level of 130 kVp. Coronary lesions with a density of 130 Hounsfield units (HU) or higher were included in the Agatston score. The area of each calcified lesion (in mm²) was multiplied by an arbitrary chosen factor related to the peak attenuation: 130–199 HU, factor 1;

200–299 HU, factor 2; 300–399 HU, factor 3; and ≥400 HU, factor 4. The final Agatston score was the weighted sum of scores for all lesions. The most important shortcoming for CAC quantification was the limited through-plane resolution: minimum scan range per heartbeat was 3mm with the first EBT systems.[13] This resulted in images with 3mm thick slices, which is large relative to the size of the coronary arteries (the diameter of epicardial coronary arteries ranges from approximately 5mm proximal to submillimeter and smaller in distal branches). Due to the limited spatial resolution, calcified structures with low density and most lesions with a smaller area than 1mm2 could not be detected in the Agatston score. Commonly, Agatston scores are classified into one of the following four risk categories: no (0), mild (1–99), moderate (100–299), or severe (≥300), respectively.[14] These categories result in different preventive measures, therefore changes in these classifications may result in differences in clinical outcome.

Computed Tomography

The evolution of contemporary 3-rd generation CT scanners evolved quickly with the introduction of slip-ring technology and multi-detector CT scanners that combine fast gantry rotation with electrocardiography triggering and dedicated reconstruction algorithms allowing high-resolution volumetric cardiac imaging. The gantry rotation speed of state- of-the-art CT scanners allows a temporal resolution between 65 ms and 130 ms, and coverage of the whole heart within a single or few heartbeats. Modern CT scanners also generate improved image quality at reduced radiation exposure due to optimized detector and image reconstruction technology.[15] Modern detector systems and powerful X-ray tubes and generators of current-generation CT scanners allow the routine acquisition of images with a slice thickness smaller than 1mm at the isocenter (*Figure 1*). Increased tube power reserves also enable imaging at lower than the traditional 120 kVp tube voltages. This enables improved 3-dimensional evaluation of CAC and theoretically allows more precise and robust quantification. Despite these technical advances, the original scoring method for CAC has not been substantially updated since the 1990s.

While several alternative methods to the Agatston scoring method have been proposed, none have been incorporated into routine clinical practice or current guidelines, which retain the traditional Agatston score.[16,17] The Agatston score has been used in all major large cohort studies, resulting in an abundance of prognostic data for the Agatston score, and a lack of data for newer, potentially more sensitive, robust, and dose efficient scoring methods.[2,3]

Thick slices

3.0 mm slice thickness

Thin slices

0.6 mm slice thickness

Small calcification
Large calcification
Artery wall
Plaque

Small calcification   Large calcification

Small calcification   Large calcification

Undetected small calcification

Small calcification is detected

6
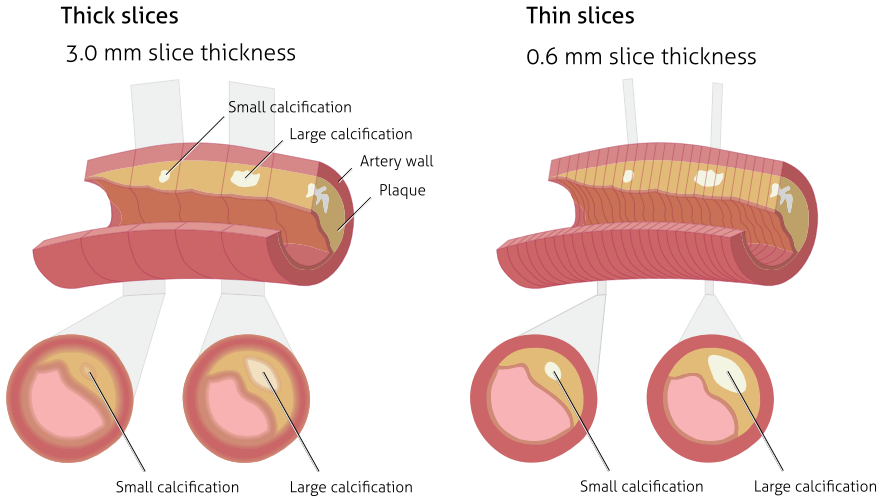
Figure 1 Partial Volume. Thick slice reconstructions (left) result in blurry images with less detail and small calcifications remain undetected. Thin slice reconstructions (right) result in sharp images with more detail and small calcifications can be detected. This results in more adequate quantification of coronary artery calcium and allows for newer quantification methods such as calcium volume and mass.

LIMITATIONS OF THE AGATSTON SCORE

Image acquisition and reconstruction: impact on image quality, noise, and radiation dose

A major limitation of the Agatston score is that it is not conducive to modern methods of radiation dose reduction (*Table 1*). The radiation dose of coronary CT angiography has decreased tremendously over the last years with the use of prospective electrocardiography (ECG) triggering, tube current modulation, low kVp acquisitions, fast (high-pitch) spiral acquisitions, and iterative reconstruction.[18] Radiation doses of CT exams for CAC quantification have also decreased, but to a lesser extend.[19] The most important methods to reduce the CT radiation dose include lowering the voltage (kVp) and current (mA) of the X-ray tube of the CT scanner. While this would result in noisier images if the same mA was used, this can be compensated for by increasing tube current combined with using iterative reconstruction algorithms that reduce image noise.[20,21] Since the Agatston score is based on counting picture elements (pixels), weighted by the peak CT attenuation (in Hounsfield units), any acquisition or reconstruction parameter affecting the dimensions of a pixel (or voxel – volume element; i.e. pixel x section thickness),

Table 1 Factors influencing the Agatston score

Factor	Disadvantage
Slice thickness (3 mm)	Small and low-density calcifications may not reach 130 Hounsfield unit threshold due to partial volume effect. Reproducibility decreases with thicker slices. Less suitable for 3D volumetric evaluation than thin slice images.
Tube voltage (120 kVp)	Not as suitable as low kVp imaging for radiation dose reduction. Detectability of low-density calcifications is reduced compared to low tube voltage imaging.
Peak density within lesion	Affected by noise. If the noise is high, the maximum attenuation of a lesion increases, whereas the average attenuation stays the same. This prohibits the use of low dose acquisitions and reconstruction of images with thin slices and iterative reconstruction.
Thresholds (130 / 200 / 300 / 400 Hounsfield Units)	Arbitrary chosen and systematic research is needed to find more optimal thresholds.
Summarizing lesions	Equal scores for differently located lesions does not allow for evaluation of distribution of calcifications throughout the coronary tree.

and any changes affecting the CT attenuation values will alter the Agatston score for a given patient. Iterative reconstruction algorithms, however, should be used with caution since they reduce noise, which affects Agatston scores resulting in misclassification rates ranging from 3% up to 31%.[22–24] Reduced kVp scanning is a powerful dose reduction method but it increases the CT attenuation of calcium. Reduced kVp scanning combined with iterative reconstruction has not been validated in large prospective multicenter trials and impact on patient management and outcomes. The 2017 CAC expert consensus from the Society of Cardiovascular Computed Tomography states that iterative reconstruction combined with 100 kVp acquisition may be used as an alternative to FBP with 120 kVp, however this should be utilized with caution after site-based or literature-based validation for each scanner vendor.[14] Early studies on radiation dose reduction mainly evaluated correlation between routinely used radiation dose scans and reduced radiation dose scans and found excellent correlations. However, recent studies have shown misclassification rates ranging from 15% up to 29% after 80% dose reduction.[10,25] This is mainly due to the dependency of the Agatston score on the peak density of calcified plaques, which is substantially influenced by noise, thus resulting in different Agatston scores.

Sensitivity

The second limitation of the Agatston score is that it is based on a relatively large section thickness (for today's standards) of 3 mm. The resulting poor spatial resolution, limits the detection and quantification of small calcifications due to partial volume effects. Smaller and less dense calcifications may be underestimated or completely missed depending on the relative location within a 3 mm thick slice (*Figure 2*). We evaluated 1871 clinical CAC CT scans acquired at our institution over the last 5 years (unpublished data). In total 959 scans (51.3%) showed a zero Agatston score. Additional thin slices were routinely reconstructed to evaluate lung parenchyma in the majority of these scans (N=826/ 959, 86.1%). Despite zero Agatston scores, we found small coronary calcifications in 118/826 patients (14.3%) on thin slices. This number is similar to a study conducted by Urabe et al., who detected coronary calcifications in 18/132 patients (13.6%) with a zero Agatston score.[26] Aslam and colleagues detected coronary calcifications in 24 patients on 0.5 mm images of 48 patients with a zero Agatston score on 3mm images.[27] It is currently unclear if the detection of small calcification is prognostically important since current acquisition and reconstruction methods do not allow for evaluation of sub-threshold small calcifications. However, there is a dose-response curve for the quantity of calcium and the occurrence of adverse events. Patients with a low calcium score (1–10) are at higher risk compared to those with zero calcium.[28]

6

Figure 2 False-negative Calcium Scores. Calcium scoring CT scan of a 61-year old female acquired with 120 kVp and reconstructed with standard 3mm slice thickness (**A**) and 0.6 mm thin slices (**B**). The calcification proximal in the left anterior descending artery did not reach the 130 Hounsfield unit threshold on the 3mm image resulting in a false-negative score, while it did reach the threshold on the 0.6 mm image.

While the entire cohort of zero calcium is at low risk, those events that do happen may do so in patients with calcifications just below the threshold of detection. Improved spatial resolution and sensitivity to smaller lesions is also desirable for studying higher order information about calcified plaque: it is very well possible that not only the total amount of coronary calcium, but also the shape and the density of individual lesions, and their distribution within the coronary arterial tree convey additional or modulate prognostic information based on density alone. Another factor that limits the detection and quantification of small and less dense calcifications is the threshold of 130 HU used for CAC quantification. This threshold is arbitrary and systematic research is needed to find a more optimal threshold.

6

Reproducibility

While the Agatston score is a successful test since it is relatively easy to obtain, the third limitation is its inherently suboptimal reproducibility. Since scoring of individual lesions is based on the maximum pixel attenuation with an arbitrary threshold of 130 HU, the method is vulnerable to image noise and beam hardening. Standard acquisition parameters for modern day CAC scoring across different CT manufacturers were proposed in a consensus paper by McCollough et al., in 2007.[29] Similar CT scanner settings should result in similar Agatston scores. In 2014, however, Willemink and colleagues showed that clinically used vendor recommended CT scanning protocols resulted in substantially different Agatston scores between CT manufacturers.[30] Median Agatston scores of 14 ex-vivo hearts for the CT system with the highest scores were 469.0 (182.8–1381.0) compared to 332.1 (114.3–1134.6) for the CT system with the lowest scores. The median relative difference between the same systems was 43.9 (21.9–55.1). If an individual undergoes a CAC CT scan at one site, he or she could be classified differently compared to another site with a different CT scanner; this was the case in up to 6.5% of subjects. Treatment recommendations for the same patient could thus differ between sites. Several studies have reported differences in CAC scoring between CT systems[19,30,31] and high inter-scan variability of Agatston scoring.[6,7,32–34] Rutten and colleagues showed that a small variation in patient position in the scanner leads to different Agatston scores with median absolute percentage differences ranging from 147 (0–200) in patients with low Agatston scores (ranging from 1 to 10) to 7 (4–22) in patients with high Agatston scores (> 400).[33] This resulted in potentially misclassification of approximately 10% of patients. The lack of reproducibility is multifactorial including motion artifacts and volume averaging. Another

important factor limiting the reproducibility is again the conventional use of 3mm slices for Agatston scoring. Simply reducing the slice thickness alone, however, will also affect Agatston scores.[8,9,27,35–38] With current acquisition protocols, CAC scores also differ between larger and smaller patients.[39,40] Compared to EBT – for which CAC quantification was originally developed – modern CT scanners systematically underestimate CAC scores, which can be explained by overall higher noise levels in EBT due to limited photon flux.[41,42]

From quantification to characterization

The fourth shortcoming of the Agatston score is the equal weighting of calcifications in different coronary artery segments. The Agatston score is calculated regardless of the CAC location or distribution within the coronary artery tree (*Figure 3*). However, distribution – such as proximal versus distal– of CAC may influence the prognostic value of CAC quantification.[43] Moreover, the Agatston score is based on the assumption that both higher volume and higher density of CAC are as- sociated with increased cardiovascular risk. However, Criqui et al. showed that CAC density was inversely associated with cardiovascular risk.[44] Their study also showed that CAC volume was positively and independently associated with cardiovascular risk. Therefore, the prognostic value of CAC might be improved if
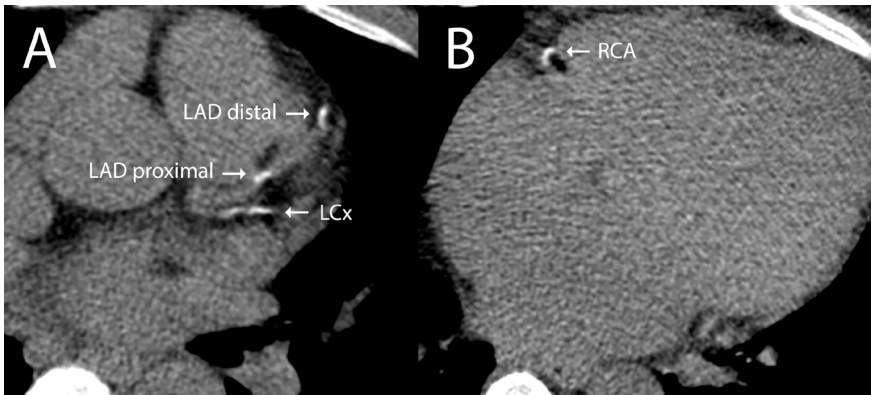
Figure 3 Distribution of Calcifications. Calcium scoring CT scan of a 49-year old male with a total Agatston score of 265. The proximal and distal calcifications in the left anterior descending artery (LAD), as well as the calcifications proximally in the circumflex (LCx), and distally in the right coronary artery (RCA) contributed equally to the total risk score, whereas some of these calcifications may be prognostically more important.

scores accounted for differences in higher CAC volumes and lower CAC densities. Furthermore, current dual-energy CT and future photon-counting CT systems may allow for improved quantification of CAC densities due to mate- rial-decomposition imaging capabilities and high spatial resolution imaging for the latter technique.[45]

ALTERNATIVE SCORING METHODS

The limitations of the Agatston score have not gone unnoticed, and several alternative CAC quantification methods have been proposed, including calcium volume and mass scores.[16,17] Three-dimensional data allow the evaluation of the volume of calcifications in cubic millimeters. With appropriate calibration, facilitated by a calibration phantom CT table insert, the actual mass of CAC can be expressed in milligrams. For feasibility reasons, volume and mass scores are developed for the same CT acquisition settings as the Agatston score. Therefore, these methods are still affected by the same partial volume effects as the Agatston score itself. An advantage of both, volume and mass scoring methods is the linearity of CAC measurements, which in contrast to the ordinal Agatston score allows for more precise quantification. Another advantage is that both measures are not substantially affected by small variations in image noise. The volume score is the only score that is not weighted by density, which may be beneficial considering the equivocal association between CAC density and cardiovascular risk. Disadvantages of the mass score include the complicated acquisition method since a calibration phantom is needed for every exam. Also, CAC mass is underestimated on CT systems from all major vendors and results differ per vendor, worsening the inter-scanner re- producibility.[19,39,42] With these methods, calcifications with areas smaller than $1mm^2$ can be quantified in a more reproducible manner. Both volume and mass scores improve the reproducibility compared to Agatston scores.[6,35,38,46,47] However, to this date, these quantification methods are not used in a clinical and screening setting due to the lack of prognostic data and the need for phantom calibration.

SUGGESTED SOLUTIONS

An updated scoring method should entail a modification of CAC acquisition, reconstruction and quantification, which addresses the limits of reproducibility, spatial resolution, and allows taking advantage of modern radiation dose reduction techniques. First, the X-ray tube voltage should be reduced (e.g. to 100 kVp), which

6

may not only result in lower radiation doses, but also increases the CT attenuation of less dense calcifications, allowing for improved detection and quantification of low-density CAC. Second, images should be reconstructed with thin slices (1 mm or less) resulting in isotropic images, allowing for precise three-dimensional CAC quantification, which may improve re- producibility. Reducing slice thickness and X-ray tube voltage will result in noisier images, particularly in larger patients; third, we re- commend using noise reducing reconstruction methods, such as iterative reconstruction, to reduce image noise. Fourth, volumetric CAC scores should be quantified taking into account both volume and density (calibrated X-ray attenuation) of calcified plaque. Last, these higher quality datasets may allow refining risk estimates in the future by assessing the distribution of calcifications throughout the coronary tree, allowing for differentiation for example between diffuse and concentrated calcifications or proximal and distal calcifications. Machine learning algorithms may be helpful for these evaluations. Since the Agatston score has already shown to be a strong and independent predictor of cardiovascular risk, an updated quantification method may not improve the predictive value substantially. Potentially, adding CAC distribution and shape may improve the predictive value. Also, using an inverse relation of CAC density and cardiovascular risk may be of incremental value. CAC distribution, shape, and inverse density should be evaluated extensively in multiple cohorts before implementation in a new CAC scoring method. Replacing the Agatston score would require an update of current CAC reference values. This can potentially be done by retrospective recalculation of large CAC studies, such as the MESA database, or if not possible a new prospective study should be set up. Future research is thus indicated before the CAC quantification method can be improved. Independent of the added value of these features, the most important aim of a new score would be to allow for radiation dose reduction and improved reproducibility without compromised predictive value. Recently, evaluation of progression of CAC has gained interest.[48] Improving reproducibility of CAC quantification will inherently result in improved evaluation of CAC progression. One should keep in mind that reductions in tube voltage and slice thickness and the application of iterative reconstruction will affect CAC attenuation and noise. Moreover, iterative reconstruction algorithms from different vendors result in different noise levels.[49] This may affect the ease CAC score of interpretation. Ideally, CAC attenuation and noise levels should be similar between CT systems from different vendors. This could be achieved by setting a single allowed slice thickness and a range of allowed noise levels and CAC attenuation using a standardized commercially available

phantom. Another approach could be to determine conversion factors for every CT system by using a standardized commercially available phantom. This should be done for both small- sized patients and large patients since body size influences CAC attenuation and image noise.

CONCLUSION

The CAC quantification method has not been updated substantially since its introduction on EBT scanners in the early 1990s. In the meantime modern multiple-detector row CT systems have replaced EBT technology and current evidence indicates that the traditional Agatston score lacks in reproducibility resulting in subjects being incorrectly risk classified. Recent studies indicate that low-density CAC is associated with a higher cardiovascular risk and that the number, distribution and shape of CAC may be of prognostic value as well. The effects of these limitations on outcome prediction are not evaluated yet in large cohorts, however, we expect that more robust CAC quantification will enhance cardiovascular risk quantification. Therefore, we suggest optimizing CAC quantification for modern CT scanners using lower radiation doses and higher spatial resolution. Optimized images should be acquired at a low tube voltage allowing for detection of small and lower density calcifications and thin slices are reconstructed with iterative reconstruction allowing for more precise 3-dimensional CAC assessment to improve reproducibility. Lower tube voltage also decreases radiation exposure. Further research is needed to derive optimized CT acquisition and CAC quantification methods. Once optimized CAC quantification methods are developed, they should be evaluated in large observational studies with long follow-up time. We expect that optimized protocols will result in improved reproducibility and will allow for evaluation of CAC density, shape and distribution within the coronary artery tree at a reduced radiation dose. In summary, the Agatston score has important limitations, including suboptimal reproducibility, spatial resolution and does not account for new image acquisition and reconstruction modes for current MDCT scanners. Therefore, we propose image acquisition, reconstruction and means to quantify coronary artery calcium should be optimized.

REFERENCES

1. Greenland P, Alpert JS, Beller GA, et al. 2010 ACCF/AHA guideline for assessment of cardiovascular risk in asymptomatic adults: Executive summary: A report of the American College of cardiology foundation/American Heart association task force on practice guidelines. Circulation. 2010;122(25):2748-2764. doi:10.1161/CIR.0b013e3182051bab

2. Detrano R, Guerci AD, Carr JJ, et al. Coronary calcium as a predictor of coronary events in four racial or ethnic groups. N Engl J Med. 2008;358(13):1336-1345. doi:10.1056/NEJMoa072100

3. Elias-Smale SE, Proença RV, Koller MT, et al. Coronary calcium score improves classification of coronary heart disease risk in the elderly: The Rotterdam study. J Am Coll Cardiol. 2010;56(17):1407-1414. doi:10.1016/j.jacc.2010.06.029

4. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. J Am Coll Cardiol. 1990;15(4):827-832. doi:10.1016/0735-1097(90)90282-T

5. Blaha MJ, Mortensen MB, Kianoush S, Tota-Maharaj R, Cainzos-Achirica M. Coronary Artery Calcium Scoring: Is It Time for a Change in Methodology? JACC Cardiovasc Imaging. Published online 2017. doi:10.1016/j.jcmg.2017.05.007

6. Hoffmann U, Siebert U, Bull-Stewart A, et al. Evidence for lower variability of coronary artery calcium mineral mass measurements by multi-detector computed tomography in a community-based cohort--consequences for progression studies. Eur J Radiol. 2006;57(3):396-402. doi:10.1016/j.ejrad.2005.12.027

7. Yamamoto H, Budoff MJ, Lu B, Takasu J, Oudiz RJ, Mao S. Reproducibility of three different scoring systems for measurement of coronary calcium. Int J Cardiovasc Imaging. 2002;18(5):391-397. doi:10.1023/a:1016051606758

8. Horiguchi J, Matsuura N, Yamamoto H, et al. Variability of repeated coronary artery calcium measurements by 1.25-mm- and 2.5-mm-thickness images on prospective electrocardiograph-triggered 64-slice CT. Eur Radiol. 2008;18(2):209-216. doi:10.1007/s00330-007-0734-7

9. Achenbach S, Meissner F, Ropers D, et al. Overlapping cross-sections significantly improve the reproducibility of coronary calcium measurements by electron beam tomography: a phantom study. J Comput Assist Tomogr. 2001;25(4):569-573. doi:10.1097/00004728-200107000-00010

10. Willemink MJ, den Harder AM, Foppen W, et al. Finding the optimal dose reduction and iterative reconstruction level for coronary calcium scoring. J Cardiovasc Comput Tomogr. 2016;10(1):69-75. doi:10.1016/j.jcct.2015.08.004

11. Vonder M, van der Werf NR, Leiner T, et al. The impact of dose reduction on the quantification of coronary artery calcifications and risk categorization: A systematic review. J Cardiovasc Comput Tomogr. 2018;12(5):352-363. doi:10.1016/j.jcct.2018.06.001

12. Blaha MJ, Budoff MJ, Tota-Maharaj R, et al. Improving the CAC Score by Addition of Regional Measures of Calcium Distribution: Multi-Ethnic Study of Atherosclerosis. JACC Cardiovasc Imaging. 2016;9(12):1407-1416. doi:10.1016/j.jcmg.2016.03.001

13. Nieman K. Evolve or perish for coronary calcium imaging. Eur Heart J Cardiovasc Imaging. 2015;16(4):354-355. doi:10.1093/ehjci/jeu220
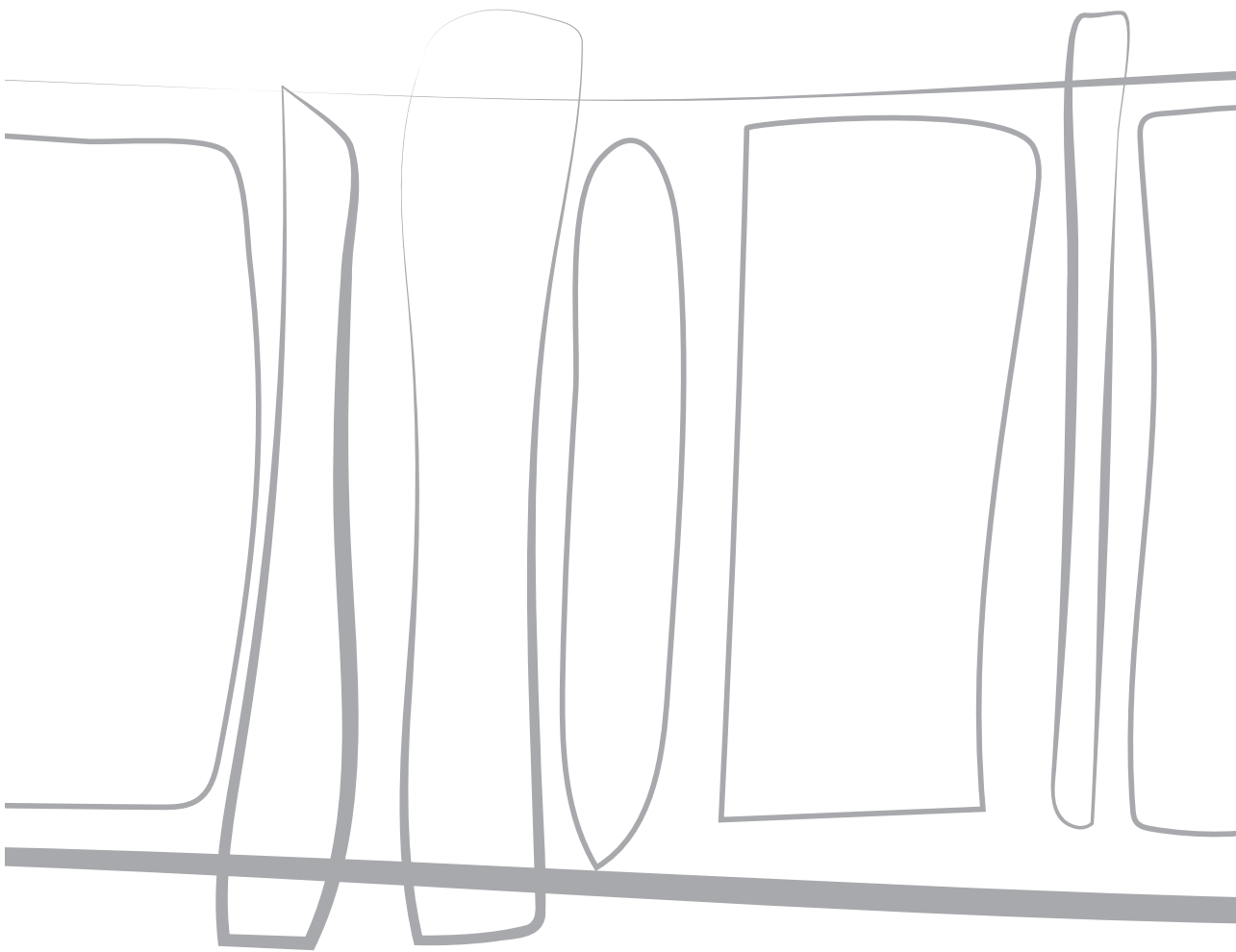
6

14. Hecht H, Blaha MJ, Berman DS, et al. Clinical indications for coronary artery calcium scoring in asymptomatic patients: Expert consensus statement from the Society of Cardiovascular Computed Tomography. J Cardiovasc Comput Tomogr. 2017;11(2):157-168. doi:10.1016/j.jcct.2017.02.010

15. den Harder AM, Willemink MJ, de Jong PA, et al. New horizons in cardiac CT. Clin Radiol. 2016;71(8):758-767. doi:10.1016/j.crad.2016.01.022

16. Detrano RC, Anderson M, Nelson J, et al. Coronary calcium measurements: effect of CT scanner type and calcium measure on rescan reproducibility--MESA study. Radiology. 2005;236(2):477-484. doi:10.1148/radiol.2362040513

17. Hong C, Bae KT, Pilgram TK. Coronary Artery Calcium: Accuracy and Reproducibility of Measurements with Multi–Detector Row CT—Assessment of Effects of Different Thresholds and Quantification Methods. Radiology. 2003;227(3):795-801. doi:10.1148/radiol.2273020369

18. den Harder AM, Willemink MJ, Bleys RL a W, et al. Dose reduction for coronary calcium scoring with hybrid and model-based iterative reconstruction: an ex vivo study. Int J Cardiovasc Imaging. 2014;30(6):1125-1133. doi:10.1007/s10554-014-0434-8

19. van der Werf NR, Willemink MJ, Willems TP, Greuter MJW, Leiner T. Influence of dose reduction and iterative reconstruction on CT calcium scores: a multi-manufacturer dynamic phantom study. Int J Cardiovasc Imaging. 2017;33(6):899-914. doi:10.1007/s10554-017-1061-y

20. Willemink MJ, de Jong P a, Leiner T, et al. Iterative reconstruction techniques for computed tomography Part 1: Technical principles. Eur Radiol. Published online January 2013. doi:10.1007/s00330-012-2765-y

21. Willemink MJ, Leiner T, De Jong PA, et al. Iterative reconstruction techniques for computed tomography part 2: Initial results in dose reduction and image quality. Eur Radiol. 2013;23(6):1632-1642. doi:10.1007/s00330-012-2764-z

22. Kurata A, Dharampal A, Dedic A, et al. Impact of iterative reconstruction on CT coronary calcium quantification. Eur Radiol. 2013;23(12):3246-3252. doi:10.1007/s00330-013-3022-8

23. Schindler A, Vliegenthart R, Schoepf UJ, et al. Iterative Image Reconstruction Techniques for CT Coronary Artery Calcium Quantification: Comparison with Traditional Filtered Back Projection in Vitro and in Vivo. Radiology. 2014;270(2):387-393. doi:10.1148/radiol.13130233

24. Van Osch JAC, Mouden M, Van Dalen JA, et al. Influence of iterative image reconstruction on CT-based calcium score measurements. Int J Cardiovasc Imaging. 2014;30(5):961-967. doi:10.1007/s10554-014-0409-9

25. den Harder AM, Wolterink JM, Willemink MJ, et al. Submillisievert coronary calcium quantification using model-based iterative reconstruction: A within-patient analysis. Eur J Radiol. 2016;85(11):2152-2159. doi:10.1016/j.ejrad.2016.09.028

26. Urabe Y, Yamamoto H, Kitagawa T, et al. Identifying Small Coronary Calcification in Non-Contrast 0.5-mm Slice Reconstruction to Diagnose Coronary Artery Disease in Patients with a Conventional Zero Coronary Artery Calcium Score. J Atheroscler Thromb. Published online 2016:1-10. doi:10.5551/jat.35634

27. Aslam A, Khokhar US, Chaudhry A, et al. Assessment of isotropic calcium using 0.5-mm reconstructions from 320-row CT data sets identifies more patients with non-zero Agatston score and more subclinical atherosclerosis than standard 3.0-mm coronary artery calcium scan and CT angiography. J Cardiovasc Comput Tomogr. 2014;8(1):58-66. doi:10.1016/j.jcct.2013.12.007

6

28. Blaha M, Budoff MJ, Shaw LJ, et al. Absence of Coronary Artery Calcification and All-Cause Mortality. JACC Cardiovasc Imaging. 2009;2(6):692-700. doi:10.1016/j.jcmg.2009.03.009

29. McCollough CH, Ulzheimer S, Halliburton SS, Shanneik K, White RD, Kalender WA. Coronary Artery Calcium: A Multi-institutional, Multimanufacturer International Standard for Quantification at Cardiac CT. Radiology. 2007;243(2):527-538. doi:10.1148/radiol.2432050808

30. Willemink MJ, Vliegenthart R, Takx RAP, et al. Coronary Artery Calcification Scoring with State-of-the-Art CT Scanners from Different Vendors Has Substantial Effect on Risk Classification. Radiology. 2014;273(3):695-702. doi:10.1148/radiol.14140066

31. Greuter MJW, Groen JM, Nicolai LJ, Dijkstra H, Oudkerk M. A model for quantitative correction of coronary calcium scores on multidetector, dual source, and electron beam computed tomography for influences of linear motion, calcification density, and temporal resolution: A cardiac phantom study. Med Phys. 2009;36(11):5079-5088. doi:10.1118/1.3213536

32. Ulzheimer S, Kalender WA. Assessment of calcium scoring performance in cardiac computed tomography. Eur Radiol. 2003;13(3):484-497. doi:10.1007/s00330-002-1746-y

33. Rutten A, Isgum I, Prokop M. Coronary calcification: effect of small variation of scan starting position on Agatston, volume, and mass scores. Radiology. 2008;246(1):90-98. doi:10.1148/radiol.2461070006

34. Lu B, Budoff MJ, Zhuang N, et al. Causes of interscan variability of coronary artery calcium measurements at electron-beam CT. Acad Radiol. 2002;9(6):654-661. doi:10.1016/s1076-6332(03)80310-0

35. Groen JM, Greuter MJ, Schmidt B, Suess C, Vliegenthart R, Oudkerk M. The influence of heart rate, slice thickness, and calcification density on calcium scores using 64-slice multidetector computed tomography: A systematic phantom study. Invest Radiol. 2007;42(12):848-855. doi:10.1097/RLI.0b013e318154c549

36. Muhlenbruch G, Thomas C, Wildberger JE, et al. Effect of varying slice thickness on coronary calcium scoring with multislice computed tomography in vitro and in vivo. Invest Radiol. 2005;40(11):695-699.

37. Van Der Bijl N, De Bruin PW, Geleijns J, et al. Assessment of coronary artery calcium by using volumetric 320-row multi-detector computed tomography: Comparison of 0.5 mm with 3.0 mm slice reconstructions. Int J Cardiovasc Imaging. 2010;26(4):473-482. doi:10.1007/s10554-010-9581-8

38. Mahnken AH, Wildberger JE, Sinha AM, et al. Variation of the coronary calcium score depending on image reconstruction interval and scoring algorithm. Invest Radiol. 2002;37(9):496-502. doi:10.1097/00004424-200209000-00004

39. Willemink MJ, Abramiuc B, den Harder AM, et al. Coronary calcium scores are systematically underestimated at a large chest size: A multivendor phantom study. J Cardiovasc Comput Tomogr. 2015;9(5):415-421. doi:10.1016/j.jcct.2015.03.010

40. Sevrukov A, Pratap A, Doss C, Jelnin V, Hoff JA, Kondos GT. Electron beam tomography imaging of coronary calcium: the effect of body mass index on radiologic noise. J Comput Assist Tomogr. 2002;26(4):592-597. doi:10.1097/00004728-200207000-00021

41. Groen JM, Greuter MJW, Vliegenthart R, et al. Calcium scoring using 64-slice MDCT, dual source CT and EBT: A comparative phantom study. Int J Cardiovasc Imaging. 2008;24(5):547-556. doi:10.1007/s10554-007-9282-0

42. Greuter MJW, Dijkstra H, Groen JM, et al. 64 Slice MDCT generally underestimates coronary calcium scores as compared to EBT: A phantom study. Med Phys. 2007;34(9):3510-3519. doi:10.1118/1.2750733

43. Tota-Maharaj R, Al-Mallah MH, Nasir K, Qureshi WT, Blumenthal RS, Blaha MJ. Improving the relationship between coronary artery calcium score and coronary plaque burden: addition of regional measures of coronary artery calcium distribution. Atherosclerosis. 2015;238(1):126-131. doi:10.1016/j.atherosclerosis.2014.11.008

44. Criqui MH, Denenberg JO, Ix JH, et al. Calcium Density of Coronary Artery Plaque and Risk of Incident Cardiovascular Events. Jama. 2014;311(3):271. doi:10.1001/jama.2013.282535

45. Willemink MJ, Persson M, Pourmorteza A, Pelc NJ, Fleischmann D. Photon-counting CT: Technical Principles and Clinical Prospects. Radiology. 2018;289(2):293-312. doi:10.1148/radiol.2018172656

46. Callister TQ, Cooil B, Raya SP, Lippolis NJ, Russo DJ, Raggi P. Coronary artery disease: improved reproducibility of calcium scoring with an electron-beam CT volumetric method. Radiology. 1998;208(3):807-814. doi:10.1148/radiology.208.3.9722864

47. van Ooijen PMA, Vliegenthart R, Witteman JC, Oudkerk M. Influence of scoring parameter settings on Agatston and volume scores for coronary calcification. Eur Radiol. 2005;15(1):102-110. doi:10.1007/s00330-004-2479-x

48. Mori H, Torii S, Kutyna M, Sakamoto A, Finn A V, Virmani R. Coronary Artery Calcification and its Progression: What Does it Really Mean? JACC Cardiovasc Imaging. 2018;11(1):127-142. doi:10.1016/j.jcmg.2017.10.012

49. Willemink MJ, Jong PA De, Prokop M, Mey J De, Leiner T, Schilham AMR. Computed Tomography Radiation Dose Reduction: Effect of Different Iterative Reconstruction Algorithms on Image Quality. J Comput Assist Tomogr. Published online 2014.
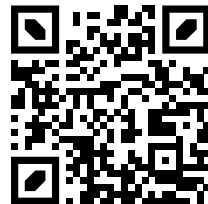
Online material is available via:

CHAPTER 7

Influence of iterative reconstruction on coronary calcium scores at multiple heart rates: a multivendor phantom study on state-of-the-art CT systems

Niels R. van der Werf, MSc

Martin J. Willemink, MD PhD

Tineke P. Willems, MD PhD

Marcel J.W. Greuter, PhD

Tim Leiner, MD PhD

ABSTRACT

Purpose

The objective of this study was to evaluate the influence of iterative reconstruction on coronary calcium scores (CCS) at different heart rates for four state-of-the-art CT systems.

Methods

Within an anthropomorphic chest phantom, artificial coronary arteries were translated in a water-filled compartment. The arteries contained three different calcifications with low (38 mg), medium (80 mg) and high (157 mg) mass. Linear velocities were applied, corresponding to heart rates of 0, <60, 60-75 and >75 bpm. Data were acquired on four state-of-the-art CT systems (CT1-CT4) with routinely used CCS protocols. Filtered back projection (FBP) and three increasing levels of iterative reconstruction (L1-L3) were used for reconstruction. CCS were quantified as Agatston score and mass score. An iterative reconstruction susceptibility (IRS) index was used to assess susceptibility of Agatston score (IRSAS) and mass score (IRSMS) to iterative reconstruction. IRS values were compared between CT systems and between calcification masses. For each heart rate, differences in CCS of iterative reconstructed images were evaluated with CCS of FBP images as reference, and indicated as small (<5%), medium (5-10%) or large (>10%). Statistical analysis was performed with repeated measures ANOVA tests.

Results

While subtle differences were found for Agatston scores of low mass calcification, medium and high mass calcifications showed increased CCS up to 77% with increasing heart rates. IRSAS of CT1-CT4 were 17%, 41%, 130% and 22% higher than IRSMS. Not only were IRS significantly different between all CT systems, but also between calcification masses. Up to a fourfold increase in IRS was found for the low mass calcification in comparison with the high mass calcification. With increasing iterative reconstruction strength, maximum decreases of 21% and 13% for Agatston and mass score were found. In total, 21 large differences between Agatston scores from FBP and iterative reconstruction were found, while only 5 large differences were found between FBP and iterative reconstruction mass scores.

Conclusion

Iterative reconstruction results in reduced CCS. The effect of iterative reconstruction on CCS is more prominent with low-density calcifications, high heart rates and increasing iterative reconstruction strength.

7

INTRODUCTION

Iterative reconstruction for computed tomography (CT) is a powerful technique which can be used to reduce CT radiation dose.[1–10] The effect of iterative reconstruction on quantitative measurements in cardiovascular imaging has been subject of recent studies.[11–18] One of the most commonly used cardiovascular risk assessment tools in CT is coronary calcium scoring (CCS). CCS can be expressed as Agatston score, which can be used to estimate the risk of future cardiovascular events.[19,20] In addition, the coronary calcium mass score has been introduced which is known to have a better stability under the influence of multiple parameters, including varying heart rates and high image noise.[21] According to the most recent guidelines from the American Heart Association (AHA), Agatston score risk categories can guide treatment strategies for asymptomatic adults.[22] Consequently, accurate determination of CCS is important to assign individual patients to correct risk categories.

Currently, CT acquisition is synchronized to the diastole resting period, where, with appropriate electrocardiograph (ECG) triggering, it is assumed that coronary arteries are imaged at relative rest. Nevertheless, the coronary arteries are known to move at velocities of up to 30 mm/s even during the rest phase, depending on both heart rate and coronary artery.[23,24] These velocities strongly influence the stability of CCS, together with other factors including calcification mass and density or imaging factors such as image noise and use of iterative reconstruction.[21,24–27] However, a systematic understanding of the mutual dependence of these factors and their influence on CCS is still lacking. We hypothesize that the influence of iterative reconstruction at different heart rates will differ between different calcification densities. Therefore, the aim of this study was to evaluate the influence of iterative reconstruction on CCS for various calcifications at different heart rates using four state-of-the-art CT systems.

MATERIALS AND METHODS

Within an anthropomorphic chest phantom (QRM-Chest, QRM, Moehrendorf, Germany), artificial coronary arteries were moved in a water-filled compartment (*Figure 1*). The artificial coronary arteries contained hydroxyapatite (HA) calcifications of low (38.5±1.7 mg), medium (80.1±3.3 mg) and high (157.1±6.5 mg) mass. The cylindrical calcifications were 5.0±0.1 mm in diameter, with a length of 10.0±0.1 mm resulting in densities of 196±3, 408±2 and 800±2 HA/cm^3 for the low, medium and high mass calcification, respectively. The chest phantom contained
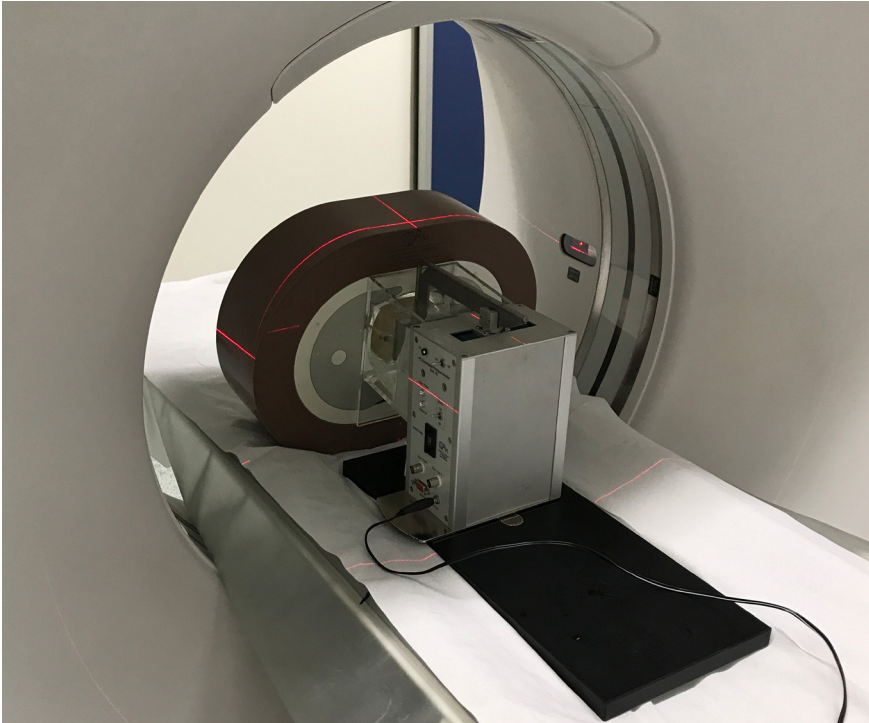
Figure 1 Anthropomorphic chest phantom with extension ring and motion controller (QRM, Moehrendorf, Germany). An artificial coronary artery moved in the horizontal plane inside the water tank in the center of the chest phantom

artificial lungs, a spine insert and a shell of soft tissue equivalent material. To mimic an averaged sized patient, an extension ring of fat equivalent material was used to increase the size of the phantom to 400 x 300 mm.[13]

The artificial coronary arteries were moved with a computer-controlled lever (QRM-Sim2D, QRM, Moehrendorf, Germany) at constant linear velocities of 0, 10, 20 and 30 mm/s in a horizontal plane perpendicular to the scan direction. The linear velocities used corresponded to heart rates of 0, <60, 60-75 and >75 bpm.[24]

Raw data were acquired with routinely used CCS protocols on four CT systems (CT1-CT4): Discovery CT 750 HD (GE Healthcare, Waukesha, Wisconsin, USA), Brilliance iCT (Philips Healthcare, Best, The Netherlands), Somatom Definition Flash (Siemens Healthcare, Forchheim, Germany) and Aquilion One Vision (Toshiba Medical Systems, Otawara, Japan), respectively (*Table 1*).

Table 1. Acquisition and reconstruction parameters used on CT system CT1 to CT4

CT system	CT1	CT2	CT3	CT4
Tube voltage [kV]	120	120	120	120
Tube current per rotation [mAs]	175	50	80	80
Collimation [mm]	64 x 0.625	128 x 0.625	128 x 0.6	320 x 0.5
Rotation time [s]	0.35	0.27	0.28	0.35
Temporal resolution* [ms]	175	135	75	175
Slice thickness [mm]	2.5	3.0	3.0	3.0
Increment [mm]	2.5	3.0	3.0	3.0
Kernel	Standard	XCA	B35f	FC12
Iterative reconstruction	ASIR	iDose[4]	SAFIRE	AIDR 3D
iterative reconstruction levels (L1-L3)	20, 60, 100%	1, 5, 7	1, 3, 5	weak, standard, strong
Calcium scoring software	Smartscore 4.0	Heartbeat-CS	Syngo	Vitrea FX 6.5.0
Noise level (HU)	26	22	28	24
Mass calibration factor	0.74	0.78	0.74	0.84

* As defined in the isocenter.

Acquisition was performed five times for each heart rate, with a small deviation and rotation between the scans by manual repositioning of the phantom. The ECG trigger output of the computer-controlled lever was used as an ECG trigger input of the CT scanner to ensure data acquisition during linear movement of the artificial artery. Acquisition was performed at 60% of the movement, so that turning points of the artificial coronary artery were not included.

FBP and (hybrid) iterative reconstruction were used to reconstruct raw data on each CT system (*Table 1*).[28] In addition, for each CT system, the lowest (L1), an intermediate (L2) and the highest (L3) available level of iterative reconstruction were used. CT-system specific noise levels in FBP images were assessed by calculating the standard deviation in the average Hounsfield value in a uniform water region (*Table 1*). CCS was quantified for each reconstruction as Agatston and mass score, with a default threshold of 130 Hounsfield Units (HU). CCS was assessed using the vendor specific software (*Table 1*), where one observer selected each calcification. A CT-specific calibration factor was used for the mass score, which was calculated as described previously.[29] For this, a dedicated, stationary, insert (CCI, QRM, Moehrendorf, Germany) for the QRM thorax phantom with a water cylinder was used. From the five measurements, mean CCS and standard deviation were calculated per calcification mass for each combination of CT system, heart rate and reconstruction.

A root mean square measure was used to quantify the susceptibility of CCS to iterative reconstruction for each CT system. The iterative reconstruction susceptibility (IRS) index was defined as:

$$IRS = \frac{1}{N-1} \left. \sqrt{\sum_{i=1}^{N}(x_0 - x_i)^2} \, \frac{1}{x_0} \right| \quad (1)$$

in which x_0 is the CCS at FBP, x_i is the CCS at iterative reconstruction level i and N is the total number of reconstructions. Low IRS indicates a low susceptibility of CCS to differences in reconstruction type. IRS values were compared between CT systems and between calcification masses for each CT system. For each heart rate, differences in CCS of iterative reconstructed images with respect to the reference of CCS of FBP images were calculated and indicated as small (<5%), medium (5-10%) or large (>10%).

Statistical analysis was performed with SPSS for Windows (IBM SPSS Statistics, Armonk, New York, USA), version 22.0, where a p-value smaller than 0.05 was used to determine significance. Normal distribution of the data was analyzed with the Shapiro-Wilk test. Significant differences between CCS of iterative reconstructed images and FBP images were assessed with repeated measures ANOVA tests, as well as differences in IRS.

RESULTS

Accuracy of CCS

The routinely used CCS protocols resulted in comparable noise levels in the reference CCS images for the used CT systems (*Table 1*). For the low mass calcification, only minor differences were shown between reference Agatston scores from different CT systems (*Figure 2*, *top row*). In comparison with reference Agatston scores at 0 bpm, the average differences with reference Agatston scores at >75 bpm were -18%, -9%, 5% and -13% for CT1-CT4, respectively. These differences in reference Agatston scores between 0 and >75 bpm were more pronounced for medium and high mass calcifications. For medium mass scores increased with 38%, 36%, 11% and 37% for CT1-CT4, respectively, while the high mass calcification even showed increases in reference Agatston score of 77%, 49%, 23% and 63%.

On all CT systems, reference mass scores generally underestimated physical mass (*Figure 2*, *bottom row*). For the low mass calcification, and averaged over all heart rates, mean reference mass scores were 28, 24, 20 and 26 for CT1-CT4, respectively.

For the medium mass calcification mean reference scores were 73, 76, 62 and 83 and for the high mass 134, 166, 131 and 178. For CT1-CT4, differences in reference mass scores between 0 and >75 bpm were -25%, -14%, -8% and -18%, respectively. Relative differences in reference CCS between 0 and >75 bpm for the medium and high mass calcification were smaller for mass scores than for Agatston scores. Medium mass calcification differences were -8%, 10% -2% and 1% for CT1-CT4, respectively, while for the high mass calcification differences of -18%, 13%, 4% and 10% were found .
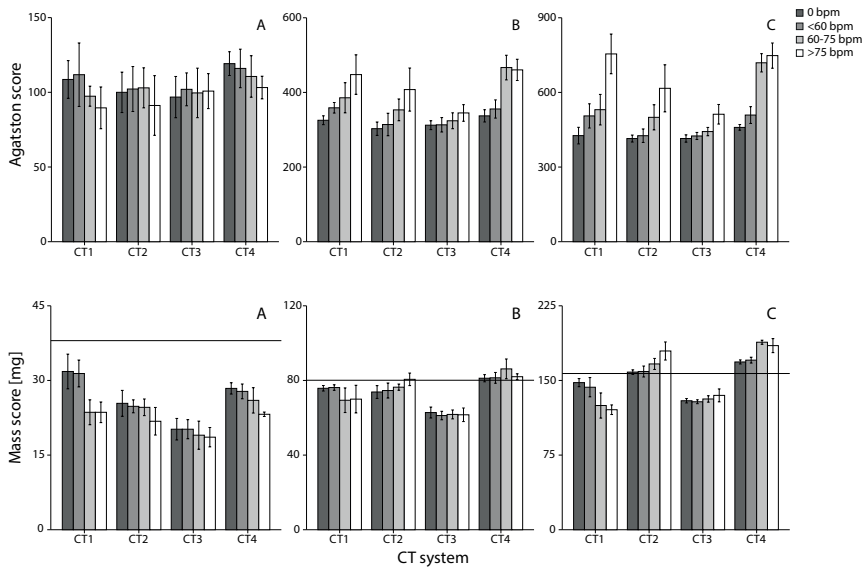
Figure 2 Mean and standard deviation reference Agatston score (top) and reference mass score (bottom) for all CT systems and heart rates for low (**A**), medium (**B**) and high (**C**) calcification mass. Physical mass, for the reference mass score graphs, is indicated with a solid line

Influence of reconstruction method on CCS

On average, Agatston scores were more susceptible to application of iterative reconstruction than mass scores, (*Figure 3*). Averaged over all heart rates and calcifications, IRS_{AS} for CT1-CT4 were 17%, 41%, 130% and 22% higher than IRS_{MS}. Both IRS_{AS} and IRS_{MS} were found to differ significantly between CT systems (Agatston score: p=0.001, mass score: p<0.000). Also, for each individual CT system, significant differences in IRS_{AS} and IRS_{MS} between all calcification masses were found. The IRS was significantly larger for low mass calcifications compared to medium and high mass calcifications.
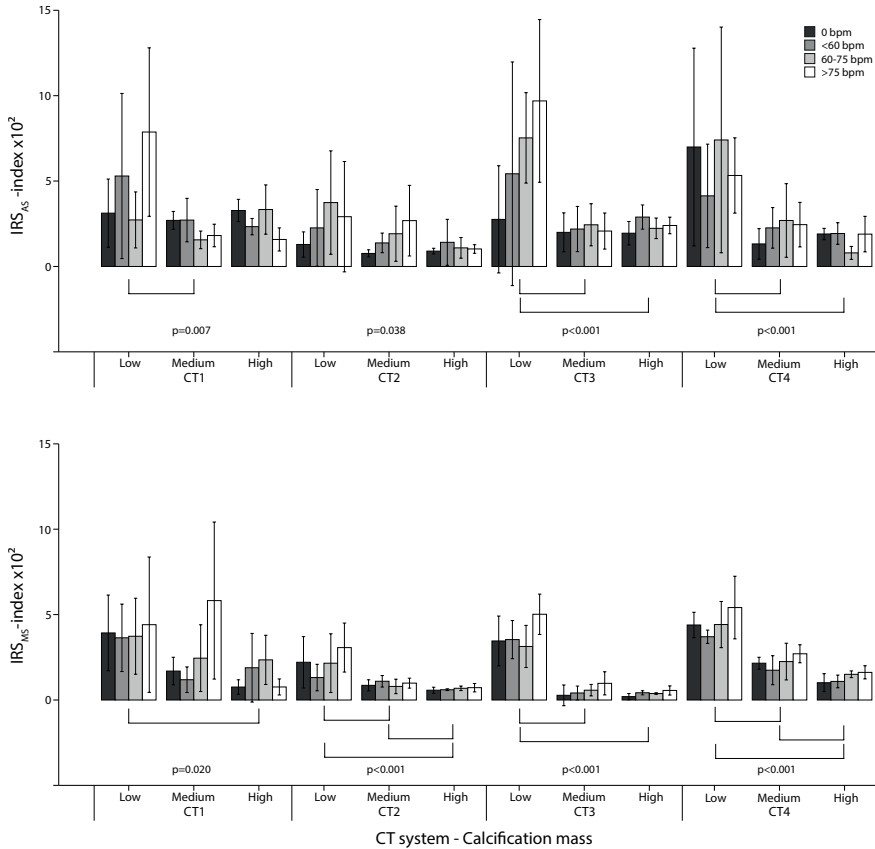
Figure 3 IRS$_{AS}$ (mean and SD) and IRS$_{MS}$ (bottom) for all CT systems at 0, <60, 60-75 and >75 bpm and for the low, medium and high mass calcification. Smaller IRS values indicate smaller susceptibility to differences in reconstruction type. P values below the bars indicate significant differences of IRS between calcification masses. If available, significantly different pairs of calcification masses are indicated with square brackets. A square bracket therefore points out that the influence of changes in reconstruction is significantly different for these two calcifications

For each reference Agatston score, differences between Agatston scores and their reference increased with increasing iterative reconstruction level (*Table 2*). These differences increased also with increasing heart rate and decreasing calcification mass. For almost all calcification masses and heart rates, significant differences were found between Agatston scores from different reconstructions. These differences resulted in large differences between iterative reconstruction and FBP Agatston scores for only the low mass calcifications. A maximum difference of -21% was found for CT3 at >75bpm and L3, where the average Agatston decreased from

Table 2. Mean percentage deviation of Agatston scores for each heart rate as a function of iterative reconstruction level (L1-L3) from the reference Agatston score at FBP.

CT	Heart rate [bpm]	Low calcification mass				Medium calcification mass				High calcification mass			
		L1	L2	L3	p	L1	L2	L3	p	L1	L2	L3	p
CT1	0	-3%	-5%	-7%	0.011	-2%	-4%	-7%	<0.001	-2%	-6%	-8%	<0.001
	<60	-5%	-8%	-13%	0.021	-1%	-3%	-7%	<0.001	-1%	-4%	-6%	<0.001
	60-75	0%	-4%	-7%	0.001	-1%	-2%	-4%	<0.001	-2%	-5%	-8%	<0.001
	>75	-3%	-14%	-19%	0.008	-1%	-3%	-4%	<0.001	-1%	-2%	-4%	<0.001
CT2	0	-1%	-2%	-2%	0.031	0%	-1%	-2%	<0.001	0%	-1%	-2%	<0.001
	<60	-1%	-2%	-6%	0.155	0%	-2%	-3%	<0.001	0%	-2%	-3%	0.018
	60-75	-4%	-7%	-7%	0.018	-1%	-3%	-5%	0.021	0%	-2%	-2%	0.001
	>75	-1%	-5%	-6%	0.185	0%	-4%	-7%	0.003	0%	-2%	-2%	<0.001
CT3	0	-1%	-5%	-6%	0.147	-2%	-3%	-4%	<0.001	-1%	-3%	-5%	<0.001
	<60	-3%	-7%	-13%	0.128	-1%	-4%	-5%	0.002	-2%	-5%	-7%	<0.001
	60-75	-1%	-11%	-19%	0.001	-1%	-4%	-6%	0.001	-1%	-3%	-6%	<0.001
	>75	-11%	-16%	-21%	0.001	-1%	-3%	-5%	0.001	-1%	-4%	-6%	<0.001
CT4	0	-10%	-13%	-13%	0.010	1%	0%	1%	0.539	2%	3%	4%	<0.001
	<60	-2%	-6%	-10%	0.024	4%	4%	0%	0.010	0%	0%	2%	0.211
	60-75	-6%	-10%	-17%	0.021	-2%	-3%	-6%	0.047	-1%	-1%	-1%	0.263
	>75	-2%	-7%	-13%	0.007	0%	-3%	-6%	0.009	-3%	-3%	-2%	0.230

Indicated are differences for low, medium and high calcification mass. Small (<5%), medium (5-10%) and large (>10%) differences are denoted by respectively green, orange and red colored cells. Statistical results of the repeated measures ANOVA test are indicated, with significant p-values in bold

7

Table 3. Mean percentage deviation of mass scores for each heart rate as a function of iterative reconstruction level (L1-L3) from the reference mass score at FBP.

CT	Heart rate [bpm]	Low calcification mass				Medium calcification mass				High calcification mass			
		L1	L2	L3	p	L1	L2	L3	p	L1	L2	L3	p
CT1	0	-2%	-1%	-7%	0.102	-1%	-2%	-2%	0.369	0%	-1%	0%	0.351
	<60	-4%	-6%	-5%	0.198	0%	0%	-2%	0.496	1%	2%	4%	0.050
	60-75	-1%	-5%	-10%	<0.001	0%	4%	1%	0.405	0%	3%	6%	0.007
	>75	-1%	-12%	-3%	0.016	7%	11%	10%	0.008	-1%	-1%	-1%	0.545
CT2	0	-2%	-3%	-5%	0.004	-1%	-1%	-2%	0.006	-1%	-1%	-1%	<0.001
	<60	0%	-2%	-3%	0.006	-1%	-2%	-2%	<0.001	-1%	-1%	-1%	<0.001
	60-75	-2%	-3%	-5%	0.028	-1%	-1%	-2%	0.038	-1%	-1%	-1%	<0.001
	>75	-4%	-6%	-6%	<0.001	-1%	-2%	-2%	<0.001	0%	-1%	-2%	<0.001
CT3	0	-3%	-6%	-7%	0.014	0%	0%	-1%	0.426	0%	0%	0%	0.873
	<60	-3%	-6%	-8%	<0.001	0%	0%	-1%	0.321	0%	-1%	-1%	<0.001
	60-75	-1%	-3%	-8%	<0.001	0%	-1%	-1%	0.089	0%	-1%	-1%	<0.001
	>75	-4%	-8%	-12%	<0.001	-1%	-2%	-2%	0.003	0%	-1%	-1%	<0.001
CT4	0	-7%	-6%	-7%	0.028	-3%	-3%	-4%	<0.001	-1%	-2%	-2%	<0.001
	<60	-4%	-5%	-9%	<0.001	-2%	-2%	-4%	0.002	-1%	-1%	-2%	0.032
	60-75	-4%	-3%	-11%	0.002	-2%	-3%	-6%	<0.001	-1%	-2%	-4%	<0.001
	>75	-4%	-6%	-13%	0.005	-2%	-4%	-6%	<0.001	-1%	-2%	-4%	<0.001

Indicated are differences for low, medium and high calcification mass. Small (<5%), medium (5-10%) and large (>10%) differences are denoted by respectively green, orange and red colored cells. Statistical results of the repeated measures ANOVA test are indicated, with significant p-values in bold

7

100.8 to 80.1. In total, 21 combinations of heart rate and calcification mass resulted in large differences in Agatston score between FBP and iterative reconstruction, while 40 medium differences were found. Large differences were not found for CT2, where only small and medium differences were given between Agatston scores from FBP and iterative reconstruction.

Similar trends were found for mass scores, although the differences were much smaller than for Agatston scores (*Table 3*). This resulted in a decrease in the number of large (n=5) and medium (n=26) differences. Despite these smaller differences, we still found large differences for low mass calcifications in combination with certain heart rates. The maximum difference of -13% was associated with the low mass calcification on CT4 at >75 bpm. While an increase in iterative reconstruction was typically associated with a decrease in mass score, an increase in mass score was shown for CT1 with the medium mass calcification. At >75 bpm and L2, an increase in mass score of 11% was shown.

DISCUSSION

In conclusion, the results of this study indicate that the effect of iterative reconstruction on CCS depends not only on CT system or calcification mass but also on heart rate. For repeated CCS measurements it was demonstrated that CT system, reconstruction type and heart rate should be similar.

The most clinically relevant finding of this study is that changes in reconstruction type as well as heart rate can cause changes in the measured amount of coronary calcium. The extent of deviation depends on the exact combination of CT system and iterative reconstruction type, as well as cardiac frequency. Our results, as stated in Tables 2 and 3, can be used to adjust CT measurements for changes in reconstruction, given a stable heart rate and CT system. Second, the impact of changes in iterative reconstruction strength on CCS was significantly different between all CT systems, as indicated by IRS analysis. In addition, the effect of iterative reconstruction on CCS was more prominent in calcification masses with lower density compared to higher density, these differences were also significant. For these lower density calcifications it was also found that the influence of iterative reconstruction is found to be comparable to the influence of heart rate. The main reason for this density-based iterative reconstruction influence is the threshold of 130 HU, which is used for CCS. With higher iterative reconstruction levels, more voxels inside the calcified region of interest fall below this threshold, because the calcification edge becomes less clear at increased heart rates, which in turn results

in a decreased number of voxels above the threshold of 130 HU. The importance of adequate quantification of low density calcifications was emphasized by Criqui et al, who showed that coronary calcification density was inversely associated with coronary heart disease and cardiovascular events.[30] Third, differences in CCS between FBP and iterative reconstruction were shown to increase with increasing iterative reconstruction level, increasing heart rate and decreasing calcification mass. This resulted in significant differences in CCS between all used reconstructions per heart rate for most combinations of CT systems and calcification masses. As stated before, CCS are used for treatment strategies according to recent AHA guidelines.[22] Not only CCS at baseline is important, also progression of CCS is gaining more interest.[31–35] For this reason, accurate determination of CCS in follow-up studies is necessary. The use of the more stable mass score, in comparison with the Agatston, could be helpful in this context. However, due to a lack of accepted reference values, this score is not widely used yet. In light of the results found in the current study, we recommend that consecutive CCS measurements should therefore be performed with the use of the same CT systems, the same reconstruction and, preferably, at a similar heart rate. Future research is needed to develop specific acquisition protocol recommendations for CCS assessment.

Previous studies have also focused on the influence of heart rate on CCS.[26,27,36–38] However, these studies did not assess the influence of iterative reconstruction at different heart rates. Our study is, to the best of our knowledge, the first study to assess this influence. We have shown for low mass calcifications that the influence of heart rate is comparable with the influence of iterative reconstruction for CT1 (-18% and -19%), CT2 (-9% and -6%) and CT4 (-13% and -13%), respectively. In contrast with CT3, where the influence of heart rate (5%) was not comparable with the influence of IR (-21%). The difference between CT3 and the other CT systems arises from the smaller temporal resolution of this system. In the current study, significant differences in CCS from different CT systems were shown. This finding is in agreement with a study from Willemink et al, who showed differences in median Agatston scores up to 43.9% between CT systems from four manufacturers.[14] However, this study assessed neither the influence of iterative reconstruction or heart rate, as it was performed on cadaveric human hearts. In addition, the underestimation of physical mass with the mass score was also demonstrated in previous studies.[39,40] Our study showed that the influence on CCS of changes in reconstruction is smaller for mass scores than for Agatston scores. IRS_{AS} were up to 130% higher than IRS_{MS}. This finding is in agreement with recent studies, which show smaller variability for mass scores.[26,38,41,42] Many recent studies

have shown that increasing levels of iterative reconstruction causes a reduction in CCS.[12,16,40,43–48] This result was also established in our study.

Limitations of this study have to be taken into consideration. First, although the motion of the phantom used in this study was linear and perpendicular to the scan direction, and the in-vivo motion of the coronary arteries exhibits a complex movement in 3D, we estimate that due to the fast rotation time and relatively short scan times, the motion in our phantom is a reliable first approximation of this complex motion. Second, for acquisition only sequential scans were performed. With new, high-pitch spiral mode scanning for coronary calcium appearing, this method could further improve temporal resolution over longer scan lengths. Third, only mild to severe coronary plaque burden was evaluated by the used calcium inserts. Therefore, our results are not directly applicable to very low to near-zero CCS. Fourth, the used calcium inserts had a relatively large size (diameter 5 mm, length 10 mm). However, quantification of the inserts still resulted in clinically relevant Agatston scores ranging from 100 to 400. Finally, acquisition protocols were not optimized for CT system comparisons.

In conclusion, the impact of iterative reconstruction on CCS is large for low mass calcifications. For medium and high mass calcifications, this impact is relatively small. The impact of iterative reconstruction on CCS further increases with increasing heart rate. In addition, IRS were significantly different between all CT systems and calcification masses. Therefore, this phantom study shows that for repeated CCS measurements, reconstruction type should be kept constant, especially for low mass calcifications. Also, heart rate and CT system should be kept constant for follow-up studies. These findings underscore that repeated CCS measurements should be acquired on the same CT system using the same reconstruction type and, ideally, at a similar heart rate. If the same setup is not available for repeated CCS measurements, the possible change in CCS as a result of a change in setup should be taken into account when interpreting the results.

REFERENCES

1. Sato J, Akahane M, Inano S, et al. Effect of radiation dose and adaptive statistical iterative reconstruction on image quality of pulmonary computed tomography. Jpn J Radiol. 2012;30(2):146-153. doi:10.1007/s11604-011-0026-7 [doi]

2. Moscariello A, Takx RA, Schoepf UJ, et al. Coronary CT angiography: image quality, diagnostic accuracy, and potential for radiation dose reduction using a novel iterative image reconstruction technique-comparison with traditional filtered back projection. Eur Radiol. 2011;21(10):2130-2138. doi:10.1007/s00330-011-2164-9 [doi]

3. Funama Y, Taguchi K, Utsunomiya D, et al. Combination of a low-tube-voltage technique with hybrid iterative reconstruction (iDose) algorithm at coronary computed tomographic angiography. J Comput Assist Tomogr. 2012;35(4):480-485. doi:10.1097/RCT.0b013e31821fee94

4. Gervaise A, Osemont B, Lecocq S, et al. CT image quality improvement using adaptive iterative dose reduction with wide-volume acquisition on 320-detector CT. Eur Radiol. 2012;22(2):295-301. doi:10.1007/s00330-011-2271-7

5. Korn A, Fenchel M, Bender B, et al. Iterative reconstruction in head CT: Image quality of routine and low-dose protocols in comparison with standard filtered back-projection. Am J Neuroradiol. 2012;33(2):218-224. doi:10.3174/ajnr.A2749

6. Prakash P, Kalra MK, Digumarthy SR, et al. Radiation Dose Reduction With Chest Computed Tomography Using Adaptive Statistical Iterative Reconstruction Technique: Initial Experience. J Comput Assist Tomogr. 2010;34(1):40-45. doi:10.1097/RCT.0b013e3181b26c67

7. Winklehner A, Karlo C, Puippe G, et al. Raw data-based iterative reconstruction in body CTA: Evaluation of radiation dose saving potential. Eur Radiol. 2011;21(12):2521-2526. doi:10.1007/s00330-011-2227-y

8. Leipsic J, LaBounty TM, Heilbron B, et al. Estimated radiation dose reduction using adaptive statistical iterative reconstruction in coronary CT angiography: The ERASIR study. Am J Roentgenol. 2010;195(3):655-660. doi:10.2214/AJR.10.4288

9. Singh S, Kalra MKM, Gilman MMD, et al. Adaptive statistical iterative reconstruction technique for radiation dose reduction in chest CT: a pilot study. Radiology. 2011;259(2):565-573. doi:10.1148/radiol.11101450

10. Nelson RC, Feuerlein S, Boll DT. New iterative reconstruction techniques for cardiovascular computed tomography: How do they work, and what are the advantages and disadvantages? J Cardiovasc Comput Tomogr. 2011;5(5):286-292. doi:10.1016/j.jcct.2011.07.001

11. Schindler A, Vliegenthart R, Schoepf UJ, et al. Iterative Image Reconstruction Techniques for CT Coronary Artery Calcium Quantification: Comparison with Traditional Filtered Back Projection in Vitro and in Vivo. Radiology. 2014;270(2):387-393. doi:10.1148/radiol.13130233

12. Van Osch JAC, Mouden M, Van Dalen JA, et al. Influence of iterative image reconstruction on CT-based calcium score measurements. Int J Cardiovasc Imaging. 2014;30(5):961-967. doi:10.1007/s10554-014-0409-9

13. Willemink MJ, Abramiuc B, den Harder AM, et al. Coronary calcium scores are systematically underestimated at a large chest size: A multivendor phantom study. J Cardiovasc Comput Tomogr. 2015;9(5):415-421. doi:10.1016/j.jcct.2015.03.010
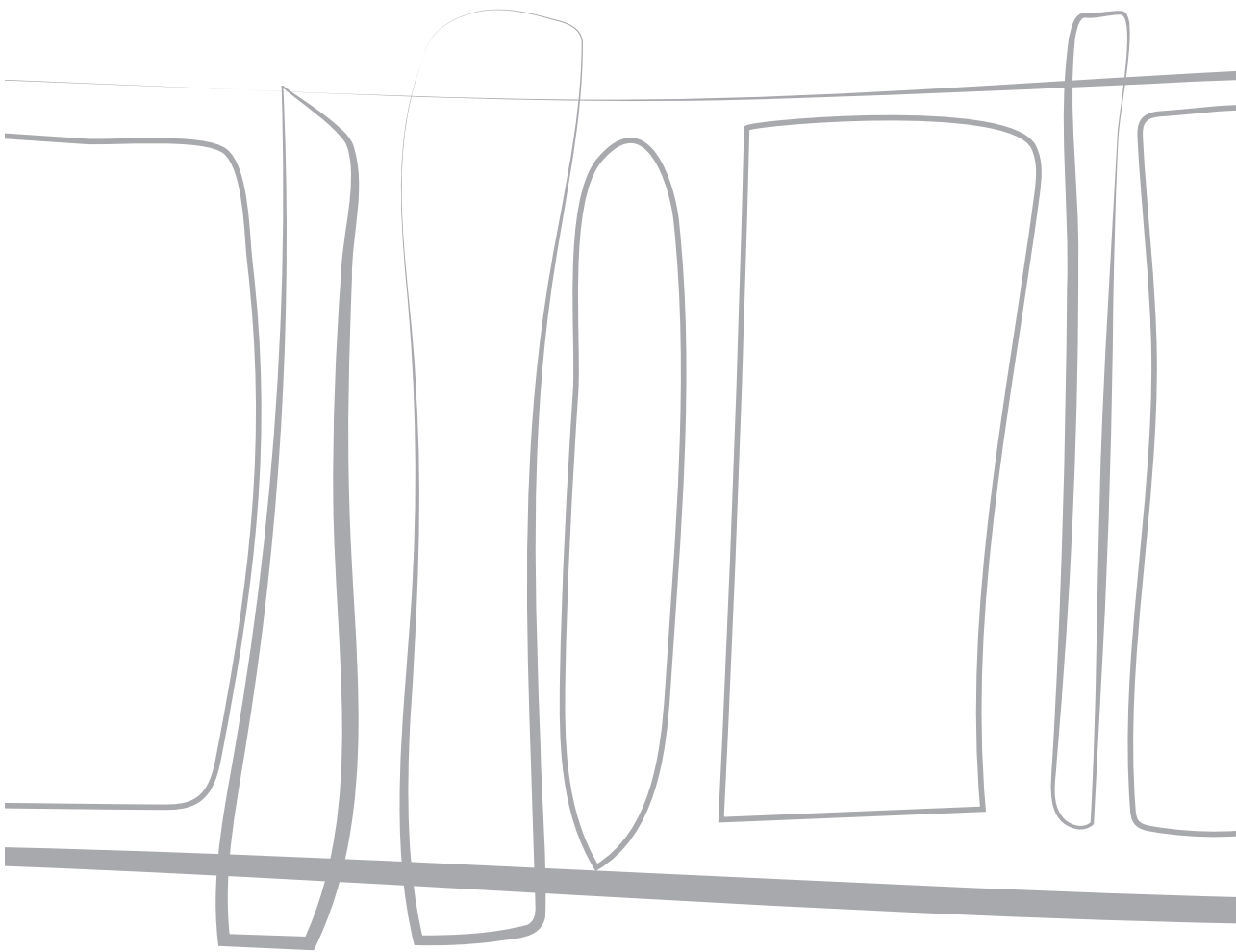
7

14. Willemink MJ, Vliegenthart R, Takx RAP, et al. Coronary Artery Calcification Scoring with State-of-the-Art CT Scanners from Different Vendors Has Substantial Effect on Risk Classification. Radiology. 2014;273(3):695-702. doi:10.1148/radiol.14140066

15. Murazaki H, Funama Y, Hatemura M, Fujioka C, Tomiguchi S. Quantitative evaluation of calcium (content) in the coronary artery using hybrid iterative reconstruction (iDose) algorithm on low-dose 64-detector CT: comparison of iDose and filtered back projection. Nihon Hoshasen Gijutsu Gakkai Zasshi. 2011;67(4):360-366. doi:10.6009/jjrt.67.360

16. Kurata A, Dharampal A, Dedic A, et al. Impact of iterative reconstruction on CT coronary calcium quantification. Eur Radiol. 2013;23(12):3246-3252. doi:10.1007/s00330-013-3022-8

17. Gebhard C, Fiechter M, Fuchs T a, et al. Coronary artery calcium scoring: Influence of adaptive statistical iterative reconstruction using 64-MDCT. Int J Cardiol. Published online September 2012:6-11. doi:10.1016/j.ijcard.2012.08.003

18. Szilveszter B, Elzomor H, Károlyi M, et al. The effect of iterative model reconstruction on coronary artery calcium quantification. Int J Cardiovasc Imaging. 2016;32(1):153-160. doi:10.1007/s10554-015-0740-9

19. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. J Am Coll Cardiol. 1990;15(4):827-832. doi:10.1016/0735-1097(90)90282-T

20. Wilson PWF, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation 1998. Circulation. 1998;97:1837-1847.

21. Greuter MJW, Groen JM, Nicolai LJ, Dijkstra H, Oudkerk M. A model for quantitative correction of coronary calcium scores on multidetector, dual source, and electron beam computed tomography for influences of linear motion, calcification density, and temporal resolution: A cardiac phantom study. Med Phys. 2009;36(11):5079-5088. doi:10.1118/1.3213536

22. Goff DC, Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: A report of the American college of cardiology/American heart association task force on practice guidelines. Circulation. 2014;129(25 SUPPL. 1):49-76. doi:10.1161/01.cir.0000437741.48606.98

23. Achenbach S, Ropers D, Holle J, Muschiol G, Daniel WG, Moshage W. In-plane coronary arterial motion velocity: measurement with electron-beam CT. Radiology. 2000;216(2):457-463. doi:10.1148/radiology.216.2.r00au19457

24. Husmann L, Leschka S, Desbiolles L, et al. Coronary artery motion and cardiac phases: dependency on heart rate -- implications for CT image reconstruction. Radiology. 2007;245(2):567-576. doi:10.1148/radiol.2451061791

25. Brodoefel H, Burgstahler C, Tsiflikas I, et al. Dual-Source CT: Effect of Heart Rate, Heart Rate Variability, and Calcification on Image Quality and Diagnostic Accuracy. Radiology. 2008;247(2):346-355. doi:10.1148/radiol.2472070906

26. Groen JM, Greuter MJW, Vliegenthart R, et al. Calcium scoring using 64-slice MDCT, dual source CT and EBT: A comparative phantom study. Int J Cardiovasc Imaging. 2008;24(5):547-556. doi:10.1007/s10554-007-9282-0

27. Tigges S, Arepalli CD, Tridandapani S, et al. A phantom study of the effect of heart rate, coronary artery displacement and vessel trajectory on coronary artery calcium score: Potential for risk misclassification. J Cardiovasc Comput Tomogr. 2012;6(4):260-267. doi:10.1016/j.jcct.2012.01.005

28. Willemink MJ, de Jong P a, Leiner T, et al. Iterative reconstruction techniques for computed tomography Part 1: Technical principles. Eur Radiol. Published online January 2013. doi:10.1007/s00330-012-2765-y

29. McCollough CH, Ulzheimer S, Halliburton SS, Shanneik K, White RD, Kalender WA. Coronary Artery Calcium: A Multi-institutional, Multimanufacturer International Standard for Quantification at Cardiac CT. Radiology. 2007;243(2):527-538. doi:10.1148/radiol.2432050808

30. Criqui MH, Denenberg JO, Ix JH, et al. Calcium Density of Coronary Artery Plaque and Risk of Incident Cardiovascular Events. Jama. 2014;311(3):271. doi:10.1001/jama.2013.282535

31. Youssef G, Kalia N, Darabian S, Budoff MJ. Coronary Calcium: New Insights, Recent Data, and Clinical Role. Curr Cardiol Rep. 2014;15(1):1-13. doi:10.1007/s11886-012-0325-3.Coronary

32. Budoff MJ, Hokanson JE, Nasir K, et al. Progression of coronary artery calcium predicts all-cause mortality. JACC Cardiovasc Imaging. 2010;3(12):1229-1236. doi:10.1016/j.jcmg.2010.08.018

33. Raggi P, Shaw LJ, Berman DS, Callister TQ. Prognostic value of coronary artery calcium screening in subjects with and without diabetes. J Am Coll Cardiol. 2004;43(9):1663-1669. doi:10.1016/j.jacc.2003.09.068

34. Raggi P, Cooil B, Shaw LJ, et al. Progression of coronary calcium on serial electron beam tomographic scanning is greater in patients with future myocardial infarction. Am J Cardiol. 2003;92(7):827-829.

35. Gopal A, Nasir K, Liu ST, Flores FR, Chen L, Budoff MJ. Coronary calcium progression rates with a zero initial score by electron beam tomography. Int J Cardiol. 2007;117(2):227-231. doi:10.1016/j.ijcard.2006.04.081

36. Funabashi N, Irie R, Namihira Y, et al. Influence of tube voltage and heart rate on the Agatston calcium score using an in vitro, novel ECG-gated dual energy reconstruction 320 slice CT technique. Int J Cardiol. 2015;180:218-220. doi:10.1016/j.ijcard.2014.11.164

37. Hong C, Bae KT, Pilgram TK, Zhu F. Coronary artery calcium quantification at multi-detector row CT: influence of heart rate and measurement methods on interacquisition variability initial experience. Radiology. 2003;228(1):95-100. doi:10.1148/radiol.2281020685

38. Groen JM, Greuter MJ, Schmidt B, Suess C, Vliegenthart R, Oudkerk M. The influence of heart rate, slice thickness, and calcification density on calcium scores using 64-slice multidetector computed tomography: A systematic phantom study. Invest Radiol. 2007;42(12):848-855. doi:10.1097/RLI.0b013e318154c549

39. Groen JM, Kofoed KF, Zacho M, Vliegenthart R, Willems TP, Greuter MJW. Calcium score of small coronary calcifications on multidetector computed tomography: Results from a static phantom study. Eur J Radiol. 2013;82(2):e58-e63. doi:10.1016/j.ejrad.2012.09.018

40. van der Werf NR, Willemink MJ, Willems TP, Greuter MJW, Leiner T. Influence of dose reduction and iterative reconstruction on CT calcium scores: a multi-manufacturer dynamic phantom study. Int J Cardiovasc Imaging. 2017;33(6):899-914. doi:10.1007/s10554-017-1061-y

7

41. Fuchs A, Groen JM, Arnold BA, et al. Assessment of coronary calcification using calibrated mass score with two different multidetector computed tomography scanners in the Copenhagen General Population Study. Eur J Radiol. 2017;88:21-25. doi:10.1016/j.ejrad.2016.12.033

42. Hoffmann U, Siebert U, Bull-Stewart A, et al. Evidence for lower variability of coronary artery calcium mineral mass measurements by multi-detector computed tomography in a community-based cohort--consequences for progression studies. Eur J Radiol. 2006;57(3):396-402. doi:10.1016/j.ejrad.2005.12.027

43. Rodrigues MA, ChB M, Williams MC, et al. Iterative reconstruction can permit the use of lower X-ray tube current in CT coronary artery calcium scoring 1. Br J Radiol. 2016;89. doi:10.1259/bjr.20150780

44. Matsuura N, Urashima M, Fukumoto W, et al. Radiation Dose Reduction at Coronary Artery Calcium Scoring by Using a Low Tube Current Technique and Hybrid Iterative Reconstruction. J Comput Assist Tomogr. 2014;00(00):1-6. doi:10.1097/RCT.0000000000000168

45. Li Q, Liu S, Myers KJ, et al. Impact of Reconstruction Algorithms and Gender-Associated Anatomy on Coronary Calcium Scoring with CT: An Anthropomorphic Phantom Study. Acad Radiol. 2016;23(12):1470-1479. doi:10.1016/j.acra.2016.08.014

46. Oda S, Utsunomiya D, Nakaura T, et al. The Influence of Iterative Reconstruction on Coronary Artery Calcium Scoring - Phantom and Clinical Studies. Acad Radiol. 2017;24(3):295-301. doi:10.1016/j.acra.2016.11.003

47. Takahashi M, Kimura F, Umezawa T, Watanabe Y, Ogawa H. Comparison of adaptive statistical iterative and filtered back projection reconstruction techniques in quantifying coronary calcium. J Cardiovasc Comput Tomogr. 2016;10(1):61-68. doi:10.1016/j.jcct.2015.07.012

48. Ode S, Kobayashi Y, Nozu Y, Ogawa Y. The Impact of Iterative Reconstruction on Reducing the Radiation Dose for Coronary Calcium Scoring : An Investigation Using Pulsating Calcified Coronary Phantom. 2016;7:95-103.

7
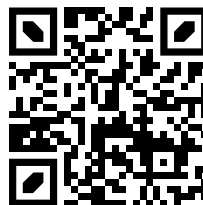
Online material is available via:

CHAPTER 8

Influence of dose reduction and iterative reconstruction on CT calcium scores: a multi-manufacturer dynamic phantom study

Niels R. van der Werf, MSc

Martin J. Willemink, MD PhD

Tineke P. Willems, MD PhD

Marcel J.W. Greuter, PhD

Tim Leiner, MD PhD

8

ABSTRACT

Purpose

To evaluate the influence of dose reduction in combination with iterative reconstruction (IR) on coronary calcium scores (CCS) in a dynamic phantom on state-of-the-art CT systems from different manufacturers.

Methods

Calcified inserts in an anthropomorphic chest phantom were translated at 20mm/s corresponding to heart rates between 60 and 75 bpm. The inserts were scanned five times with routinely used CCS protocols at reference dose and 40% and 80% dose reduction on four high-end CT systems. Filtered back projection (FBP) and increasing levels of IR were applied. Noise levels were determined. CCS, quantified as Agatston and mass scores, were compared to physical mass and scores at FBP reference dose.

Results

For the reference dose in combination with FBP, noise level variation between CT systems was less than 18%. Decreasing dose almost always resulted in increased CCS, while at increased levels of IR, CCS decreased again. The influence of IR on CCS was smaller than the influence of dose reduction. At reference dose, physical mass was underestimated 3 to 30%. All CT systems showed similar CCS at 40% dose reduction in combinations with specific reconstructions. For some CT systems CCS was not affected at 80% dose reduction, in combination with IR.

Conclusion

This multivendor study showed that radiation dose reductions of 40% did not influence CCS in a dynamic phantom using state-of-the-art CT systems in combination with specific reconstruction settings. Dose reduction resulted in increased noise and consequently increased CCS, whereas increased IR resulted in decreased CCS.

INTRODUCTION

The coronary calcium score (CCS) is known to be a strong predictor for major adverse cardiovascular events.[1,2] Computed tomography (CT) is the first modality of choice for assessment of the presence and quantification of calcium in the coronary arteries. The number of CCS examinations with CT is expanding rapidly.[3] However, due to the expanding use of ionizing radiation in medicine, CT has become the main source of increased population dose in Western countries.[4] This dose issue is especially important when considering the 2013 guidelines from the American Heart Association that recommend CCS measurements if, after quantitative risk assessment, the risk-based treatment decision is uncertain in asymptomatic adults at intermediate and low-to-intermediate risk.[5]

Recently, iterative reconstruction (IR) has become widely available on commercially available CT systems. IR allows for a dose reduction without the typical decrease in image quality.[6–8] It may therefore be possible to quantify CCS at lower dose levels, when using IR. Recent studies found that application of IR can result in spurious decreases in CCS in comparison with conventionally used filtered back projection (FBP).[9–11] These effects of dose reduction and IR on CCS can be explained by their effect on image noise. At decreased dose an increase in noise is expected. This increase in noise can be associated with an increase in voxels above the calcium threshold of 130 Hounsfield Units (HU), which in turn increases CCS. Conversely, a decrease in CCS is expected with IR since it reduces noise.[12–15]

Moreover, cardiac motion imposes problems for the stability of CCS since calcium can be blurred and CCS can be over- or underestimated, depending on the density of the calcification.[16–18] The combined effects of dose reduction, IR and heart rate on CCS for all major manufacturers have not been investigated before in a phantom study.

Therefore, the objective of this study was to evaluate the influence of dose reduction in combination with IR on CCS of moving calcifications in coronary CT on state-of-the-art CT systems from different manufacturers.

MATERIALS AND METHODS

An anthropomorphic chest phantom (Thorax, QRM, Moehrendorf, Germany) with artificial lungs, a spine insert and a shell of soft tissue equivalent material was used.[16,17] An extension ring of tissue equivalent material was placed around the

8

chest to simulate an averaged sized patient of 400 x 300 mm (QRM-Extensionring, QRM, Germany).[19] The center compartment of the phantom was filled with water in which a motion simulator (Sim2D, QRM, Moehrendorf, Germany) translated an artificial coronary artery with two calcium hydroxyapatite (HA) inserts. The inserts had densities of 196±3, 380±2, 408±2 and 800±2 mg HA / cm^3 and masses of 38.5±1.7, 74.6±3.1, 80.1±3.3 and 157.1±6.5 mg HA, respectively (Appendix B and C).

All inserts had equal dimensions; length 10.0±0.1 mm, diameter 5.0±0.1 mm, volume 196.3±8.1 mm^3. The artificial arteries were linearly translated in the horizontal plane at a velocity of 20 mm/s perpendicular to the scan direction. This velocity is comparable to typical velocities of the left anterior descending and right coronary arteries during the late diastolic scan phase of the R-R interval, at heart rates between 60 and 75 bpm.[20,21]

In order to assess the influence of IR and dose reduction on CCS in a clinical setting, daily used clinical CT protocols for coronary calcium scoring were used. These protocols were equal to the vendor recommended protocols if available or were adapted based on recommendation by the specific manufacturer consultants. Four different state-of-the-art CT systems (referred to as S1 to S4) were used: Discovery CT 750 HD (GE Healthcare, Waukesha, Wisconsin, USA), Brilliance iCT (Philips Healthcare, Best, The Netherlands), Somatom Definition Flash (Siemens Healthcare, Forchheim, Germany) and Aquilion One (Toshiba Medical Systems, Otaware, Japan), respectively (*Table 1*).

The phantom was scanned at three dose levels by reduction of the tube current: a reference dose at 100% tube current, and at reduced dose levels of 40% and 80% reduced tube current. Each scan was repeated five times with a small translation (2 mm) and rotation (2 degrees) between each scan by manually repositioning the phantom. The internal ECG signal of the motion controller was used to simulate the heart rate of the patient and used as ECG trigger on all four CT systems. The triggering was carefully timed so that data acquisition was during linear motion of the phantom.

Images were reconstructed with FBP, and three increasing levels of IR: the lowest (L1), an intermediate (L2) and the highest level available on the CT system (L3) (*Table 1*). For each data set the noise level in the images was assessed by calculating the standard deviation in the average Hounsfield value in a uniform water region. The amount of calcium of each insert was quantified as Agatston and mass scores

Table 1. Acquisition and reconstruction parameters used on CT system S1 to S4.

CT system	S1	S2	S3	S4
Tube voltage	120 kV	120 kV	120 kV	120 kV
Tube charge per rotation	500 mA	185 mA	285 mA	230 mA
Collimation	64 x 0.625 mm	128 x 0.625 mm	128 x 0.6 mm	320 x 0.5 mm
Rotation time	0.35 s	0.27 s	0.28 s	0.35 s
Temporal resolution[a]	175 ms	135 ms	75 ms	175 ms
Slice thickness	2.5 mm	3.0 mm	3.0 mm	3.0 mm
Increment	2.5 mm	3.0 mm	3.0 mm	3.0 mm
Kernel	Standard	XCA	B35f	FC12
Levels of IR	20, 60, 100%	1, 5, 7	1, 3, 5	weak, standard, strong
Noise level	26 HU	22 HU	28 HU	24 HU
$CTDI_{vol}$	10.6 mGy	3.2 mGy	2.8 mGy	6.5 mGy
Software	Smartscore 4.0	Heartbeat-CS	Syngo	Vitrea FX 6.5.0

[a] As defined in the isocenter

with manufacturer-recommended software (*Table 1*) with a default threshold of 130 Hounsfield units (HU). A semi-automatic method was used for selecting the calcification by one observer. On each CT system, the mass score calibration factors were determined as described by McCollough et al.[19] Although mass scores were not used clinically, they were included for this study because of its potential to compare the score to the physical mass.

The design of this study resulted in 480 calcium scores per CT system (5 acquisitions at 3 dose levels with 4 reconstruction types for 4 calcifications and 2 calcium scores).

Agatston score and mass score were expressed as median and 25th to 75th percentile for each calcification insert and CT system. For each insert, CCS from both the iteratively reconstructed and FBP reconstructed data sets for reduced dose levels were compared to the CCS from the FBP reconstructed data sets at reference dose using a Wilcoxon signed rank test. All statistical analyses were performed with SPSS for Windows, version 22.0. A p value of 0.05 was used to determine significant differences.

RESULTS

Influence of dose reduction and iterative reconstruction on noise

For all CT systems and all reconstructions, a decrease in dose resulted in a vendor dependent increase in noise, whereas IR led to a decrease in noise (*Figure 1*). Also, although the $CTDI_{vol}$ differed at most with a factor of 3.8 between the CT systems, the noise levels varied less than 18% at FBP reference dose.
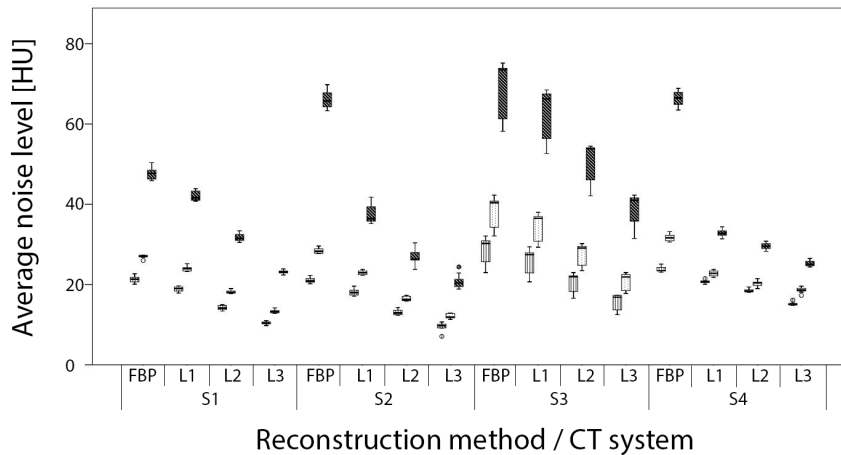

Figure 1 Average noise levels in a uniform water region for images reconstructed with filtered back projection (FBP) and increasing levels of IR L1, L2 and L3 as measured on CT systems S1, S2, S3 and S4. For each combination of reconstruction method and CT system boxplots are shown for the average noise level at reference dose, and 40% and 80% reduced dose

Influence of dose reduction on Agatston score with FBP

Dose reduction resulted in significant increases in Agatston scores for almost all calcifications and CT systems (*Figure 2*). This increase, in combination with an increase in noise, is depicted in the top row of *Figure 3*.

For S1 at FBP and averaged over all inserts, Agatston scores increased by 8% and 25% at 40% and 80% reduced dose respectively. For the other CT systems similar increases in Agatston scores at FBP were observed at reduced dose with a corresponding average increase of 7% and 64% for S2, 4% and 26% for S3, and 1% and 23% for S4. The largest increase in Agatston score at reduced dose was observed for the 38 mg insert at 80% dose reduction: 58%, 160%, 48%, and 71% for S1 to S4, respectively.

Figure 2 Influence of dose reduction on Agatston score for S1 to S4 with FBP. The movement of the calcification corresponds to displacements seen with heart rates of 60 – 75 bpm. For calcifications of 38, 74, 80 and 157 mg boxplots of the Agatston score at reference dose, and at 40% and 80% reduced dose are shown. Agatston scores are compared with the Agatston score at reference dose using the Wilcoxon signed rank test. Significant different Agatston scores are indicated by brackets

Influence of dose reduction on mass score with FBP

Also, dose reductions resulted in significantly increased mass scores at FBP for almost all inserts and CT systems, albeit that the increase was smaller than the increase in Agatston scores (*Figure 4*).

At 40% reduced dose, mass scores increased on average by 0%, 3%, 1% and 0% for S1 to S4 respectively in comparison with the mass score at reference dose. At 80% reduced dose, mass scores increased 35%, 15% and 13% for S2 to S4, whereas for S1 the mass score decreased 11%.

151

8
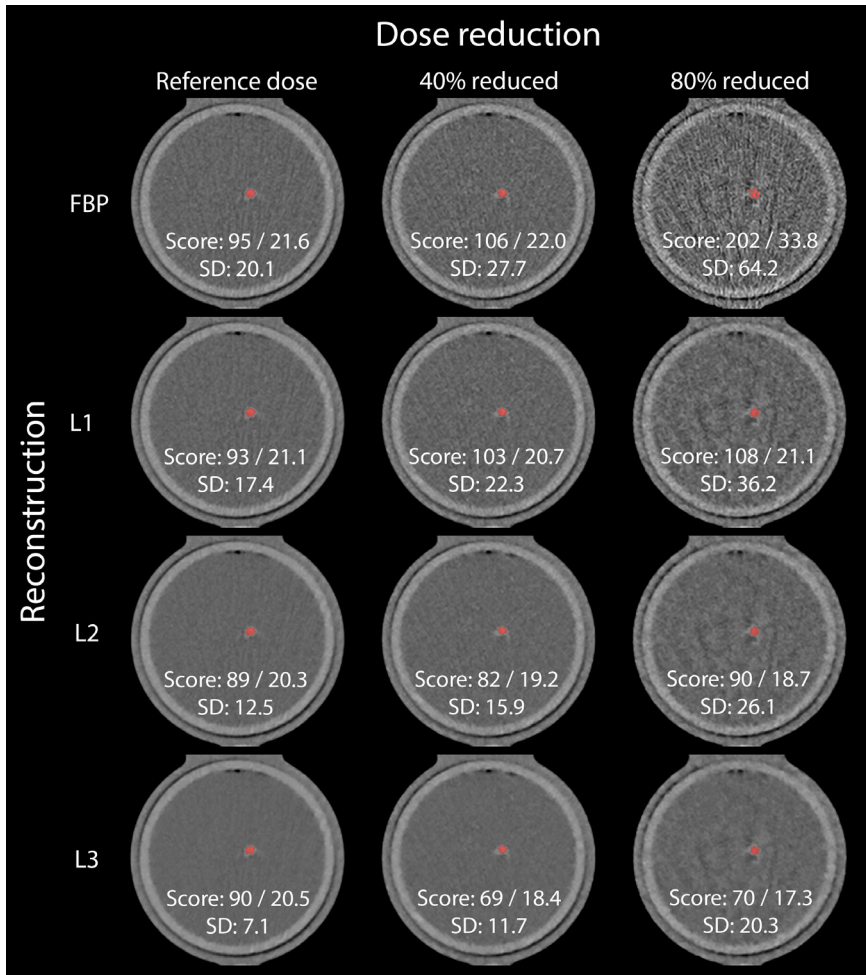
Figure 3 Reconstructed images of a 3.0 mm slice of the 38 mg insert moving at 20 mm/s on S2. Data was acquired at reference dose, and at 40% and 80% dose reduction (from left to right) and reconstructed with filtered back projection and increasing levels of IR L1, L2 and L3 (from top to bottom). CCS were included as Agatston score / Mass score. Noise levels (SD) are expressed as Hounsfield Units. Window center was 90 HU and window width was 750 HU

Influence of iterative reconstruction on Agatston scores

With increased IR levels, a significant decrease in Agatston scores was observed for almost all calcifications and CT systems (*Figure 5*). This decrease in Agatston score was accompanied by decrease in noise, as can be seen from the left column in *Figure 3*.
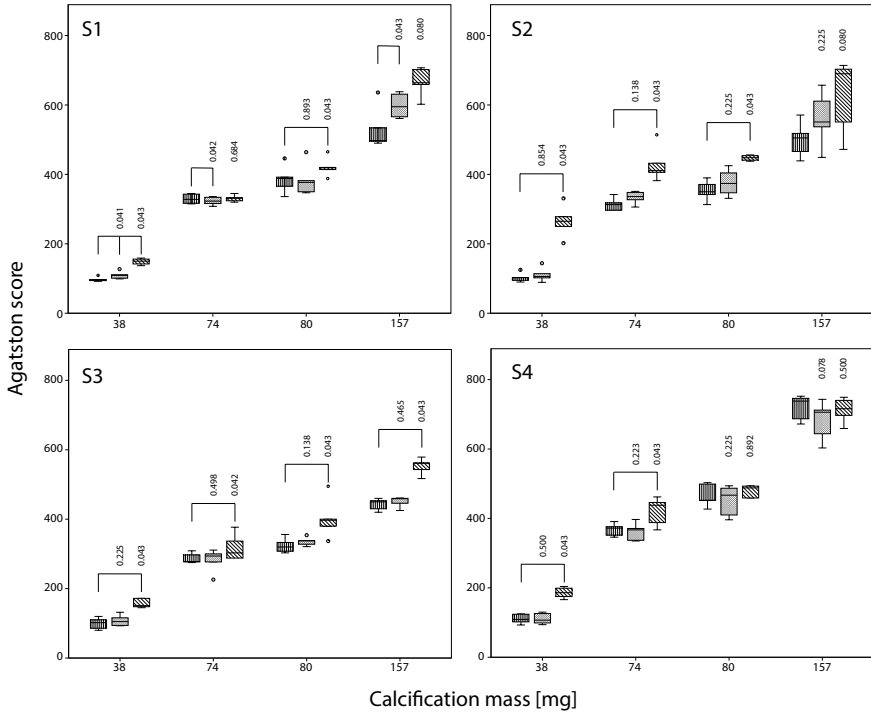
Figure 4 Influence of dose reduction on mass score for S1 to S4 with FBP. The movement of the calcification corresponds to displacements seen with heart rates of 60 – 75 bpm. For calcifications of 38, 74, 80 and 157 mg boxplots of the mass score at reference dose, and at 40% and 80% reduced dose are shown. Mass scores are compared with the Mass score at reference dose using the Wilcoxon signed rank test. Significant different Mass scores are indicated by brackets

Averaged over all inserts, Agatston scores for S1 decreased on average 0%, 2% and 5% at L1 to L3 respectively. For S2 the corresponding decrease was 1%, 4%, and 5%; for S3 1%, 4%; and 9% and for S4 1%, 4%, and 7%. The largest decrease in Agatston score was again observed for the 38 mg calcification: 22% with L3 on S3, and 19% with L3 on S4.

Influence of iterative reconstruction on mass scores

The decrease in mass scores at increased levels of IR was smaller than the observed decrease in Agatston scores (*Figure 6*). Mass score decreased on average between 0% and 6% for all CT systems and inserts.
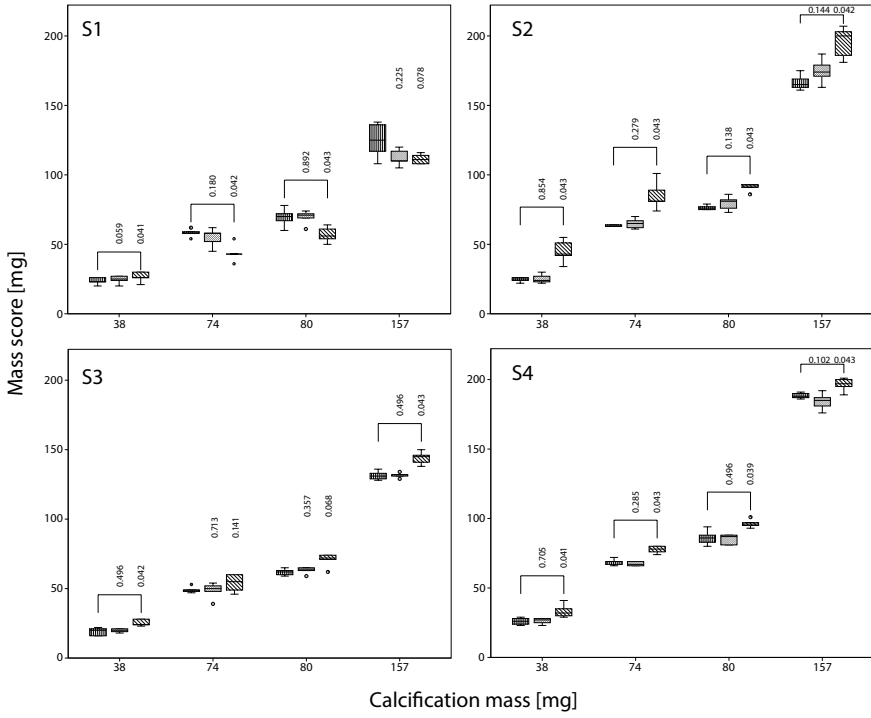
Figure 5 Influence of IR on Agatston score for S1 to S4 with FBP. The movement of the calcification corresponds to displacements seen with heart rates of 60 – 75 bpm. For calcifications of 38, 74, 80 and 157 mg boxplots of the Agatston score at FBP and increasing levels of IR (from left to right: L1, L2 and L3) are shown. Agatston scores are compared with the Agatston score at FBP using the Wilcoxon signed rank test. Significant different Agatston scores are indicated by brackets

Combination of dose reduction and iterative reconstruction on Agatston and mass scores

Representative images of the reconstructed datasets are shown in *Figure 3*.

For all four CT systems 40% dose reduction in combination with varying levels of IR did not result in significantly different Agatston and mass scores with respect to the reference dose (*Table 2*). For 80% dose reduction, only S2 in combination with L2 and L3 did not result in significantly different Agatston scores. For the other CT systems, there was no combination of investigated imaging parameters that resulted in Agatston scores which were unchanged from the reference protocol and dose.
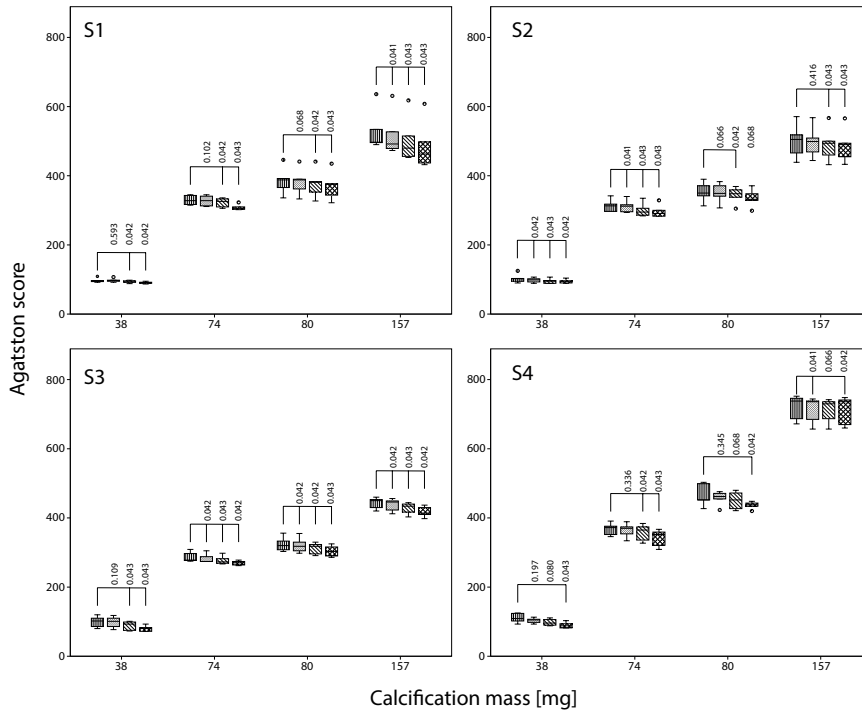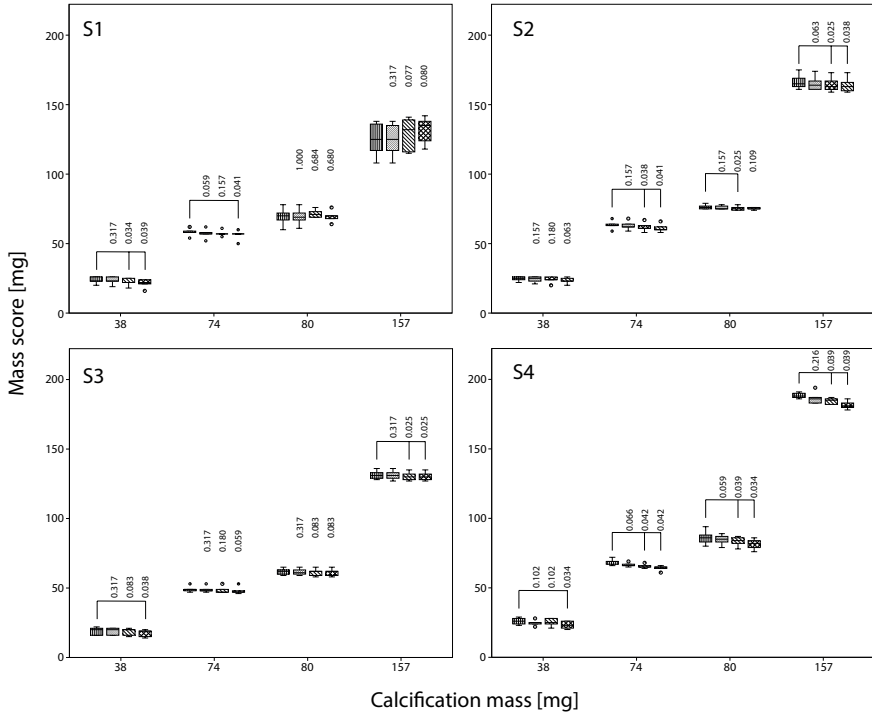
Figure 6 Influence of IR on mass score for S1 to S4 with FBP. The movement of the calcification corresponds to displacements seen with heart rates of 60 – 75 bpm. For calcifications of 38, 74, 80 and 157 mg boxplots of the mass score at FBP and increasing levels of IR (from left to right: L1, L2 and L3) are shown. Mass scores are compared with the Mass score at FBP using the Wilcoxon signed rank test. Significant different Mass scores are indicated by brackets

Table 2. Reconstructions per CT system S1 to S4 that did not result in significantly different Agatston and mass scores at 60-75 bpm and at a dose reduction of 40% and 80% with respect to the FBP-reference dose. FBP = filtered back projection; L1, L2, L3 = increasing levels of iterative reconstruction

CT system	Dose reduction	Agatston score	Mass scores
S1	40%	L1	FBP, L1
	80%	n/a	n/a
S2	40%	FBP, L1, L2, L3	FBP, L1, L2, L3
	80%	L2, L3	L1
S3	40%	FBP, L1, L2	FBP, L1, L2, L3
	80%	n/a	L2, L3
S4	40%	FBP	FBP
	80%	n/a	n/a

On all CT systems, mass scores generally underestimated the physical mass of the calcifications. Mass scores at FBP and reference dose and deviations from the physical mass are listed in *Table 3*. Averaged over all inserts the physical mass was underestimated by 23%, 12%, 30%, and 3% for S1 to S4 respectively. The largest underestimation was again observed for the 38 mg insert, where the underestimation was 39%, 33%, 30%, and 31%, respectively for S1 to S4. At 40% reduced dose the underestimation was 24%, 9%, 29%, and 3%. At 80% reduced dose the underestimation was 24% and 29% on S1 and S3, whereas S2 and S4 showed an overestimation of on average 16% and 9%. The influence of IR on mass scores was relatively small compared to the influence of dose reduction. At the maximum IR level, the underestimation of the physical mass at reference dose was 23%, 15%, 32%, and 8% for S1 to S4, respectively (averaged over all inserts).

8

Table 3. Physical mass and corresponding Mass scores for all CT systems and calcification masses. The mass scores are expressed as median and range. The difference between the median and physical mass is also given as median and range

CT system	Physical mass [mg]	Mass score [mg]	Deviation [%]
S1	38	23 (20-26)	-39% (-47%; -32%)
	74	58 (54-62)	-22% (-27%; -16%)
	80	70 (60-78)	-13% (-25%; -3%)
	157	125 (108-138)	-20% (-31%; -12%)
S2	38	25 (22-26)	-33% (-43%; -31%)
	74	63 (59-68)	-15% (-20%; -8%)
	80	76 (75-79)	-5% (-6%; -1%)
	157	165 (161-175)	5% (3%; 11%)
S3	38	20 (16-22)	-46% (-57%; -42%)
	74	49 (47-53)	-34% (-37%; -28%)
	80	62 (59-65)	-23% (-26%; -19%)
	157	131 (128-136)	-17% (-19%; -13%)
S4	38	26 (23-29)	-31% (-40%; -24%)
	74	69 (66-72)	-7% (-11%; -3%)
	80	86 (80-94)	7% (0%; 18%)
	157	188 (186-191)	20% (19%; 21%)

DISCUSSION

To our knowledge this is the first multivendor study to evaluate the effects of dose reduction and IR on CCS in a dynamic phantom. We have shown that dose reduction in dynamic coronary calcium CT can result in a substantial increase in CCS, whereas the use of IR results in modestly decreased CCS. The most important clinically relevant finding is the ability to reduce dose by 40% in routinely used clinical protocols on state-of-the-art CT systems of 4 major manufacturers, without compromising the calcium score. This result is not only valid for high plaque burden, but also for the clinically more important mild to moderate coronary plaque burden, represented by the 38 and 74 mg calcifications respectively.

Since risks of radiation dose increase with growing numbers of CT examinations, dose reduction techniques in CCS are highly relevant. Because new guidelines recommend CCS measurements if, after quantitative risk assessment, the risk-based treatment decision is uncertain, it is expected that the number of CT examinations for CCS will further increase in coming years.[5] In the current study we found for all CT systems that dose reductions of 40%, in combination with the in *Table 2* specified reconstruction methods, did not significantly affect Agatston scores. For one vendor, the Agatston scores were even similar at 80% reduced dose, and for two vendors there was no significant difference in mass scores at 80% reduced dose in combination with IR.

These results are consistent with those of Hecht et al (2015) who showed in a patient study that for one CT system (equal to S2) CCS can be performed at reduced radiation dose (50%) in combination with IR, without significantly affecting Agatston scores.[15] Ode et al showed, for a pulsating phantom at 60 bpm and one CT system (similar to S4), that increased IR resulted in decreased Agatston scores, which is in agreement with our results.[22] In comparison with full dose FBP, Agatston scores were not influenced at IR levels L2 and L3 in combination with dose reduction up to 75%, for all used calcifications combined. In our study however, Agatston scores at 40 and 80% reduced dose were found to be significantly different for all IR levels. The reason for this difference is that we only included combinations of dose reduction and IR, when valid for all calcifications separately. Our results also correspond well with a recent study which showed that IR has the potential to reduce radiation dose with 27-54% using a non-dynamic phantom and the same CT systems.[23] With non-dynamic ex vivo human hearts it was shown that a dose reduction of 80% was possible for the four CT systems.[23,24] This study, however, used static calcifications, did not report on a reference standard of true

calcification mass, and used a small-sized phantom. In our dynamic study, we found that a dose reduction of 80% was only feasible for one CT system, and a dose reduction of 40% was possible for all four CT systems, even for low-density calcifications in combination with specific reconstruction methods. Because iterative CT reconstruction significantly reduces calcium scores which potentially alters perceived cardiovascular risk, this effect may be counter balanced by the use of reduced dose levels.[10,25,26] Moreover, it has been shown that the application of IR significantly improves objective image quality, and does not alter quantitative analysis of coronary plaque volume, composition and luminal area.[12,27]

Our results showed a relatively large variation in calcium scores between the CT systems, with Agatston scores ranging from 450 to 738, for the 157 mg calcification. This is in line with previous studies that found that state-of the-art CT scanners of different manufacturers produce substantially different Agatston scores, which can result in reclassification of patients to high- or low-risk categories in up to 6.5% of the cases.[28] Moreover, mass scores generally underestimated the physical mass of the inserts by 3 to 23% depending on the specific CT system. Underestimations of the physical mass up to 68% were also observed with a static calcium phantom.[29]

Reference dose levels, from routinely used clinical protocols of the four high-end CT systems, showed large differences (2.8 – 10.6 mGy). Despite of these differences in dose levels, similar noise levels were found (22 – 28 HU). It is important to note, however, that noise is not only determined by dose, but - among other parameters - also by reconstruction kernel. A sharper kernel results in more noise as compared to a softer kernel, if the dose levels are the same. Therefore, different CT acquisition and reconstruction settings may result in different dose levels but similar noise levels. The noise levels behaved as expected as a function of dose reduction and IR: noise levels increased at decreasing dose, and noise levels decreased at increased IR. Our findings indicate that even in the presence of comparable noise levels CCS differed up to 39% between different CT systems at full dose FBP. These differences are surprising for a relatively straightforward metric as the coronary calcium score.

This study has limitations. First, this was an in-vitro study with artificial arteries with calcified inserts. However, the inserts where embedded in an anthropomorphic phantom and were translated at a velocity that is generally observed in in-vivo studies, and the masses of the inserts were in range with calcium masses clinically detected in patients.[30] Second, movement of the calcifications was linear. In vivo, coronary arteries perform a complex movement in three dimensions, which was not feasible in our setup. However, because a linear movement can approximate

the movement in 3D during the acquisition time of the CT data, we estimate that addition of 3D movement would result in minor changes in our results. Third, analysis on the inter and intra variability for the different CT systems has not been performed. The associated CT specific correlation between noise reduction and CCS accuracy was also not within the scope of this study. However, these analysis can answer questions about current practice. For example specificity, sensitivity, variations in CCS score between different vendors and the possibility to reduce dose without impact on the metric. Finally, only sequential scan modes were used. With the current appearance of high-pitch spiral mode scanning for coronary calcium it would be interesting to assess the differences in the accuracy of coronary calcium assessment between sequential and high-pitch spiral mode. However, that was not within the scope of this study.

We conclude that for all CT systems a dose reduction of 40% in combination with specific reconstruction gives a CCS comparable for reference protocols. For several systems, even higher dose reductions are possible. Dose reduction results in increased noise and consequently increased CCS, whereas increased IR results in decreased CCS. Mass scores generally underestimated physical mass of the calcifications.
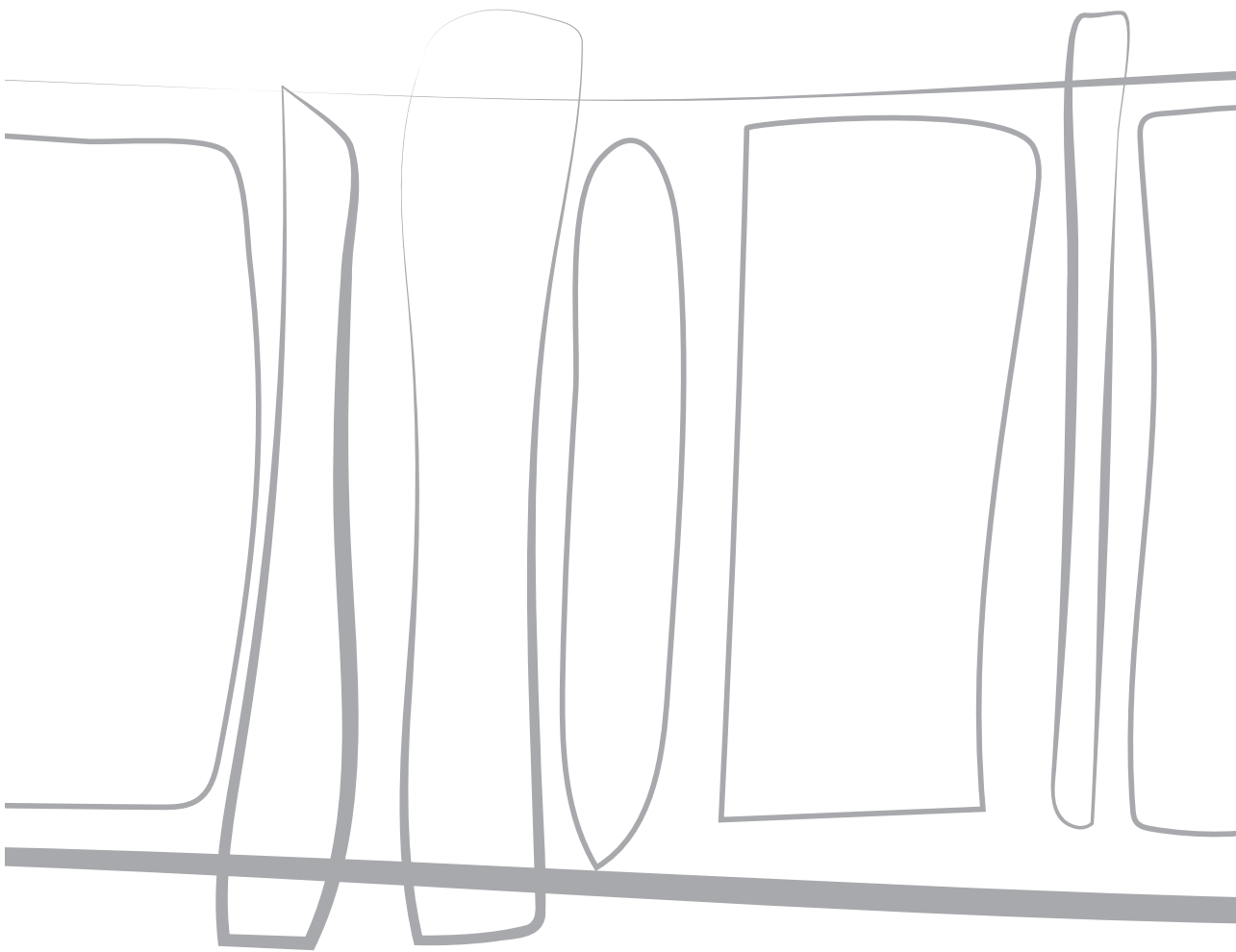
8

REFERENCES

1. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. *J Am Coll Cardiol*. 1990;15(4):827-832. doi:10.1016/0735-1097(90)90282-T

2. Versteylen MO, Joosen IA, Winkens MH, et al. Combined use of exercise electrocardiography, coronary calcium score and cardiac CT angiography for the prediction of major cardiovascular events in patients presenting with stable chest pain. *Int J Cardiol*. 2013;167(1):121-125. doi:10.1016/j.ijcard.2011.12.016

3. Brenner DJ, Hall EJ. Current concepts - Computed tomography - An increasing source of radiation exposure. *N Engl J Med*. 2007;357(22):2277-2284. doi:10.1056/NEJMra072149

4. Berrington de González A, Darby S. Risk of cancer from diagnostic X-rays: estimates for the UK and 14 other countries. *Lancet*. 2004;363(9406):345-351. doi:10.1016/S0140-6736(04)15433-0

5. Goff DC, Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: A report of the American college of cardiology/American heart association task force on practice guidelines. *Circulation*. 2014;129(25 SUPPL. 1):49-76. doi:10.1161/01.cir.0000437741.48606.98

6. Leipsic J, LaBounty TM, Heilbron B, et al. Estimated radiation dose reduction using adaptive statistical iterative reconstruction in coronary CT angiography: The ERASIR study. *Am J Roentgenol*. 2010;195(3):655-660. doi:10.2214/AJR.10.4288

7. Moscariello A, Takx RA, Schoepf UJ, et al. Coronary CT angiography: image quality, diagnostic accuracy, and potential for radiation dose reduction using a novel iterative image reconstruction technique-comparison with traditional filtered back projection. *Eur Radiol*. 2011;21(10):2130-2138. doi:10.1007/s00330-011-2164-9 [doi]

8. Sato J, Akahane M, Inano S, et al. Effect of radiation dose and adaptive statistical iterative reconstruction on image quality of pulmonary computed tomography. *Jpn J Radiol*. 2012;30(2):146-153. doi:10.1007/s11604-011-0026-7 [doi]

9. Fleischmann D, Boas FE. Computed tomography--old ideas and new technology. *Eur Radiol*. 2011;21(3):510-517. doi:10.1007/s00330-011-2056-z

10. Gebhard C, Fiechter M, Fuchs T a, et al. Coronary artery calcium scoring: Influence of adaptive statistical iterative reconstruction using 64-MDCT. *Int J Cardiol*. Published online September 2012:6-11. doi:10.1016/j.ijcard.2012.08.003

11. Willemink MJ, Takx RAP, De Jong PA, et al. The impact of CT radiation dose reduction and iterative reconstruction algorithms from four different vendors on coronary calcium scoring. *Eur Radiol*. 2014;24(9):2201-2212. doi:10.1007/s00330-014-3217-7

12. Willemink MJ, de Jong P a, Leiner T, et al. Iterative reconstruction techniques for computed tomography Part 1: Technical principles. *Eur Radiol*. Published online January 2013. doi:10.1007/s00330-012-2765-y

13. Schindler A, Vliegenthart R, Schoepf UJ, et al. Iterative Image Reconstruction Techniques for CT Coronary Artery Calcium Quantification: Comparison with Traditional Filtered Back Projection in Vitro and in Vivo. *Radiology*. 2014;270(2):387-393. doi:10.1148/radiol.13130233

8

14. Takahashi M, Kimura F, Umezawa T, Watanabe Y, Ogawa H. Comparison of adaptive statistical iterative and filtered back projection reconstruction techniques in quantifying coronary calcium. *J Cardiovasc Comput Tomogr*. 2016;10(1):61-68. doi:10.1016/j.jcct.2015.07.012

15. Hecht HS, De Siqueira MEM, Cham M, et al. Low- vs. Standard-dose coronary artery calcium scanning. *Eur Heart J Cardiovasc Imaging*. 2015;16(4):358-363. doi:10.1093/ehjci/jeu218

16. Greuter MJW, Groen JM, Nicolai LJ, Dijkstra H, Oudkerk M. A model for quantitative correction of coronary calcium scores on multidetector, dual source, and electron beam computed tomography for influences of linear motion, calcification density, and temporal resolution: A cardiac phantom study. *Med Phys*. 2009;36(11):5079-5088. doi:10.1118/1.3213536

17. Groen JM, Greuter MJW, Vliegenthart R, et al. Calcium scoring using 64-slice MDCT, dual source CT and EBT: A comparative phantom study. *Int J Cardiovasc Imaging*. 2008;24(5):547-556. doi:10.1007/s10554-007-9282-0

18. Groen JM, Dijkstra H, Greuter MJ, Oudkerk M. Threshold adjusted calcium scoring using CT is less susceptible to cardiac motion and more accurate. *Med Phys*. 2009;36(2):438-446.

19. McCollough CH, Ulzheimer S, Halliburton SS, Shanneik K, White RD, Kalender WA. Coronary Artery Calcium: A Multi-institutional, Multimanufacturer International Standard for Quantification at Cardiac CT. *Radiology*. 2007;243(2):527-538. doi:10.1148/radiol.2432050808

20. Achenbach S, Ropers D, Holle J, Muschiol G, Daniel WG, Moshage W. In-plane coronary arterial motion velocity: measurement with electron-beam CT. *Radiology*. 2000;216(2):457-463. doi:10.1148/radiology.216.2.r00au19457

21. Husmann L, Leschka S, Desbiolles L, et al. Coronary artery motion and cardiac phases: dependency on heart rate -- implications for CT image reconstruction. *Radiology*. 2007;245(2):567-576. doi:10.1148/radiol.2451061791

22. Ode S, Kobayashi Y, Nozu Y, Ogawa Y. The Impact of Iterative Reconstruction on Reducing the Radiation Dose for Coronary Calcium Scoring : An Investigation Using Pulsating Calcified Coronary Phantom. 2016;7:95-103.

23. Willemink MJ, Jong PA De, Prokop M, Mey J De, Leiner T, Schilham AMR. Computed Tomography Radiation Dose Reduction: Effect of Different Iterative Reconstruction Algorithms on Image Quality. *J Comput Assist Tomogr*. Published online 2014.

24. den Harder AM, Willemink MJ, Bleys RL a W, et al. Dose reduction for coronary calcium scoring with hybrid and model-based iterative reconstruction: an ex vivo study. *Int J Cardiovasc Imaging*. 2014;30(6):1125-1133. doi:10.1007/s10554-014-0434-8

25. Van Osch JAC, Mouden M, Van Dalen JA, et al. Influence of iterative image reconstruction on CT-based calcium score measurements. *Int J Cardiovasc Imaging*. 2014;30(5):961-967. doi:10.1007/s10554-014-0409-9

26. Kurata A, Dharampal A, Dedic A, et al. Impact of iterative reconstruction on CT coronary calcium quantification. *Eur Radiol*. 2013;23(12):3246-3252. doi:10.1007/s00330-013-3022-8

27. Takx RAP, Willemink MJ, Nathoe HM, et al. The effect of iterative reconstruction on quantitative computed tomography assessment of coronary plaque composition. *Int J Cardiovasc Imaging*. 2014;30(1):155-163. doi:10.1007/s10554-013-0293-8

8

28. Willemink MJ, Vliegenthart R, Takx RAP, et al. Coronary Artery Calcification Scoring with State-of-the-Art CT Scanners from Different Vendors Has Substantial Effect on Risk Classification. *Radiology*. 2014;273(3):695-702. doi:10.1148/radiol.14140066

29. Groen JM, Kofoed KF, Zacho M, Vliegenthart R, Willems TP, Greuter MJW. Calcium score of small coronary calcifications on multidetector computed tomography: Results from a static phantom study. *Eur J Radiol*. 2013;82(2):e58-e63. doi:10.1016/j.ejrad.2012.09.018

30. Hong C, Pilgram TK, Zhu F, Bae KT. Is coronary artery calcium mass related to Agatston score? *Acad Radiol*. 2004;11(3):286-292. doi:10.1016/s1076-6332(03)00714-1

8
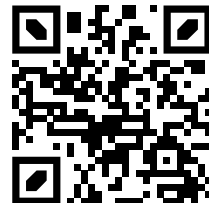
Online material is available via:

CHAPTER 9

Coronary calcium scoring potential of large field-of-view spectral photon counting CT: a phantom study

Niels R. van der Werf, MSc*

Salim Si-Mohamed, MD PhD*

Pierre-Antoinne Rodesch, MSc

Robbert W. van Hamersvelt, MD PhD

Marcel J.W. Greuter, PhD

Sara Boccalini, MD PhD

Joël Greffier, PhD

Tim Leiner, MD PhD

Loïc Boussel, PhD

Martin J. Willemink, MD PhD

Philippe Douek, MD PhD

* Authors contributed equally to this work

ABSTRACT

Objective

The aim of the current study was, first, to assess the coronary artery calcium (CAC) scoring potential of spectral photon counting CT (SPCCT) in comparison with computed tomography (CT) for routine clinical protocols. Second, improved CAC detection and quantification at reduced slice thickness was assessed.

Methods

Raw data was acquired and reconstructed with several combinations of reduced slice thickness and increasing strengths of iterative reconstruction (IR) for both CT systems with routine clinical CAC protocols for CT. Two CAC containing cylindrical inserts, consisting of CAC of different densities and sizes, were placed in an anthropomorphic phantom. A specific CAC was detectable when 3 or more connected voxels exceeded the CAC scoring threshold of 130 Hounsfield units (HU). For all reconstructions, total CAC detectability was compared between both CT systems. Significant differences in CAC quantification (Agatston and volume scores) were assessed with Mann Whitney U tests. Furthermore, volume scores were compared with the known CAC physical.

Results

CAC scores for routine clinical protocols were comparable between SPCCT and CT. SPCCT showed 34% and 4% higher detectability of CAC for the small and large phantom, respectively. At reduced slice thickness, CAC detection increased by 142% and 169% for CT and SPCCT, respectively. In comparison with CT, volume scores from SPCCT were more comparable with the physical volume of the CAC.

Conclusion

CAC scores using routine clinical protocols are comparable between conventional CT and SPCCT. The increased spatial resolution of SPCCT allows for increased detectability and more accurate CAC volume estimation.

INTRODUCTION

Spectral photon counting computed tomography (SPCCT) is a novel emerging technology within the field of X-ray diagnostic radiology.[1-7] This technology employs energy discriminating photon counting detectors (PCDs) to detect individual photons in more than 2 energy bins. Due to high photon flux in CT, small pixel detectors are required to allow for individual photons to be counted without pulse pile-up effects.[8-10] In turn, the smaller PCD pixels result in superior spatial resolution in comparison with standard conventional energy integrating detector (EID) CT, which can be a major benefit for assessment of coronary artery calcifications (CAC).[5,11-14]

CAC is traditionally quantified on CT using the Agatston methodology (e.g. 120 peak kilovolt (kVp) acquisition; 3 mm slice thickness reconstruction).[15] Quantification with Agatston scores are recommended by several guidelines to evaluate risk assessment for coronary artery disease.[16-18] The increased in-plane spatial resolution of SPCCT may result in reclassification of risk categories, as partial volume effects are decreased.[19] Especially small and low-density coronary calcifications might not be resolved on current EID CT system. This can potentially lead to the erroneous conclusion of a zero Agatston score, and correspondingly a mis-classification to the lowest risk category. With the increased in-plane spatial resolution of SPCCT, certainty of zero Agatston scores and Agatston score reproducibility can both potentially be increased. Through-plane increased spatial resolution will result in the same advantages, when data is reconstructed at small slice thickness. Furthermore, Agatston scores resulting from larger or higher density CAC can be impacted by this increased spatial resolution as well because of reduced blooming artefacts.

In addition to an increase in spatial resolution, SPCCT also decreases the impact of electronic noise. By setting the lowest energy bin threshold just above the electronic noise signal, the majority of noise can be successfully filtered out.[1,20] This effect reduces the resulting total image noise.[21-23] This feature can potentially be used to acquire and reconstruct data at reduced slice thicknesses, so that both in-plane and through-plane CAC detection can be increased.

Because differences in Agatston scores between CT systems with EID or PCD elements are largely unknown the aim of the current study was twofold. First, the CAC scoring potential of SPCCT in comparison with conventional EID CT for routine clinical protocols was assessed. Second, the potential for improved CAC

9

detection and quantification at reduced slice thickness will be assessed for SPCCT in comparison with EID CT.

MATERIALS AND METHODS

Phantom

An anthropomorphic (cardio)thoracic CT phantom (QRM Thorax, QRM GmbH) in combination with two different cardiac inserts was used. These inserts were a D100 insert and a cardiac calcification insert (CCI, QRM GmbH). Both inserts include cylindrical calcifications composed of hydroxyapatite (HA) powder. The D100 phantom contains 100 small calcifications of different sizes (ranged from 0.5 to 2.0 mm) and densities (ranged from 90 to 540 mgHAcm^{-3}) and was used for the assessment of calcification detectability.[24] The CCI insert consists of nine calcifications with three different amounts of HA (200, 400 and 800 mgHAcm^{-3}) and three different lengths and diameter (1.0, 3.0 and 5.0 mm) for each amount of HA. Additionally, to evaluate the effect of patient size, acquisitions were performed with and without a fat tissue-equivalent extension ring (QRM-Extension ring, QRM) simulating a small and large sized patient, respectively.[25]

Acquisition and reconstruction parameters

Data acquisition was performed on two CT systems from one manufacturer: a dual layer CT (DLCT) (IQon Spectral CT, Philips Healthcare) and a clinical spectral photon counting CT (SPCCT) prototype (SPCCT, Philips Healthcare). The DLCT system was equipped with EID, while the SPCCT system was equipped with novel PCD.[26]

Both devices were equipped with the same X-ray source and had the same source-to-isocenter and source-to-detector distances. Apart from the X-ray detection technology, the size of the detector pixels at iso-center was different between both systems, with 0.625 x 0.625 mm for EID, and 0.275 x 0.275 mm for PCD. Further technical details concerning the prototype system and its performances are provided in previous studies.[27,28]

For both aims of the current study, routine clinical CAC scoring protocols were used for data acquisition and reconstruction (*Table 1*). For SPCCT, acquisition and reconstruction parameters were based on DLCT protocols recommended by the manufacturer. For the second aim, raw data were reconstructed at several

9

Table 1 Acquisition and reconstruction parameters for all used systems for the CAC scoring potential at routine clinical protocols

Parameter	DLCT	SPCCT
CT system	IQon	SPCCT
Technique	Sequential	Sequential
Tube voltage [kVp]	120	120
Tube current time product [mAs]	Small phantom: 40 Large phantom: 80	Small phantom: 40 Large phantom: 80
Automatic exposure correction	Off	Off
Focal spot	Standard	Small[1]
Collimation [mm]	64x0.625	64x0.275
Energy bin threshold [keV]	Not applicable	30 (lower) / 120 (upper)[1]
Field of View [mm]	220	220
Rotation time [s]	0.27	0.33
Slice thickness - increment [mm]	- 0.67 – 0.67 1.0 – 0.5 1.0 – 1.0 3.0 – 1.5 3.0 – 3.0	0.67 – 0.335 0.67 – 0.67 1.0 – 0.5 1.0 – 1.0 3.0 – 1.5 3.0 – 3.0
Reconstruction kernel	IQon-Std-B	SPCCT-Std-B[2]
Reconstruction matrix [pixels]	512x512	512x512
Reconstruction [iDose level]	0 / 3 / 5	0 / 3 / 5 [3]
Repetitions	5	5

1 SPCCT was operated in conventional imaging mode, with only 2 thresholds to either suppress electronic noise (lower threshold) or to suppress pile-up counts (upper threshold)
2 Despite differences in detector element size, reconstruction kernel and reconstruction algorithm for SPCCT, reconstruction parameters for SPCCT were optimized by the manufacturer to get comparable results as with DLCT
3 The small focal spot is the only available option for the current clinical SPCCT prototype

combinations of slice thicknesses and increments, to assess the potential of improved detectability and quantification for both CT systems (*Table 1*). To counteract increased image noise at reduced slice thickness, several iterative reconstruction (IR) levels (iDose[4] algorithm, Philips Healthcare) were added. Each scan was repeated five times, with manual repositioning between each scan (2 mm translation, 2 degrees rotation).

Analysis

General

Agatston scores were determined from the resulting reconstructed images using a previously validated, in-house developed Python script (Python version 3.7).[29] To discriminate calcium-containing voxels from background material, a calcium scoring threshold of 130 HU was used. In addition, in line with the vendor-specific implementation for the Agatston score, a minimum connected area of 0.5 mm^2 was used to include a group of voxels in the Agatston score of a specific calcification. For the used combination of field-of-view (220 mm) and reconstruction matrix (512 x 512), this results in a minimum of three connected voxels. In order to compare CAC quantification with physical volume, the volume score was also determined using the same in-house developed Python script.[30]

In addition to the CAC scores, several image quality metrics were determined. First, mean HU values and noise levels (standard deviation (SD)) of the background material were calculated. Second, mean HU values of the largest calcifications of the CCI insert (5 mm diameter and length) were calculated and compared between both CT systems for the routine clinical protocol. Third, signal-to-noise ratios (SNR) were determined for these same calcifications and reconstructions. SNR was calculated as:

$$SNR = \frac{CAC\ HU_{mean}}{Background\ HU_{SD}}$$

Where CAC HU is the mean attenuation of the CAC, and Background HU$_{SD}$ is the SD of the mean attenuation of the background. Fourth, contrast-to-noise ratios (CNR) were also determined for these calcifications and reconstructions. CNR were calculated as:

$$CNR = \frac{CAC\ HU_{mean} - Background\ HU_{mean}}{Background\ HU_{SD}}$$

With Background HU$_{mean}$ the mean of the attenuation of the background. And fifth, a background Agatston score (BAS) was evaluated for the D100 insert, whereby an Agatston score was calculated in the CAC containing slices with the CAC themselves automatically masked, resulting in a BAS score based on only noise.[29] CAC scores for slices with nonzero BAS were excluded, as it was unknown if actual CAC was measured, or if noise led to an Agatston score (*Supplemental Figure 1*).

Detectability (D100)

Detectability, assessed with the D100 insert, was defined as the ability to determine an Agatston score for a calcification for at least four out of the five repetitions. An Agatston was determined for a calcifications if at least three adjacent (horizontally or vertically) voxels were above the 130HU threshold. For the routine CAC protocol, detectability was assessed using previously described visibility curves.[24] The potential of CAC detection for both CT systems at reduced slice thickness was assessed with the number of detected calcifications.

Quantification (CCI)

For quantification of CAC, evaluated with the CCI insert, median CAC scores and range were calculated from the five repeated measurements. Because DLCT images could not be reconstructed at 0.67 / 0.34 mm slice thickness / increment, comparison between SPCCT and DLCT scores was not possible for this slice thickness and increment. Comparisons with physical volume (98.2 mm^3) were performed for the volume scores obtained with both CT systems.

Statistical analysis

Mean HU, SNR and CNR were compared between DLCT and SPCCT using a Mann-Whitney U signed rank test, with a significance level of $p < 0.05$. Routine CAC protocol agreement between DLCT and SPCCT for Agatston scores was assessed using Bland-Altman plots.[31] Differences in CAC quantification potential between DLCT and SPCCT at reduced slice thickness were assessed on the largest calcifications (5 mm diameter and length). For each combination of slice thickness and increment, CAC scores were compared with the reference (DLCT, reconstructed IR level 0) using a Mann-Whitney U signed rank test, with a significance level of $p < 0.05$.

All statistical analysis were performed with SPSS version 27 (IBM SPSS Statistics, Armonk, New York, United States of America).

RESULTS

Image quality

Background mean CT number and image noise for both phantoms sizes and CT systems is shown in *Table 2* for routine clinical protocols (3/3 mm slice thickness / increment, iDose level 0). Mean image noise was lower for SPCCT in comparison

Table 2 Background mean CT number (median (range)) and image noise (median (range)) for both phantom sizes and both CT systems for routine clinical protocols (3/3 mm slice thickness / increment, iDose level 0.

CT system	Phantom size	Mean	Noise
DLCT	Small	35.9 (35.7-38.6)	15.4 (15.2-16.9)
	Large	49.3 (40.7-82.8)	28.8 (27.7-33.1)
SPCCT	Small	32.6 (32.3-33.0)	14.1 (13.9-14.3)
	Large	27.9 (27.1-28.3)	28.4 (28.1-28.8)

with DLCT, and for the small phantom size in comparison with the large phantom. Mean HU values and SNR for the largest calcifications of the CCI insert were comparable (p > 0.05) between both CT systems (*Figure 1*). Only low density CAC resulted in significantly different (p = 0.008) SNR between both CT systems. A significant increase (p < 0.05) in CNR for SPCCT was shown for the medium and high density CAC. SNR and CNR were, in general, higher for the small phantom size for both CT systems.

Detectability (D100)

Routine CAC protocols
For routine CAC protocols (3/3 mm slice thickness / increment, iDose level 0), representative images for the D100 insert and detectability curves are shown in *Supplemental Figure 2 and 3*, respectively. In comparison with DLCT, more CAC were detected with SPCCT for the small phantom. This effect decreased for increased phantom dimensions.

CAC potential at reduced slice thickness
The percentage of detected CAC, with the total of 500 calcifications (five repetitions of D100 insert) as the denominator, is presented in *Table 3*. In comparison with 3 mm slice thickness and increment, detection of CAC increased, as expected, with overlapping slices and reduced slice thickness for both DLCT and SPCCT. For DLCT, detection increased by 142% from 12.8% to a maximum of 31% detected calcifications for reconstructions with 1 mm slice thickness, 0.5 mm slice increment and IR level 0. At these reconstruction settings, SPCCT CAC detection was even 39% higher. SPCCT CAC detection increased by 169% from 17% to a maximum of 46% detected calcifications for reconstructions with 0.67 mm slice thickness, 0.335 mm slice increment and IR level 3.
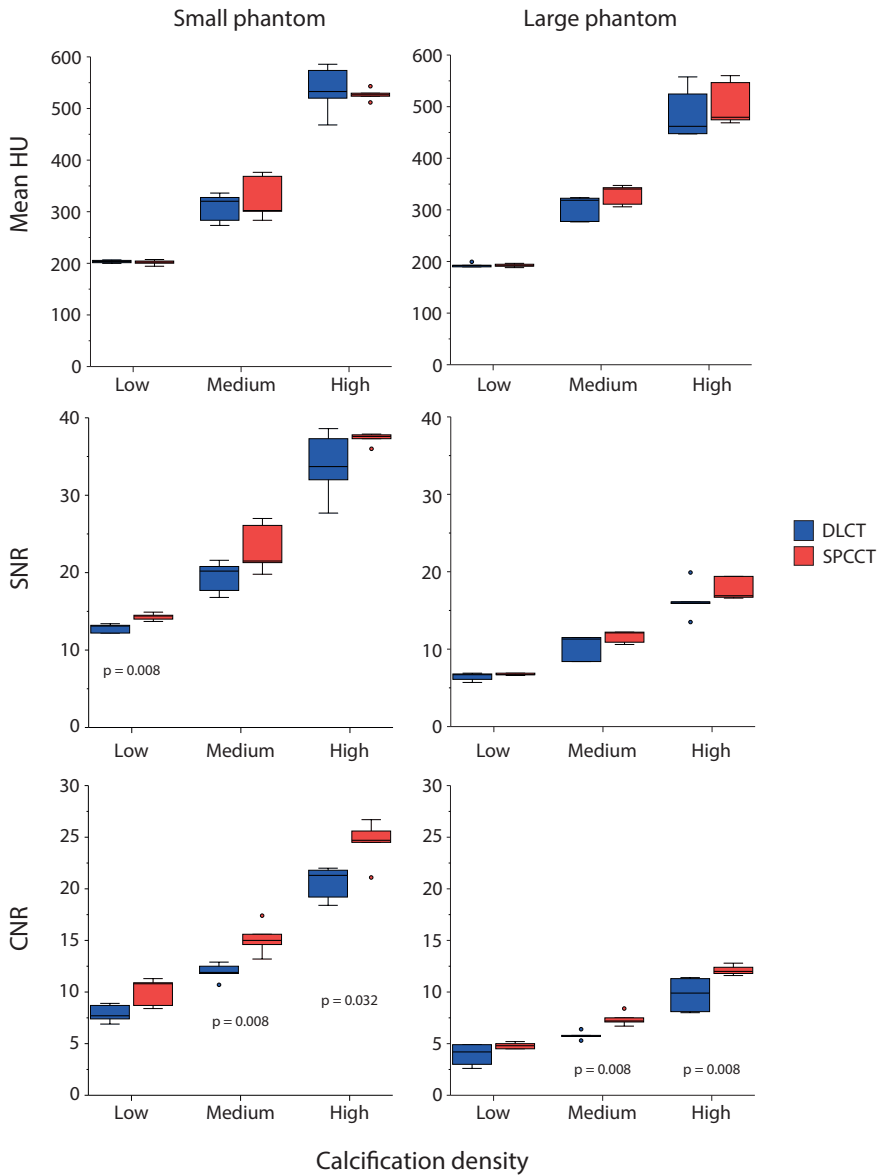
Figure 1 Mean HU, SNR and CNR for the small and large phantom for the large (5 mm diameter and length) calcifications in the CCI insert for both dual-layer CT (DLCT) and spectral photon-counting CT (SPCCT)
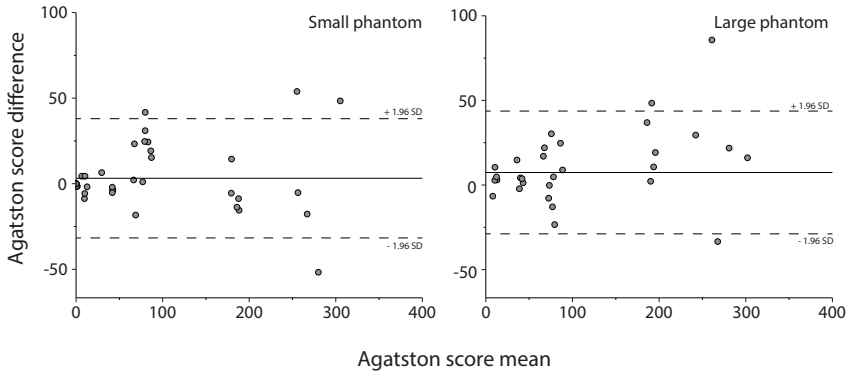
Figure 2 Bland Altman plots for routine CAC protocols for the small (left) and large (right) phantom, comparing dual-layer CT (DLCT) and spectral photon-counting CT (SPCCT) Agatston scores. A positive difference indicates a higher Agatston score for DLCT

9

Figure 3 Agatston scores of the large calcification (5 mm diameter and length) with high density (800 mgHAcm^{-3}), for acquisitions at different combinations of slice thickness and increment, reconstructed with different levels of IR, on both spectral photon-counting CT (SPCCT) and dual-layer CT (DLCT). Results are shown for the small (upper) and large (lower) phantom. For each combination of slice thickness and increment, P-values from significant differences in comparison with the reference (DLCT and iDose 0) are indicated

Table 3 Percentage of detected calcifications, with total of 500 calcifications (five repetitions of D100 insert) as the denominator, for all combinations of slice thickness, slice increment, phantom size and IR level, for both DLCT and SPCCT. Green cells indicate that a system has detected a higher number of calcifications compared to the other system for the same acquisition and reconstruction parameters. Orange indicates that the number of detected calcifications is equal, while red indicates that less calcifications are detected by that system

CT system	Phantom size	IR level	Slice thickness / Slice increment [mm]					
			3/3	3/1.5	1/1	1/0.5	0.67/0.67	0.67/0.335
DLCT	Small	0	12.8%	17.2%	29.4%	31.0%	10.0%	n/a
		3	12.6%	16.6%	27.2%	28.8%	26.2%	n/a
		5	12.6%	16.4%	27.2%	27.6%	25.2%	n/a
	Large	0	14.8%	15.6%	0.0%	0.0%	0.0%	n/a
		3	14.4%	14.6%	1.0%	5.6%	0.0%	n/a
		5	12.6%	13.0%	18.8%	19.4%	0.4%	n/a
SPCCT	Small	0	17.2%	21.4%	40.6%	43.0%	45.4%	43.8%
		3	17.0%	20.8%	38.4%	39.4%	44.4%	46.2%
		5	16.6%	20.2%	36.2%	37.2%	41.8%	44.2%
	Large	0	15.4%	16.4%	0.8%	1.8%	0.0%	0.0%
		3	13.8%	14.4%	32.4%	19.0%	2.2%	0.0%
		5	12.8%	13.6%	28.0%	23.0%	27.2%	25.2%

9

Quantification (CCI)

Routine CAC protocols
Agreement in CAC scores for the CCI insert between DLCT and SPCCT for routine CAC protocols is shown in *Figure 2*. For the small phantom, the mean ± SD difference in Agatston score between both systems was very small at 3.2 ± 17.7. This difference in Agatston score was slightly higher for the large phantom, at 7.4 ± 13.5. Differences increased with increasing Agatston scores.

CAC potential at reduced slice thickness
High density CAC Agatston scores showed significant differences ($p < 0.05$) between DLCT and SPCCT for almost all combinations of slice thickness and increment, irrespective of applied IR level or patient size (*Figure 3*). Low density CAC Agatston scores for the large phantom again show significant differences ($p < 0.05$) between DLCT and SPCCT (*Figure 4*). However, Agatston scores were comparable for the small phantom size, when appropriate IR levels were applied.
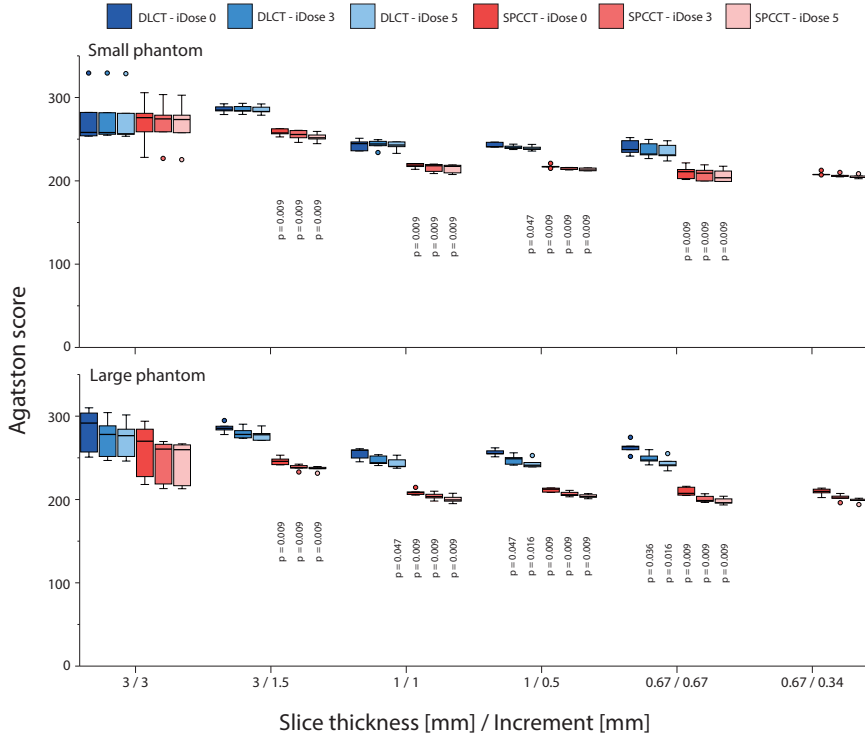
Figure 4 Agatston scores of the large calcification (5 mm diameter and length) with low density (200 mgHAcm^{-3}), for acquisitions at different combinations of slice thickness and increment, reconstructed with different levels of IR, on both spectral photon-counting CT (SPCCT) and dual-layer CT (DLCT). Results are shown for the small (upper) and large (lower) phantom. For each combination of slice thickness and increment, P-values from significant differences in comparison with the reference (DLCT and iDose 0) are indicated

Volume scores showed similar trends as described above for the Agatston score (*Figures 5 and 6*). When compared to the physical volume, high density volume scores showed large overestimations (up to 150%) for all reconstructions. These overestimations decreased at smaller slice thicknesses because of reduced partial volume and blooming artefacts. For all reconstructions, overestimation of physical mass was smaller for SPCCT then for DLCT. Low density volume scores showed better agreement with physical volume. For the large phantom, physical volume was overestimated by DLCT at reduced slice thickness due to noise effects.

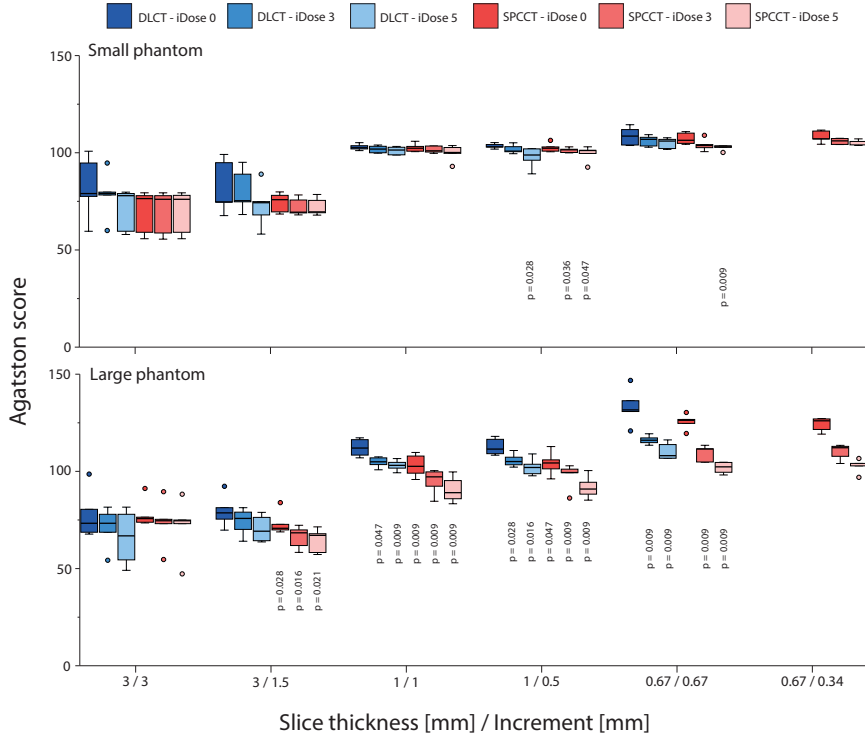***Figure 5*** Volume scores of the large calcification (5 mm diameter and length) with high density (800 mgHAcm^{-3}), for acquisitions at different combinations of slice thickness and increment, reconstructed with different levels of IR, on both spectral photon-counting CT (SPCCT) and dual-layer CT (DLCT). Results are shown for the small (upper) and large (lower) phantom. For each combination of slice thickness and increment, P-values from significant differences in comparison with the reference (DLCT and iDose 0) are indicated. The dashed line indicates the physical volume of the calcification (98.2 mm^3)
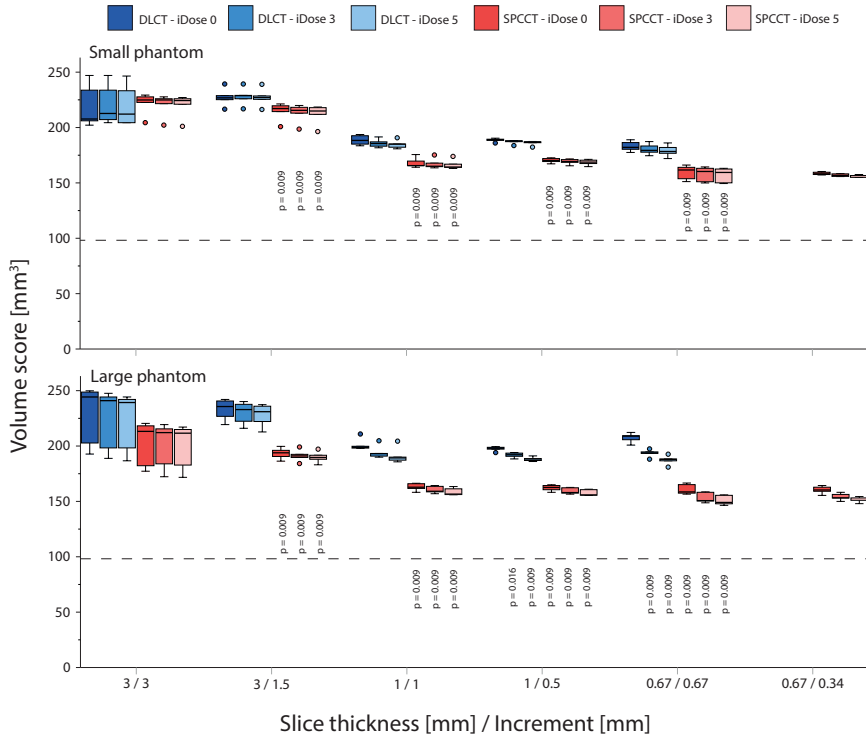
9
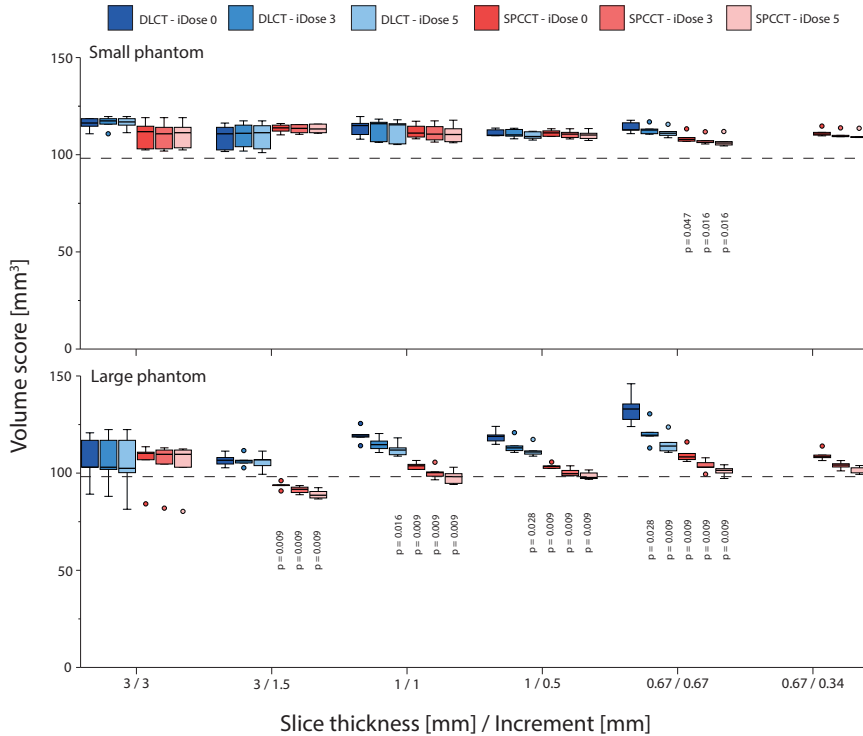
Figure 6 Volume scores of the large calcification (5 mm diameter and length) with low density (200 mgHAcm^{-3}), for acquisitions at different combinations of slice thickness and increment, reconstructed with different levels of IR, on both spectral photon-counting CT (SPCCT) and dual-layer CT (DLCT). Results are shown for the small (upper) and large (lower) phantom. For each combination of slice thickness and increment, P-values from significant differences in comparison with the reference (DLCT and iDose 0) are indicated. The dashed line indicates the physical volume of the calcification (98.2 mm^3)

DISCUSSION

In the present study, we found that SPCCT Agatston scores are comparable with conventional DLCT Agatston scores for routine CAC protocols. Furthermore, we found SPCCT to be more sensitive for detection of CAC at reduced slice thickness acquisitions. Finally, we demonstrated that CAC quantification with SPCCT at reduced slice thickness using volume scores was more accurate than DLCT when compared to the actual physical volume of CAC.

Agatston scores are inherently associated with calcification density, due to the maximum voxel-based weighting factor. In addition, blooming artefacts, including partial volume artefacts, further increase the apparent size of medium and high

density CAC.[32] Also, very small calcifications might potentially be missed due to partial volume effects. This is clinically important because small or low density CAC may be more vulnerable compared to large or high density CAC.[33] One solution to reduce blooming and partial volume artefacts is to increase spatial resolution. In the current study, we have shown that the effect of this increased spatial resolution is only minor for clinical CAC protocols, were only the in-plane resolution was improved, while the slice thickness was still set at 3 mm. This resulted in comparable CAC scores for these protocols on both scanners. For reduced slice thickness and/or overlapping slices, however, significant differences between DLCT and SPCCT were shown. For low-density CAC, the blooming artefact is inherently small. However, for the high-density calcification, reduced blooming artefacts resulted in more accurate CAC scores because of smaller deviations between the volume score and physical CAC volume for SPCCT. Furthermore, increased spatial resolution of SPCCT resulted in increased detectability of small or low density calcifications for SPCCT. Finally, CAC visualization, as determined with SNR and CNR, increased for SPCCT due to reduced image noise in comparison with DLCT. This may be also the effect of a more important energy weighting of the lower energy photons due to the energy-resolving capabilities of the PCDs compared to the EIDs.[1] Altogether, our results are in-line with a recent study by Symons et al, who also showed improved CAC CNR for a different SPCCT system, in comparison with conventional EID CT.[23]

The strength of our study is that we systematically evaluated CAC scoring potential of SPCCT for routine and reduced slice thickness and slice increment, which provides a basis for future research and potential clinical application. In combination with the key findings of previous studies using SPCCT in the cardiovascular field, this modality is an exciting and promising tool for coronary artery disease with potential great expectations for patient management.[6,12,34,35] Our study also has some limitations. First, we used a non-commercial SPCCT system for our evaluation. Second, we used a static anthropomorphic phantom. Despite the fact that the linear attenuation coefficients of the phantoms were in line with human materials at the used tube potential (120 kVp), a phantom does not completely simulate an actual human, with all internal organs. Also, coronary motion was not taken into account. Third, increased noise levels for reduced slice thickness or increased phantom size resulted in BAS > 0. With this, the possibility to assess CAC detectability was reduced, as it was unclear if a group of voxels above the CAC threshold contained CAC or noise. CAC detectability could therefore potentially be further increased, at the cost of increased radiation dose.

Finally, current volume grid parameters were limited to the specifications of the used DLCT. Future studies can assess further improvements for SPCCT, such as other field-of-view and reconstruction matrix combinations, or increased IR strengths as recently reported.[27]

In conclusion, CAC scores using routine clinical protocols are comparable between conventional CT and SPCCT. The increased spatial resolution of SPCCT allows for increased detectability and more accurate CAC volume estimation at reduced slice thickness.
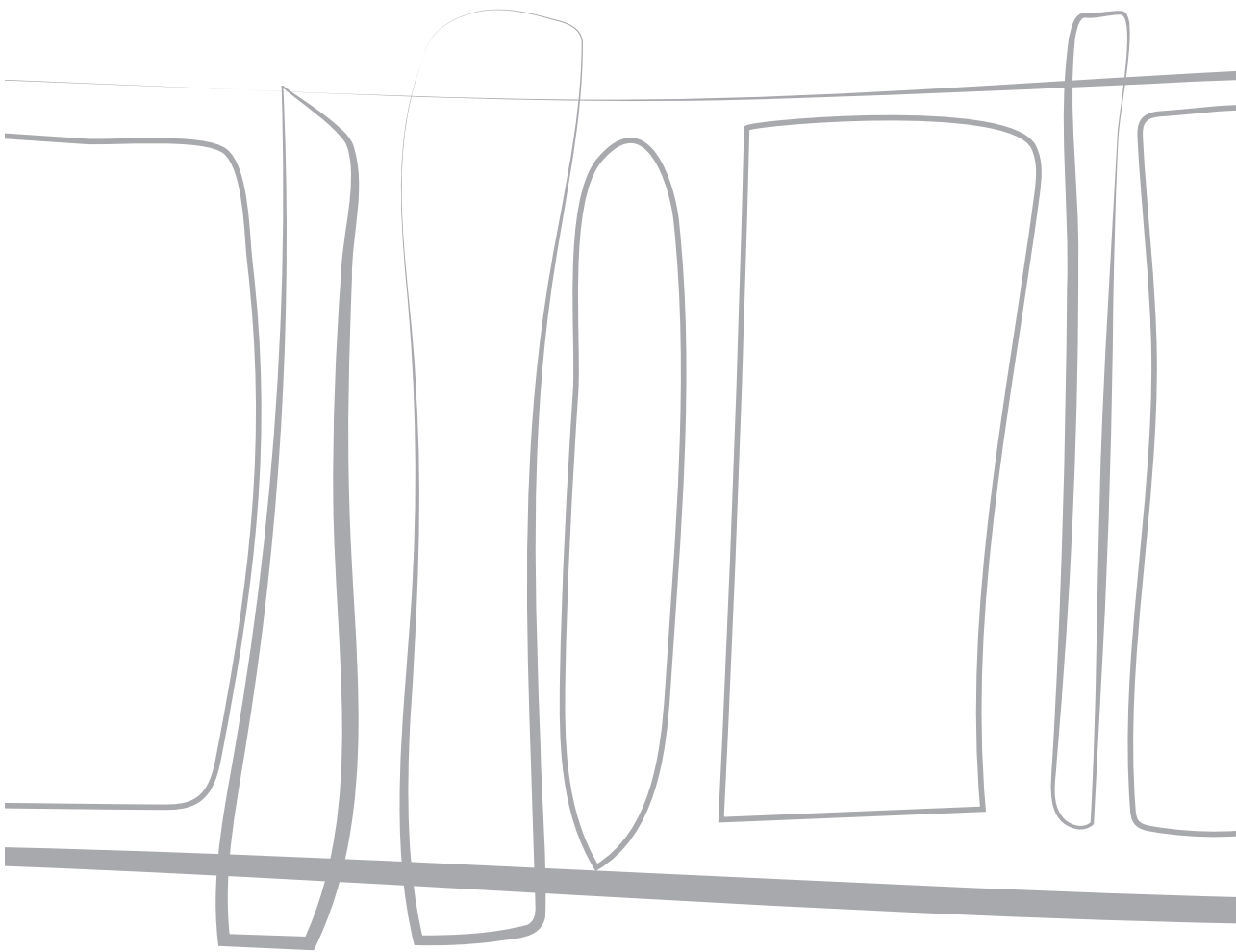
9

REFERENCES

1.   Willemink MJ, Persson M, Pourmorteza A, Pelc NJ, Fleischmann D. Photon-counting CT: Technical principles and clinical prospects. *Radiology*. 2018;289(2):293-312. doi:10.1148/radiol.2018172656

2.   Leng S, Bruesewitz M, Tao S, et al. Photon-counting detector CT: System design and clinical applications of an emerging technology. *Radiographics*. 2019;39(3):729-743. doi:10.1148/rg.2019180115

3.   Si-Mohamed S, Bar-Ness D, Sigovan M, et al. Review of an initial experience with an experimental spectral photon-counting computed tomography system. *Nucl Instruments Methods Phys Res Sect A Accel Spectrometers, Detect Assoc Equip*. 2017;873:27-35. doi:10.1016/j.nima.2017.04.014

4.   Si-Mohamed S, Bar-Ness D, Sigovan M, et al. Multicolour imaging with spectral photon-counting CT: a phantom study. *Eur Radiol Exp*. 2018;2(1). doi:10.1186/s41747-018-0063-4

5.   Taguchi K, Iwanczyk JS. Vision 20/20: Single photon counting x-ray detectors in medical imaging. *Med Phys*. 2013;40(10):100901. doi:10.1118/1.4820371

6.   Sandfort V, Persson M, Pourmorteza A, Noël PB, Fleischmann D, Willemink MJ. Spectral photon-counting CT in cardiovascular imaging. *J Cardiovasc Comput Tomogr*. 2020;(September). doi:10.1016/j.jcct.2020.12.005

7.   Greffier J, Frandon J. Spectral photon-counting CT system: Toward improved image quality performance in conventional and spectral CT imaging. *Diagn Interv Imaging*. Published online February 2021. doi:10.1016/j.diii.2021.02.003

8.   Roessl E, Daerr H, Engel KJ, Thran A, Schirra C, Proksa R. Combined effects of pulse pile-up and energy response in energy-resolved, photon-counting computed tomography. In: *2011 IEEE Nuclear Science Symposium Conference Record*. ; 2011:2309-2313. doi:10.1109/NSSMIC.2011.6153869

9.   Taguchi K, Frey EC, Wang X, Iwanczyk JS, Barber WC. An analytical model of the effects of pulse pileup on the energy spectrum recorded by energy resolved photon counting x-ray detectors. *Med Phys*. 2010;37(8):3957-3969. doi:10.1118/1.3429056

10. Wang AS, Harrison D, Lobastov V, Tkaczyk JE. Pulse pileup statistics for energy discriminating photon counting x-ray detectors. *Med Phys*. 2011;38(7):4265-4275. doi:10.1118/1.3592932

11. Leng S, Gutjahr R, Ferrero A, et al. Ultra-High Spatial Resolution, Multi-Energy CT using Photon Counting Detector Technology. *Proc SPIE Int Soc Opt Eng*. 2017;(12):139-148. doi:10.1016/j.physbeh.2017.03.040

12. Mannil M, Hickethier T, Von Spiczak J, et al. Photon-Counting CT: High-Resolution Imaging of Coronary Stents. *Invest Radiol*. 2018;53(3):143-149. doi:10.1097/RLI.0000000000000420

13. Kopp FK, Daerr H, Si-Mohamed S, et al. Evaluation of a preclinical photon-counting CT prototype for pulmonary imaging. *Sci Rep*. 2018;8(1):1-9. doi:10.1038/s41598-018-35888-1

14. Sigovan M, Si-Mohamed S, Bar-Ness D, et al. Feasibility of improving vascular imaging in the presence of metallic stents using spectral photon counting CT and K-edge imaging. *Sci Rep*. 2019;9(1):1-9. doi:10.1038/s41598-019-56427-6

9

15. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. *J Am Coll Cardiol*. 1990;15(4):827-832. doi:10.1016/0735-1097(90)90282-T

16. Goff DC, Lloyd-Jones DM, Bennett G, et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: A report of the American college of cardiology/American heart association task force on practice guidelines. *Circulation*. 2014;129(25 SUPPL. 1):49-76. doi:10.1161/01.cir.0000437741.48606.98

17. Greenland P, Alpert JS, Beller GA, et al. 2010 ACCF/AHA guideline for assessment of cardiovascular risk in asymptomatic adults: Executive summary: A report of the American College of cardiology foundation/American Heart association task force on practice guidelines. *Circulation*. 2010;122(25):2748-2764. doi:10.1161/CIR.0b013e3182051bab

18. Hecht H, Blaha MJ, Berman DS, et al. Clinical indications for coronary artery calcium scoring in asymptomatic patients: Expert consensus statement from the Society of Cardiovascular Computed Tomography. *J Cardiovasc Comput Tomogr*. 2017;11(2):157-168. doi:10.1016/j.jcct.2017.02.010

19. Blaha MJ, Cainzos-Achirica M, Greenland P, et al. Role of Coronary Artery Calcium Score of Zero and Other Negative Risk Markers for Cardiovascular Disease: The Multi-Ethnic Study of Atherosclerosis (MESA). *Circulation*. 2016;133(9):849-858. doi:10.1161/CIRCULATIONAHA.115.018524

20. Hsieh SS. Design considerations for photon-counting detectors: connecting detectors characteristics to system performances. In: *Spectral, Photon Counting Computed Tomography: Technology and Applications*. ; 2020:326 41.

21. Leng S, Yu Z, Halaweish A, et al. Dose-efficient ultrahigh-resolution scan mode using a photon counting detector computed tomography system. *J Med Imaging*. 2016;3(4):043504. doi:10.1117/1.jmi.3.4.043504

22. Pourmorteza A, Symons R, Henning A, Ulzheimer S, Bluemke DA. Dose Efficiency of Quarter-Millimeter Photon-Counting Computed Tomography: First-in-Human Results. *Invest Radiol*. 2018;53(6):365-372. doi:10.1097/RLI.0000000000000463

23. Symons R, Sandfort V, Mallek M, Ulzheimer S, Pourmorteza A. Coronary artery calcium scoring with photon-counting CT: first in vivo human experience. *Int J Cardiovasc Imaging*. 2019;35(4):733-739. doi:10.1007/s10554-018-1499-6

24. Groen JM, Kofoed KF, Zacho M, Vliegenthart R, Willems TP, Greuter MJW. Calcium score of small coronary calcifications on multidetector computed tomography: Results from a static phantom study. *Eur J Radiol*. 2013;82(2):e58-e63. doi:10.1016/j.ejrad.2012.09.018

25. McCollough CH, Ulzheimer S, Halliburton SS, Shanneik K, White RD, Kalender WA. Coronary Artery Calcium: A Multi-institutional, Multimanufacturer International Standard for Quantification at Cardiac CT. *Radiology*. 2007;243(2):527-538. doi:10.1148/radiol.2432050808

26. Steadman R, Herrmann C, Livne A. ChromAIX2: A large area, high count-rate energy-resolving photon counting ASIC for a Spectral CT Prototype. *Nucl Instruments Methods Phys Res Sect A Accel Spectrometers, Detect Assoc Equip*. 2017;862:18-24. doi:10.1016/j.nima.2017.05.010

9

27. Si-Mohamed S, Boccalini S, Rodesch P-A, et al. Feasibility of lung imaging with a large field-of-view spectral photon-counting CT system. *Diagn Interv Imaging*. Published online February 2021. doi:10.1016/j.diii.2021.01.001

28. Boccalini S, Si-Mohamed S, Dessouky R, Sigovan M, Boussel L, Douek P. Feasibility of human vascular imaging of the neck with a large field-of-view spectral photon-counting CT system. *Diagn Interv Imaging*. Published online January 11, 2021. doi:10.1016/j.diii.2020.12.004

29. Booij R, van der Werf NR, Budde RPJ, Bos D, van Straten M. Dose reduction for CT coronary calcium scoring with a calcium-aware image reconstruction technique: a phantom study. *Eur Radiol*. 2020;30(6):3346-3355. doi:10.1007/s00330-020-06709-9

30. Callister TQ, Cooil B, Raya SP, Lippolis NJ, Russo DJ, Raggi P. Coronary artery disease: improved reproducibility of calcium scoring with an electron-beam CT volumetric method. *Radiology*. 1998;208(3):807-814. doi:10.1148/radiology.208.3.9722864

31. Bland MJ, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res*. 1999;8:135-160.

32. Hoffmann U, Ferencik M, Cury RC, Pena AJ. Coronary CT angiography. *J Nucl Med*. 2006;57(47):797-806. doi:10.1016/j.crvasa.2015.09.008

33. Alluri K, Joshi PH, Henry TS, Blumenthal RS, Nasir K, Blaha MJ. Scoring of coronary artery calcium scans: History, assumptions, current limitations, and future directions. *Atherosclerosis*. 2015;239(1):109-117. doi:10.1016/j.atherosclerosis.2014.12.040

34. Dweck MR, Maurovich-Horvat P, Leiner T, et al. Contemporary rationale for non-invasive imaging of adverse coronary plaque features to identify the vulnerable patient: a Position Paper from the European Society of Cardiology Working Group on Atherosclerosis and Vascular Biology and the European Associ. *Eur Heart J Cardiovasc Imaging*. 2020;21(11):1177-1183. doi:10.1093/ehjci/jeaa201

35. Halttunen N, Lerouge F, Chaput F, et al. Hybrid Nano-GdF3 contrast media allows pre-clinical in vivo element-specific K-edge imaging and quantification. *Sci Rep*. 2019;9(1):1-8. doi:10.1038/s41598-019-48641-z

9
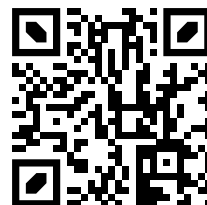
Online material is available via:

CHAPTER 10

Improved coronary calcium detection and quantification with low-dose full field-of-view photon-counting CT: a phantom study

Niels R. van der Werf, MSc*

Pierre-Antoinne Rodesch, MSc*

Salim Si-Mohamed, MD PhD

Robbert W. van Hamersvelt, MD PhD

Marcel J.W. Greuter, PhD

Tim Leiner, MD PhD

Loïc Boussel, PhD

Martin J. Willemink, MD PhD

Philippe Douek, MD PhD

* Authors contributed equally to this work

ABSTRACT

Objective

The aim of the current study was to systematically assess coronary artery calcium (CAC) detection and quantification for spectral photon counting CT (SPCCT) in comparison to conventional CT, and, in addition, to evaluate the possibility of radiation dose reduction.

Methods

Routine clinical CAC CT protocols were used for data acquisition and reconstruction of two CAC containing cylindrical inserts which were positioned within an anthropomorphic thorax phantom. In addition, data was acquired at 50% lower radiation dose by reducing tube current, and slice thickness was decreased. Calcifications were considered detectable when three adjacent voxels exceeded the CAC scoring threshold of 130 Hounsfield units (HU). Quantification of CAC (as volume and mass score) was assessed by comparison with known physical quantities.

Results

In comparison with CT, SPCCT detected 33% and 7% more calcifications for the small and large phantom, respectively. At reduced radiation dose and reduced slice thickness, small phantom CAC detection increased by 108% and 150% for CT and SPCCT, respectively. For the large phantom size, noise levels interfered with CAC detection. Although comparable between CT and SPCCT, routine protocols CAC quantification showed large deviations (up to 134%) from physical CAC volume. At reduced radiation dose and slice thickness, physical volume overestimations decreased to 96% and 72% for CT and SPCCT, respectively. In comparison with volume scores, mass score deviation from physical quantities were smaller.

Conclusion

CAC detection on SPCCT is superior to CT, and was even preserved at reduced radiation dose. Furthermore, SPCCT allows for improved physical volume estimation.

10

INTRODUCTION

All major computed tomography (CT) manufacturers are currently developing spectral photon-counting CT (SPCCT) systems.[1,2] The difference with conventional CT is the fundamentally improved detector technology. Conventional CT uses energy-integrating detectors (EID), while PCCT uses photon-counting detectors (PCD). With PCD, individual photons are counted within predefined energy specific bins characterized by thresholds. This technology reduces the influence of electronic noise by setting the lowest threshold above the electronic noise.[2] Because electronic noise is superimposed on each pulse, when the width of the energy bins are set sufficiently wide, the impact of electronic noise is only minor, reducing the resulting total image noise.[3–5]

With conventional EID, x-ray photons are converted to visual light photons, which may affect neighboring detector pixels. This phenomenon is called optical cross talk, which is limited by the highly reflective septa between EID pixels. With PCD, x-ray photons are directly converted into an electric signal without conversion to visual light photons. Therefore, no reflective septa are needed between PCD pixels, allowing for smaller detector pixels. Furthermore, small detector pixels are required for SPCCT to allow for individual photons to be counted without suffering from pulse pile-up effects.[6–8] This decreased detector pixel size enhances spatial resolution for SPCCT in comparison with conventional CT, both in-plane and through-plane. In turn, increased spatial resolution results in a decrease of partial volume effects and blooming artefacts, which are especially important for high contrast materials such as iodinated contrast, bone, and calcium.

An important use-case for these major improvements in CT technology is coronary artery calcium (CAC) detection and quantification.[9] The highly significant association between CAC as characterized with CT, total coronary atherosclerosis burden, and future adverse cardiovascular events is well known.[10] Moreover, ischemic heart disease caused by coronary plaque remains, according to the World Health Organization, the main cause of death worldwide.[11] Evaluation of CAC detection and quantification with CT imaging are thus recommended in guidelines for clinical risk prediction in appropriately selected asymptomatic individuals, resulting in a high number of CT examinations for CAC assessment.[12–14] In this screening setting, accurate and precise CAC assessment at a low radiation dose is therefore key. Clinically, CAC is assessed according to the Agatston methodology (120 kVp, 3 mm slice thickness, 130 Hounsfield units (HU) CAC threshold).[15] For routine protocols, based on the Agatston methodology, SPCCT outcome is

in agreement with conventional CT.[16] In addition to the Agatston score, newer CAC metrics have been introduced, which are related to the physical measures of volume and mass.[17,18] These metrics have been shown to improve reproducibility of CAC assessment with EID CT.[19–21]

With current conventional CT systems, the accuracy of CAC quantification is affected by blooming artefacts around CAC, which increases inter- and intra-scan variability.[9] The effect of blooming on different CAC densities differs, with underestimation and overestimation of low and high density CAC, respectively. Moreover, due to partial volume averaging effects, small CAC may not be detected, as the conventional Agatston scoring threshold of 130 HU may not be reached. A systematic assessment of the influence of PCD in comparison with EID for CAC detection and quantification of different densities of CAC is lacking.

Furthermore, due to the reduced impact of electronic noise on SPCCT images, data acquisition for CAC assessment could potentially be performed at a reduced radiation dose, while maintaining image quality in comparison with EID.[2,4,22] The combination of SPCCT acquisitions at a reduced radiation dose, in combination with iterative reconstruction (IR) could theoretically further decrease radiation dose burden. Several studies indicated that CT radiation dose reduction with the use of IR did not affect CAC scores compared to routine radiation dose and reconstruction.[23–30] However, for these studies images were acquired with conventional EID CT. The impact of IR on CAC scores for images acquired with PCD remains unknown.

The aim of the current study is therefore twofold. First, we systematically assess CAC detection and quantification for SPCCT in comparison with conventional CT, and second, we evaluate the possibility of dose reduction.

METHODS

Phantom

Anthropomorphic phantoms were used to assess the potential of SPCCT to improve CAC detectability and quantification. We used two setups, which both consisted of an anthropomorphic (cardio)thoracic CT phantom (QRM Thorax, QRM GmbH), in combination with a task-specific cylindrical insert. For the detectability task, the D100 (D100, QRM GmbH) insert was used. This insert contained 100 cylindrical calcifications, divided over four planes, where each plane consisted of a five-by-five matrix of calcifications. The diameter and length of all calcifications were identical,

and ranged from 0.5 to 2.0 mm. The densities of the calcifications ranged from 90 to 540 mg hydroxyapatite (HA) cm^{-3}. The quantification task was assessed with the Cardiac Calcification Insert (CCI, QRM GmbH), which contained three cylindrical (5.0 mm in length and diameter) calcifications with densities of 800, 400, and 200 mgHAcm^{-3}, designated as high, medium and low density, respectively.[17,31] The CCI insert also contained two large calibration rods, consisting of water-equivalent and 200 mgHAcm^{-3} materials, which were used to calculate a mass calibration factor as described by McCollough et al.[17]

To further assess the influence of radiation dose on CAC scoring with SPCCT, two phantom sizes were assessed by adding a tissue-equivalent extension ring (QRM Extension Ring, QRM GmbH) to the anthropomorphic thorax phantom. The resulting dimensions of the phantom were 300 mm x 200 mm without extension ring, and 400 mm x 300 mm with extension ring, which simulates a small and a large sized patient, respectively.[17]

Acquisition and reconstruction parameters

Data acquisition was performed on two CT systems from one manufacturer. First a conventional dual layer CT (DLCT) (IQon Spectral CT, Philips Healthcare) was used, which uses EID for image acquisition. Second, a prototype full field-of-view SPCCT system (Spectral Photon Counting CT, Philips Healthcare) was used, which uses novel PCD technology for image acquisition.

For both phantom sizes, routine clinical CAC protocols recommended by the vendor for conventional CT were used on both CT systems (*Table 1*). In addition to standard reconstruction parameters (3 mm slice thickness and IR level 0), raw data was also reconstructed at 1 mm slice thickness, and with IR levels 3 and 5. Further, to assess the potential of radiation dose reduction, tube currents were reduced by 50% for each phantom size. Each scan was repeated five times, with manual repositioning between each scan (2 mm translation, 2 degrees rotation).

Analysis

For all reconstructions, CAC scores were determined with the use of a previously validated, in-house developed open-source Python script (Python version 3.7).[32,33] To obtain CAC scores which were equal to CAC scores calculated with the vendor-specific software, vendor-specific implementations of both the volume and mass score were used.[15,18]

Table 1 Routine clinical CAC acquisition and reconstruction parameters for DLCT. SPCCT parameters were matched to this as closely as possible

Parameter	DLCT	SPCCT
CT system	IQon	SPCCT
Technique	Sequential	Sequential
Tube voltage [kVp]	120	120
Tube current time product [mAs]	Small phantom: 40[1] Large phantom: 80[1]	Small phantom: 40[1] Large phantom: 80[1]
Automatic exposure correction	Off	Off
Focal spot	Standard	Small[2]
Collimation [mm]	64x0.625	64x0.275
Field of View [mm]	220	220
Rotation time [s]	0.27	0.33
Slice thickness - Increment [mm]	3.0 – 3.0	3.0 – 3.0
Reconstruction kernel	IQon-Std-B	SPCCT-Std-B[3]
Matrix size [pixels]	512x512	512x512
Reconstruction [iDose level]	0	0
Volume CT dose index [mGy]	Small phantom: 4.5 Large phantom: 9.4	Small phantom: 4.0 Large phantom: 8.1

[1] Reference tube current from routine clinical protocol for DLCT

[2] The small focal spot is the only available option for the current clinical SPCCT prototype

[3] Despite differences in detector pixel size, reconstruction kernel and reconstruction algorithm for SPCCT, reconstruction parameters for SPCCT were optimized by the manufacturer to get comparable results as with DLCT

DLCT = Dual layer computed tomography; SPCCT = Spectral photon-counting computed tomography

For both metrics, a CAC scoring threshold of 130 HU was used. Furthermore, only groups of connected voxels with a minimum in-plane area of 0.5 mm^2 were taken into consideration for CAC scoring. For the used reconstruction parameters, with a reconstructed pixel spacing of 0.43 x 0.43 mm^2 (220 mm field-of-view and 512 x 512 matrix), this resulted in a minimum of 3 horizontally or vertically connected voxels. Finally, no interpolation was used for the volume score calculation.

Image noise (standard deviation [SD]) was determined in a large square (128 x 128 voxels) region-of-interest (ROI) of uniform background material. In addition, a false positives analysis was performed, to assess the extent of erroneous detection of CAC in locations where no CAC was present. False positives were defined as groups of voxels > 0.5 mm^2 with CT values exceeding 130 HU in the D100 phantom with the largest and highest density CAC, where the known CAC locations were masked.[32,33] This resulted in a background Agatston score (BAS). Reconstructions

with BAS > 0 were defined as non-diagnostic for CAC detection assessment, as it was uncertain if CAC was detected or if a CAC score was calculated based on only noise. As a result, all repetitions of the D100 insert with BAS > 0 were discarded for CAC detection analysis.

Due to the large size of the CCI calcifications, CAC localization for this insert was not hampered by reconstructions with BAS > 0. Therefore, all reconstructions of the CCI insert were taken into consideration for CAC quantification. Reconstructions with BAS > 0 could result in a slight increase in resulting CAC scores.

For all reconstructions, CAC scores were determined for each of the 100 calcifications of the D100 insert. For each combination of CT system, phantom size, IR level, slice thickness and radiation dose level, detectability was defined as the ability to calculate a CAC score for each individual calcification.

For each calcification of the CCI insert, volume and mass scores were determined. For each combination of CT system, phantom size, IR level, slice thickness, radiation dose level and CAC density, CAC scores were compared with the physical values, a volume of 96.2 mm^3 for all calcifications and a mass of 78.5, 39.3 and 19.6 mg for high, medium, and low-density CAC, respectively. Over- and/or underestimations of physical volume and mass were calculated by dividing the calculated value by the physical value.

RESULTS

Reference volume CT dose index (CTDI$_{vol}$) for the small phantom was 4.5 and 4.0 mGy, for DLCT and SPCCT, respectively. For the large phantom, reference CTDI$_{vol}$ was 9.4 and 8.1 mGy, respectively (*Table 1*). Resulting image noise levels for all radiation dose levels are shown in *Table 2*. Several combinations of phantom size, radiation dose reduction, and IR level resulted in false positive CAC detection (BAS > 0) (*Table 2*). The number of repetitions which were of diagnostic quality for CAC detection is shown in *Table 3*. For these repetitions, BAS did not exceed 0 and it was certain that noise did not lead to false positive CAC detection.

Detectability results (D100-insert)

Representative images and the total number of detected calcifications for both CT systems, phantom sizes, slice thickness and radiation dose levels are shown in *Figures 1* and *2*, respectively. In comparison with DLCT, SPCCT showed an overall superior CAC detectability. For the routine clinical protocol (3 mm slice thickness,

Table 2 Image noise (SD) levels in HU, presented as median (total range) for both DLCT and SPCCT, both phantom sizes, three IR levels and both radiation dose levels. Image noise levels which lead to BAS > 0 for one or more repetitions are indicated in grey

CT system	Phantom size	IR level	3 mm slice thickness		1 mm slice thickness	
			100% dose	50% dose	100% dose	50% dose
DLCT	Small	0	15.4 (15.2-16.9)	22.3 (14.7-23.0)	28.4 (17.7-29.1)	38.8 (24.4-40.1)
		3	8.2 (8.0-13.3)	17.4 (11.5-18.1)	22.2 (13.8-22.7)	30.2 (19.0-31.2)
		5	6.7 (6.7-11.0)	14.3 (9.5-14.9)	18.1 (11.3-18.7)	24.6 (15.5-25.6)
	Large	0	28.8 (27.7-33.1)	41.6 (37.9-43.2)	49.0 (48.4-53.0)	66.7 (66.2-70.2)
		3	22.7 (21.8-27.1)	33.0 (29.7-34.4)	38.1 (37.7-42.1)	52.0 (51.7-55.5)
		5	18.8 (17.9-23.4)	27.6 (24.5-29.3)	31.2 (30.9-35.2)	42.7 (42.2-45.6)
SPCCT	Small	0	14.1 (13.9-14.3)	19.8 (19.2-21.3)	23.5 (23.0-23.9)	33.1 (32.4-37.6)
		3	11.3 (11.1-11.4)	15.8 (15.3-16.9)	18.5 (18.1-18.8)	26.0 (25.5-29.4)
		5	9.5 (9.3-9.6)	13.3 (12.9-14.1)	15.3 (15.0-15.6)	21.6 (21.2-24.2)
	Large	0	28.4 (28.1-28.8)	41.3 (40.5-44.4)	46.4 (46.1-47.3)	70.5 (67.3-78.4)
		3	22.7 (22.5-23.1)	33.0 (32.4-35.3)	37.0 (36.3-37.3)	55.5 (53.0-56.3)
		5	19.1 (18.9-19.5)	27.8 (27.3-29.6)	30.8 (30.1-31.0)	46.4 (44.5-46.7)

DLCT = Dual layer CT; SPCCT = Spectral photon counting CT; BAS = Background Agatston score; IR = iterative reconstruction

10

Table 3 Number of repetitions, out of the total of five, which did not lead to false positives (BAS = 0). These repetitions were considered to be of diagnostic quality for CAC detection purposes

CT system	Phantom size	IR level	3 mm slice thickness		1 mm slice thickness	
			100% dose	50% dose	100% dose	50% dose
DLCT	Small	0	5	5	5	2
		3	5	5	5	5
		5	5	5	5	5
	Large	0	4	3	1	0
		3	5	2	0	0
		5	5	3	0	0
SPCCT	Small	0	5	5	5	5
		3	5	5	5	5
		5	5	5	5	5
	Large	0	5	2	0	0
		3	5	4	5	0
		5	5	5	5	0

DLCT = Dual layer CT; SPCCT = Spectral photon counting CT; BAS = Background Agatston score; IR = iterative reconstruction
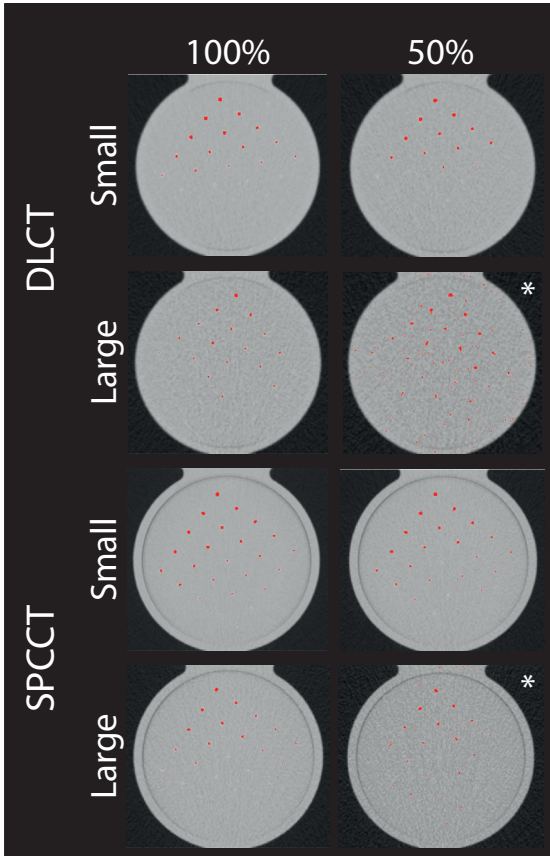
Figure 1 Representative images for the D100 insert placed in the small and large phantom, for DLCT and SPCCT for routine clinical protocols (100% radiation dose, IR level 0 and 3 mm slice thickness) and 50% reduced radiation dose. Voxels exceeding the CAC scoring threshold of 130 HU are indicated in red. An asterisk indicates reconstructions for which BAS > 0

10

100% radiation dose and IR level 0), DLCT detected a median (range) number of 12 (11-16) and 14 (13-16) out of 100 calcifications for the small and large phantom, respectively. For SPCCT, CAC detection increased to 16 (14-23) and 15 (14-17), respectively.

At a reduced slice thickness of 1 mm, the median number of detected calcifications in the small phantom increased by 142% for DLCT and by 156% for SPCCT. For the large phantom, reduced slice thickness resulted in non-diagnostic image quality for CAC detection (BAS > 0) for DLCT for all reconstructions. For SPCCT, this was only the case when IR level 0 was used. However, changing IR level to 3 enabled satisfying image quality and increased median CAC detection by 127% for the large phantom in comparison with the routine clinical protocol.
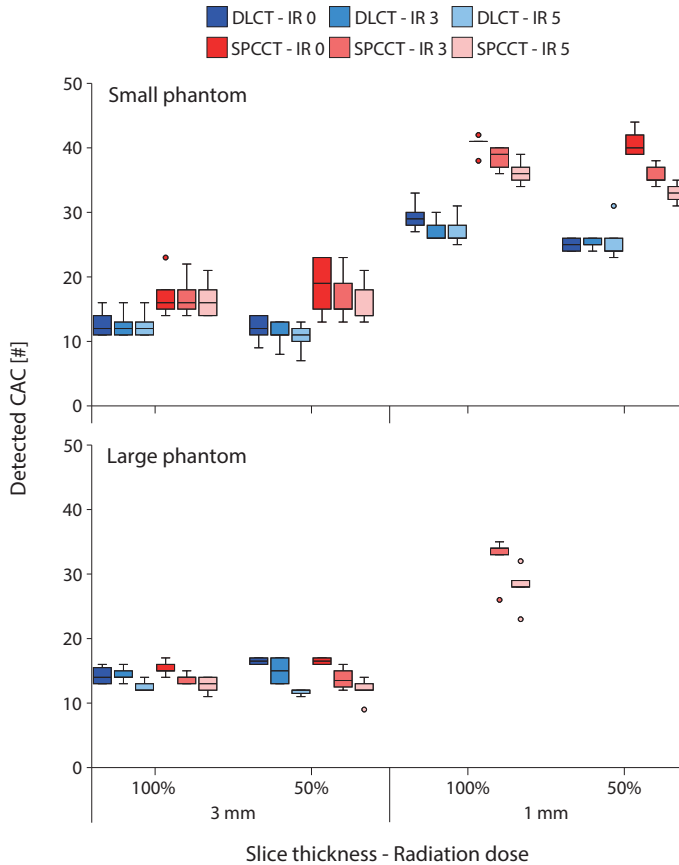
Figure 2 Box and whisker for the total number of detected CAC from the 100 present CAC in the D100 insert for the small (upper) and large (lower) phantom size. Results are shown for both used slice thicknesses and radiation dose levels

At both reduced slice thickness and reduced radiation dose, noise levels were too high for the large phantom for both DLCT and SPCCT with all IR levels, resulting in non-diagnostic image quality for CAC detection (BAS > 0). For the small phantom, however, median CAC detection was still increased in comparison with 100% dose and 3 mm slice thickness when appropriate IR levels were selected (*Figure 2*). For DLCT at 50% dose and 1 mm slice thickness with IR level 0, the median (range) number of detected calcifications was 25 (24-26), which is equal to an increased CAC detection of 108% in comparison with the routine clinical protocol. For SPCCT with IR level 0, the median (range) number of detected calcifications was 40 (39-44), which is equal to an increased CAC detection of 150% in comparison with the routine clinical protocol.

Quantification results (CCI-insert)

Routine clinical CAC protocols (100% dose, 3 mm slice thickness and IR level 0) resulted in comparable volume scores for the CCI insert between DLCT and SPCCT, independent of phantom size or CAC density (*Figures 3* and *4*). However, especially for the high-density calcifications, large deviations from the physical volume of the calcifications were shown. For DLCT, small phantom volume scores from routine clinical protocols overestimated physical volume by (median (range)) 116% (110% to 157%), 62% (56% to 106%), and 21% (15% to 23%) for high, medium, and low CAC density, respectively. For SPCCT, small phantom volume scores overestimated physical CAC volume for routine clinical protocols by 134% (131% to 138%), 93% (50% to 96%), and 16% (7% to 24%) for high, medium, and low CAC density, respectively.

10

Figure 3 Box and whisker plots of the volume scores for CCI insert CAC. Results are shown for small (upper) and large (lower) phantom dimensions, high (A and D) / medium (B and E) / low (C and F) CAC density, both CT systems, three IR levels, two slice thicknesses and both radiation dose levels. The dashed line indicates the total physical volume of the calcifications (96.2 mm3)
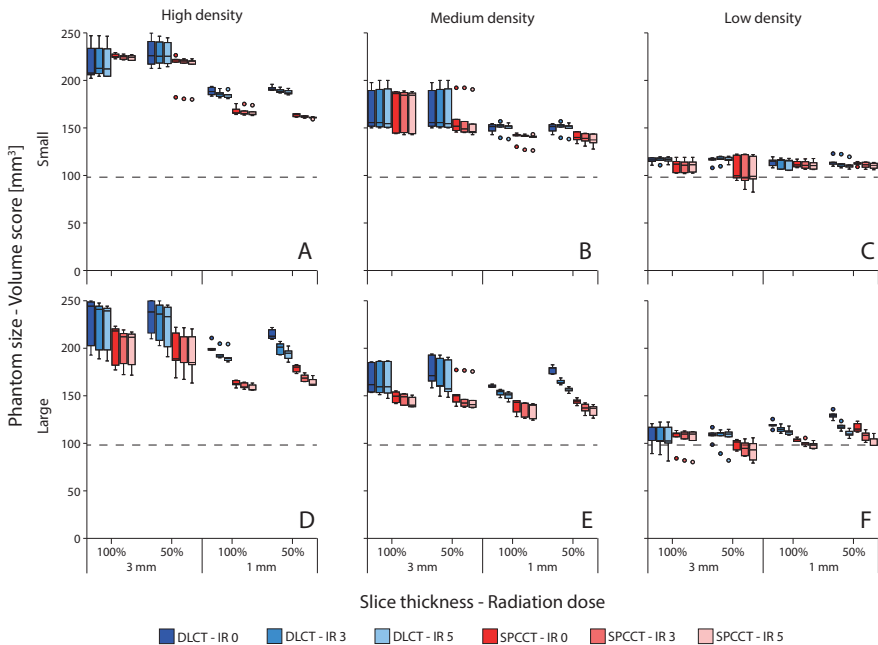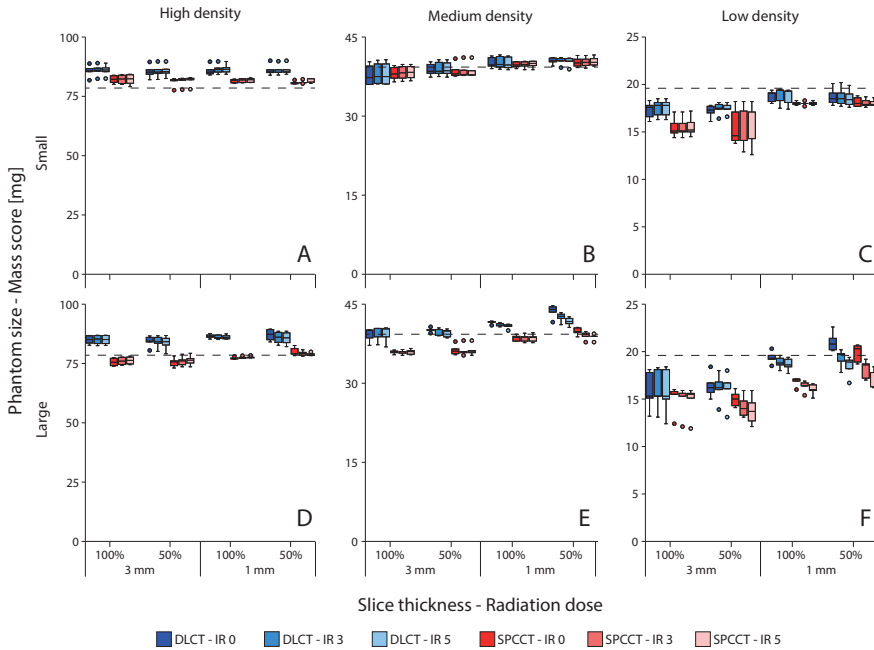
Figure 4 Box and whisker plots of the mass scores for CCI insert CAC. Results are shown for small (upper) and large (lower) phantom dimensions, high (A and D) / medium (B and E) / low (C and F) CAC density, both CT systems, three IR levels, two slice thicknesses and both radiation dose levels. The dashed line indicates the total physical mass of the calcifications (78.5, 39.3 and 19.6 mg for high, medium and low density, respectively)

In comparison to volume scores, deviations from physical quantities were in general smaller for mass scores (*Figure 4*). For DLCT, small phantom mass scores from routine clinical protocols deviated from physical mass by (median (range)) 9% (4% to 13%), -5% (-8% to 2%), and -10% (-18% to -7%) for high, medium, and low CAC density, respectively. For SPCCT, small phantom mass scores deviated from physical CAC mass by 5% (2% to 7%), -3% (-7% to -1%), and -23% (-27% to -13%) for high, medium, and low CAC density, respectively.

The influence of radiation dose, IR levels and phantom size on volume approximation was only minor (*Figure 3*). However, reconstructions at reduced slice thickness reduced partial volume effects, and therefore improved volume scores substantially, in particular for high density CAC and SPCCT. For DLCT at routine dose, 1 mm slice thickness and IR level 0, small phantom volume scores were overestimated by 96% (91% to 101%), 57% (49% to 61%), and 20% (12% to 24%), again for high, medium, and low CAC density, respectively. For SPCCT, this

10

volume overestimation was 72% (71% to 83%), 48% (35% to 50%) and 16% (12% to 22%). In comparison with DLCT, SPCCT showed superior physical volume depiction for all CAC densities at reduced slice thickness.

Physical mass approximation improved for reduced slice thickness for both CT systems and all CAC densities (*Figure 4*). For both CT systems, the largest decrease in physical mass deviation was shown for the low-density CAC. DLCT small phantom deviation from physical mass at 1 mm slice thickness, 100% radiation dose and IR level 0 in median (range) was 8% (7% to 14%), 1% (-1% to 5%), and -3% (-8% to -1%) for high, medium, and low CAC density, respectively. SPCCT mass scores deviated from physical mass by 4% (3% to 5%), 1% (-1% to 3%), and -8% (-9% to -7%) for high, medium, and low CAC density, respectively. Changes in physical mass approximation for changes in radiation dose for both CT systems, slice thickness and IR were only minor.

DISCUSSION

The current study shows superior CAC detectability for SPCCT in comparison with DLCT. SPCCT shows improved CAC detection (up to 156%) at reduced slice thickness, even at 50% radiation dose. SPCCT can more accurately measure physical volumes, especially at reduced slice thickness and for high-density CAC. Both DLCT and SPCCT show more accurate physical mass at reduced slice thickness.

Detection of small and low density calcifications is clinically relevant, due to the important role of zero CAC scores for the risk estimation of cardiovascular disease.[34] The sensitivity of CT for the detection of small or low-density CAC can be increased by thinner reconstructed slices. However, thinner slices yield an increase in image noise, when radiation dose is kept constant. This can be counteracted by the use of increased levels of IR.[35,36] The application of IR can, however, impact CAC quantification as CAC can be removed from the image.[25,37–40] In the current study, CAC detection using SPCCT improved by up to 141% with reduced slice thickness at only 50% of the clinical radiation dose level. For the large phantom size, increased CAC detection of 113% was shown for reduced slice thickness at clinical radiation dose at IR level 3. For both phantom sizes, physical volume approximation improved for large CAC.

To the best of our knowledge, our study is the first to systematically assess CAC detection and quantification at reduced slice thickness and reduced radiation

10

dose for SPCCT systems. Other studies did however assess CAC quantification on SPCCT. Our results corroborate the recent study by Sandstedt et al, who also showed improved CAC volume quantification for SPCCT at standard radiation dose.[9] In that study, however, a SPCCT system with limited FOV was used, and radiation dose reduction was not applied. Also, the exact densities of the used CAC were unknown. Our results are also in-line with a recent publication by Symons et a, who showed improved CAC CNR for SPCCT, which could potentially reduce CAC score radiation dose while maintaining diagnostic image quality.[5]

The systematic nature of our study provides insight in CAC detection and the quantification potential of SPCCT for different densities, radiation dose levels, slice thicknesses, and IR levels. Our study also has some limitations. First, a non-commercial SPCCT system was used for our evaluation. Second, our study was based on static phantom data since the rotation gantry rotation time of the used prototype SPCCT system is not optimized yet. Despite the anthropomorphic nature of the phantom with in-vivo linear attenuation coefficients, coronary motion and complex internal structures were not taken into account. This could be assessed with a dynamic anthropomorphic phantom in a follow-up study on the next version of the prototype. Using a static phantom did, however, provide us with the opportunity to systematically assess CAC detection, without any influence of motion artefacts. Finally, SPCCT parameters for this study were only based on clinical conventional CT values. Further improvements for SPCCT CAC detection and quantification are likely for SPCCT specific protocol optimizations.

In conclusion, CAC detection on SPCCT is superior to DLCT, and was even preserved at reduced radiation dose. Furthermore, SPCCT allows for improved physical volume approximation.

10

REFERENCES

1. Sandfort V, Persson M, Pourmorteza A, Noël PB, Fleischmann D, Willemink MJ. Spectral photon-counting CT in cardiovascular imaging. *J Cardiovasc Comput Tomogr*. 2020;(September). doi:10.1016/j.jcct.2020.12.005

2. Willemink MJ, Persson M, Pourmorteza A, Pelc NJ, Fleischmann D. Photon-counting CT: Technical principles and clinical prospects. *Radiology*. 2018;289(2):293-312. doi:10.1148/radiol.2018172656

3. Leng S, Yu Z, Halaweish A, et al. Dose-efficient ultrahigh-resolution scan mode using a photon counting detector computed tomography system. *J Med Imaging*. 2016;3(4):043504. doi:10.1117/1.jmi.3.4.043504

4. Pourmorteza A, Symons R, Henning A, Ulzheimer S, Bluemke DA. Dose Efficiency of Quarter-Millimeter Photon-Counting Computed Tomography: First-in-Human Results. *Invest Radiol*. 2018;53(6):365-372. doi:10.1097/RLI.0000000000000463

5. Symons R, Sandfort V, Mallek M, Ulzheimer S, Pourmorteza A. Coronary artery calcium scoring with photon-counting CT: first in vivo human experience. *Int J Cardiovasc Imaging*. 2019;35(4):733-739. doi:10.1007/s10554-018-1499-6

6. Wang AS, Harrison D, Lobastov V, Tkaczyk JE. Pulse pileup statistics for energy discriminating photon counting x-ray detectors. *Med Phys*. 2011;38(7):4265-4275. doi:10.1118/1.3592932

7. Leng S, Gutjahr R, Ferrero A, et al. Ultra-High Spatial Resolution, Multi-Energy CT using Photon Counting Detector Technology. *Proc SPIE Int Soc Opt Eng*. 2017;(12):139-148. doi:10.1016/j.physbeh.2017.03.040

8. Mannil M, Hickethier T, Von Spiczak J, et al. Photon-Counting CT: High-Resolution Imaging of Coronary Stents. *Invest Radiol*. 2018;53(3):143-149. doi:10.1097/RLI.0000000000000420

9. Sandstedt M, Marsh J, Rajendran K, et al. Improved coronary calcification quantification using photon-counting-detector CT: an ex vivo study in cadaveric specimens. *Eur Radiol*. 2021;31(9):6621-6630. doi:10.1007/s00330-021-07780-6

10. Budoff MJ, Achenbach S, Blumenthal RS, et al. Assessment of coronary artery disease by cardiac computed tomography: A scientific statement from the American Heart Association Committee on Cardiovascular Imaging and Intervention, Council on Cardiovascular Radiology and Intervention, and Committee on C. *Circulation*. 2006;114(16):1761-1791. doi:10.1161/CIRCULATIONAHA.106.178458

11. WHO Fact Sheet - The Top 10 Causes of Death. May 2018. https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

12. Greenland P, Alpert JS, Beller GA, et al. 2010 ACCF/AHA guideline for assessment of cardiovascular risk in asymptomatic adults: Executive summary: A report of the American College of cardiology foundation/American Heart association task force on practice guidelines. *Circulation*. 2010;122(25):2748-2764. doi:10.1161/CIR.0b013e3182051bab

13. Blaha MJ, Mortensen MB, Kianoush S, Tota-Maharaj R, Cainzos-Achirica M. Coronary Artery Calcium Scoring: Is It Time for a Change in Methodology? *JACC Cardiovasc Imaging*. 2017;10(8):923-937. doi:10.1016/j.jcmg.2017.05.007
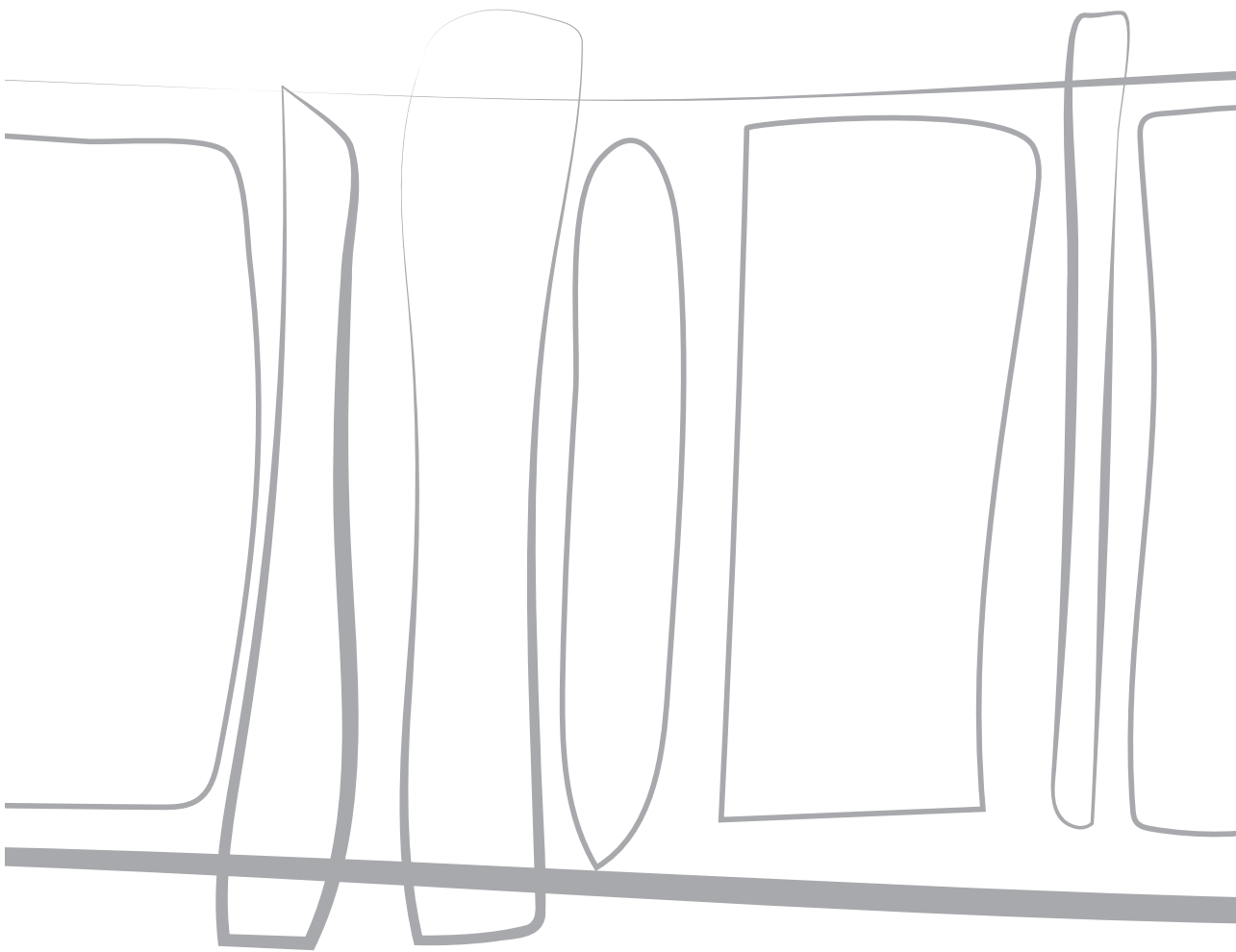
10

14. Hecht H, Blaha MJ, Berman DS, et al. Clinical indications for coronary artery calcium scoring in asymptomatic patients: Expert consensus statement from the Society of Cardiovascular Computed Tomography. *J Cardiovasc Comput Tomogr*. 2017;11(2):157-168. doi:10.1016/j.jcct.2017.02.010

15. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. *J Am Coll Cardiol*. 1990;15(4):827-832. doi:10.1016/0735-1097(90)90282-T

16. van der Werf NR, Si-Mohamed S, Rodesch PA, et al. Coronary calcium scoring potential of large field-of-view spectral photon-counting CT: a phantom study. *Eur Radiol*. Published online 2021. doi:10.1007/s00330-021-08152-w

17. McCollough CH, Ulzheimer S, Halliburton SS, Shanneik K, White RD, Kalender WA. Coronary Artery Calcium: A Multi-institutional, Multimanufacturer International Standard for Quantification at Cardiac CT. *Radiology*. 2007;243(2):527-538. doi:10.1148/radiol.2432050808

18. Callister TQ, Cooil B, Raya SP, Lippolis NJ, Russo DJ, Raggi P. Coronary artery disease: improved reproducibility of calcium scoring with an electron-beam CT volumetric method. *Radiology*. 1998;208(3):807-814. doi:10.1148/radiology.208.3.9722864

19. Hoffmann U, Siebert U, Bull-Stewart A, et al. Evidence for lower variability of coronary artery calcium mineral mass measurements by multi-detector computed tomography in a community-based cohort--consequences for progression studies. *Eur J Radiol*. 2006;57(3):396-402. doi:10.1016/j.ejrad.2005.12.027

20. Groen JM, Greuter MJ, Schmidt B, Suess C, Vliegenthart R, Oudkerk M. The influence of heart rate, slice thickness, and calcification density on calcium scores using 64-slice multidetector computed tomography: A systematic phantom study. *Invest Radiol*. 2007;42(12):848-855. doi:10.1097/RLI.0b013e318154c549

21. van Ooijen PMA, Vliegenthart R, Witteman JC, Oudkerk M. Influence of scoring parameter settings on Agatston and volume scores for coronary calcification. *Eur Radiol*. 2005;15(1):102-110. doi:10.1007/s00330-004-2479-x

22. Ren L, Rajendran K, McCollough CH, Yu L. Radiation dose efficiency of multi-energy photon-counting-detector CT for dual-contrast imaging. *Phys Med Biol*. 2019;64(24):245003. doi:10.1088/1361-6560/ab55bf

23. Vonder M, Pelgrim GJ, Meyer M, Henzler T, Oudkerk M, Vliegenthart R. Dose reduction techniques in coronary calcium scoring: The effect of iterative reconstruction combined with low tube voltage on calcium scores in a thoracic phantom. *Eur J Radiol*. 2017;93(January):229-235. doi:10.1016/j.ejrad.2017.06.001

24. Vonder M, van der Werf NR, Leiner T, et al. The impact of dose reduction on the quantification of coronary artery calcifications and risk categorization: A systematic review. *J Cardiovasc Comput Tomogr*. 2018;12(5):352-363. doi:10.1016/j.jcct.2018.06.001

25. van der Werf NR, Willemink MJ, Willems TP, Greuter MJW, Leiner T. Influence of dose reduction and iterative reconstruction on CT calcium scores: a multi-manufacturer dynamic phantom study. *Int J Cardiovasc Imaging*. 2017;33(6):899-914. doi:10.1007/s10554-017-1061-y

10

26. Ode S, Kobayashi Y, Nozu Y, Ogawa Y. The Impact of Iterative Reconstruction on Reducing the Radiation Dose for Coronary Calcium Scoring : An Investigation Using Pulsating Calcified Coronary Phantom. 2016;7:95-103.

27. Schindler A, Vliegenthart R, Schoepf UJ, et al. Iterative Image Reconstruction Techniques for CT Coronary Artery Calcium Quantification: Comparison with Traditional Filtered Back Projection in Vitro and in Vivo. *Radiology*. 2014;270(2):387-393. doi:10.1148/radiol.13130233

28. Willemink MJ, den Harder AM, Foppen W, et al. Finding the optimal dose reduction and iterative reconstruction level for coronary calcium scoring. *J Cardiovasc Comput Tomogr*. 2016;10(1):69-75. doi:10.1016/j.jcct.2015.08.004

29. den Harder AM, Wolterink JM, Willemink MJ, et al. Submillisievert coronary calcium quantification using model-based iterative reconstruction: A within-patient analysis. *Eur J Radiol*. 2016;85(11):2152-2159. doi:10.1016/j.ejrad.2016.09.028

30. Sulaiman N, Soon J, Park J kwan, et al. Comparison of low-dose coronary artery calcium scoring using low tube current technique and hybrid iterative reconstruction vs. filtered back projection. *Clin Imaging*. 2017;43(November 2015):19-23. doi:10.1016/j.clinimag.2017.01.017

31. McCollough CH, Primak AN, Saba O, et al. Dose performance of a 64-channel dual-source CT scanner. *Radiology*. 2007;243(3):775-784. doi:10.1148/radiol.2433061165

32. Booij R, van der Werf NR, Budde RPJ, Bos D, van Straten M. Dose reduction for CT coronary calcium scoring with a calcium-aware image reconstruction technique: a phantom study. *Eur Radiol*. 2020;30(6):3346-3355. doi:10.1007/s00330-020-06709-9

33. van Praagh GD, van der Werf NR, Wang J, et al. Fully Automated Quantification Method (FQM) of Coronary Calcium in an Anthropomorphic Phantom. *Med Phys*. Published online 2021:in press. doi:10.1002/mp.14912

34. Blaha MJ, Cainzos-Achirica M, Greenland P, et al. Role of Coronary Artery Calcium Score of Zero and Other Negative Risk Markers for Cardiovascular Disease: The Multi-Ethnic Study of Atherosclerosis (MESA). *Circulation*. 2016;133(9):849-858. doi:10.1161/CIRCULATIONAHA.115.018524

35. Willemink MJ, Takx RAP, De Jong PA, et al. The impact of CT radiation dose reduction and iterative reconstruction algorithms from four different vendors on coronary calcium scoring. *Eur Radiol*. 2014;24(9):2201-2212. doi:10.1007/s00330-014-3217-7

36. Takx RAP, Willemink MJ, Nathoe HM, et al. The effect of iterative reconstruction on quantitative computed tomography assessment of coronary plaque composition. *Int J Cardiovasc Imaging*. 2014;30(1):155-163. doi:10.1007/s10554-013-0293-8

37. Kurata A, Dharampal A, Dedic A, et al. Impact of iterative reconstruction on CT coronary calcium quantification. *Eur Radiol*. 2013;23(12):3246-3252. doi:10.1007/s00330-013-3022-8

38. Takahashi M, Kimura F, Umezawa T, Watanabe Y, Ogawa H. Comparison of adaptive statistical iterative and filtered back projection reconstruction techniques in quantifying coronary calcium. *J Cardiovasc Comput Tomogr*. 2016;10(1):61-68. doi:10.1016/j.jcct.2015.07.012

39. Oda S, Utsunomiya D, Nakaura T, et al. The Influence of Iterative Reconstruction on Coronary Artery Calcium Scoring—Phantom and Clinical Studies. *Acad Radiol*. 2017;24(3):295-301. doi:10.1016/j.acra.2016.11.003

10

40. van der Werf NR, Willemink MJ, Willems TP, Greuter MJW, Leiner T. Influence of iterative reconstruction on coronary calcium scores at multiple heart rates: a multivendor phantom study on state-of-the-art CT systems. *Int J Cardiovasc Imaging.* 2017;0(0):1-11. doi:10.1007/s10554-017-1292-y

10
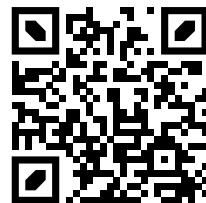
Online material is available via:

CHAPTER 11

Dose reduction for CT coronary calcium scoring with a calcium-aware image reconstruction technique: a phantom study

Ronald Booij, PhD

Niels R. van der Werf, MSc

Ricardo P.J. Budde, MD PhD

Daniel Bos, MD PhD

Marcel van Straten, PhD

ABSTRACT

Objective

To assess the dose reduction potential of a calcium-aware reconstruction technique, which aims at tube voltage-independent computed tomography (CT) numbers for calcium.

Methods

A cardiothoracic phantom, mimicking different patient sizes, was scanned with two calcium inserts (named D100 and CCI), containing calcifications varying in size and density. Tube voltage was varied both manually (range 70 – 150 and Sn100 kVp) and automatically. Tube current was automatically adapted to maintain reference image quality defined at 120 kVp. Data was reconstructed with the standard reconstruction technique (kernel Qr36) and the calcium-aware reconstruction technique (kernel Sa36). We assessed the radiation dose reduction potential (volumetric CT dose index values (CTDI$_{vol}$)), noise (standard deviation (SD)), mean CT number (HU) of each calcification, and Agatston scores for varying kVp. Results were compared with the reference acquired at 120 kVp and reconstructed with Qr36.

Results

Automatic selection of the optimal tube voltage resulted in a CTDI$_{vol}$ reduction of 22%, 15%, and 12% compared with the reference for the small, medium, and large phantom, respectively. CT numbers differed up to 64% for the standard reconstruction and 11% for the calcium-aware reconstruction. Similarly, Agatston scores deviated up to 40% and 8% for the standard and calcium-aware reconstruction technique, respectively.

Conclusion

CAC scores using routine clinical protocols are comparable between conventional CT and SPCCT. The increased spatial resolution of SPCCT allows for increased detectability and more accurate CAC volume estimation.

INTRODUCTION

Ischemic heart diseases remain one of the leading causes of death worldwide.[1,2] Within the framework of individual risk prediction for these diseases, the assessment of coronary artery calcium has become increasingly important. Currently, the most common strategy for quantification of the coronary artery calcium score (CACS) is on computed tomography (CT) examinations using the Agatston method.[3] Despite the excellent prognostic value of this CT-based strategy, the Agatston scoring method has some limitations.[4,5] Recent guidelines demand a fixed tube voltage of 120 peak kilo voltage (kVp) in combination with filtered back projection (FBP) or iterative reconstruction with 100 kVp acquisition after site- and literature-based validation.[5,6] However, there is a main argument for the use of lower, or even patient-specific, tube voltages: the need to reduce radiation dose given the increase in the number of CT examinations.[7]

Lowering tube voltage potentially reduces radiation dose in CACS at the cost of inconsistent scores because CT numbers, expressed in Hounsfield units (HU), are energy dependent. In this case, the standard calcium scoring threshold should be made tube voltage or patient-specific.

Recently, a calcium-aware reconstruction technique was introduced via the application of a new reconstruction kernel (Sa36f). The technique is also known by the name "Agatston score equivalent calcium scoring," "artificial 120 kV equivalent CT images," or "artificial 120." Please refer to the vendor's whitepaper for a detailed explanation.[8] With this technique, CT numbers of calcium are scaled to match the CT numbers that would have been measured at 120 kVp, enabling the use of the standard 130 HU threshold.[9] The technique might enable acquiring images at reduced radiation dose, while preserving the Agatston score and its risk assessment potential. In contrast to tube voltage–dependent threshold adjustments, the calcium-aware reconstruction technique seems an easy tool to implement clinically.

The purpose of our phantom study was to evaluate the calcium-aware reconstruction technique with regard to coronary calcium quantification for a wide range of tube voltages and calcifications varying in size and density and for different chest sizes. Moreover, the radiation dose reduction by automatic tube voltage selection was assessed for these cases.

11

METHODS

Phantom

An anthropomorphic (cardio) thoracic CT phantom (QRM Thorax, QRM GmbH) in combination with two different inserts was used for quantitative assessment of CACS both for the standard and the calcium-aware reconstruction technique. One insert (D100, QRM GmbH) contained 100 calcifications of different diameters (0.5 to 2.0 mm) and hydroxyapatite (HA) densities (90 to 540 mg HA/cm³).[10] The other insert was a cylindrical cardiac calcification insert (CCI, QRM GmbH) with nine calcifications varying in size (1.0 to 5.0 mm) and density (200 to 800 mg HA/cm³). To simulate different chest sizes, the thorax phantom was scanned with and without fat-equivalent extension rings (QRM GmbH) resulting in three different chest sizes: small (300 × 200 mm), medium (350 × 250 mm), and large (400 × 300 mm). To ensure a realistic translation of the results from different phantom sizes to human chest sizes, the water equivalent diameter (Dw) was used. Dw reflects the x-ray attenuation of the patient and is therefore a preferred patient size metric.[11] Retrospective analysis of Dw's in 41 patient scans for CACS performed in our hospital showed that these diameters mostly matched with the Dw of the medium and large extension rings.

Acquisition and reconstruction parameters

Scans were performed on a dual source CT (DSCT) system (SOMATOM Force, Siemens Healthineers, Syngo CT VB10). A reference tube voltage of 120 kVp in combination with automated tube current modulation (ATCM) CARE Dose4D was used for both inserts (*Table 1*). The calcium-aware reconstruction technique was assessed by acquiring data with varying tube voltages of 70–150 kVp, in steps of 10 kVp. Additionally, automatic tube voltage selection ("kVon") was set to keep the contrast to noise ratio for calcium constant when selecting the optimal tube voltage for radiation dose optimization. Finally, a scan was performed using a dedicated CACS Tin filtration protocol with an adaptation of the reference tube voltage to Sn100 in combination with ATCM CARE Dose4D (*Table 1*). All scans were repeated five times after manual repositioning (approximately 2 mm translation and 2 degrees rotation) of the phantom to assess positioning influence and interscan variation.

Images were reconstructed with the conventional calcium scoring reconstruction technique (kernel Qr36) and the dedicated calcium-aware reconstruction

11

Table 1 Acquisition and reconstruction parameters

Scanner*	SOMATOM Force	SOMATOM Force-tin filtration
Acquisition mode	Sequential	Sequential
Scan length (mm)	100.5	100.5
Reference tube voltage	120	Sn100
Reference tube current product	80	534
Manual tube voltage settings	70–150	Sn100
CARE kV dose optimization slider**	5 (bone/calcium)	5 (bone/calcium)
Collimation (mm)	32 × 1.2	32 × 1.2
Rotation time (sec)	0.25	0.25
Image reconstruction (FBP)	Qr36 and Sa36	Qr36 and Sa36
Slice thickness (mm)	3.0	3.0
Increment (mm)***	1.5	1.5
FoV (mm)	180	180
Reconstruction matrix	512 × 512	512 × 512

*Siemens Healthineers, Syngo CT VB10
**The dose optimization slider from the default calcium scoring protocol was retained
***Increment of 1.5 mm is the standard for calcium scoring with Siemens equipment

technique (kernel Sa36), both based on FBP. For the latter technique, calcium is identified in preliminary reconstructed images and a lookup table is used to correct the CT numbers of calcium in the finally reconstructed images.[8] The exact working of the algorithm is proprietary information of the vendor. The algorithm is fully integrated within the standard image reconstruction interface and can be activated by selecting the corresponding reconstruction kernel (Sa36). It does not need an additional workstation or increased reconstruction times.

Image and dose analysis

The volumetric CT dose index values ($CTDI_{vol}$) in mGy were noted to assess potential radiation dose reduction. Consistency of CT numbers (mean and standard deviation (SD)) was determined in the central calcium insert (200HA) of the CCI insert. Noise SD was determined within a homogeneous region of the CCI insert. Agatston score, together with different image quality metrics, was computed using an in-house developed Python script (Python version 3.7) for the D100 and CCI insert. Resulting Agatston scores of the Python script were validated against the standard vendor-specific scoring software (Syngo.via, Siemens Healthineers) with the aid of CCI data and proven equal (maximum deviation 0.1%).

This study addresses directly the CT number or CT value in Hounsfield units (HU) of calcifications. CT numbers are related to the linear x-ray attenuation coefficients and depend on the density, the effective atomic number, and x-ray tube voltage.[12] The attenuation coefficient of the phantom base material does not resemble the attenuation coefficient of human soft tissue equally well at all tube voltages. Allmendinger et al previously described a base material-specific correction, necessary for correct Agatston scores at varying tube voltages by adjustment of the standard 130 HU threshold.[8] This correction was applied automatically in our study as well as for all reconstructions.

Image noise was compared with recommended noise targets (in HU) for calcium scoring CT scans defined for different chest sizes (small, medium, large chest width): 20 HU for the small and medium chest width, and 23 HU for the large chest width.[13]

Additionally, an Agatston score was determined in a non-calcium region (55×55 mm), therefore depending purely on noise. This score was called the background Agatston score (BAS). For acquisitions with a non-zero BAS, the Agatston scores of calcifications could be less reliable, as it was uncertain if a calcification was seen at a specific location, or just noise. These scores were noted.

Reference values for both inserts were the Agatston scores acquired with a tube voltage of 120 kVp and reconstructed with the standard technique (Qr36). Each deviation in acquisition or reconstruction was compared against this reference

Statistical analyses

SPSS (version 25, IBM Corp) was used for statistical analysis. Normality of data was tested with the Shapiro-Wilk test. Wilcoxon signed-rank test was performed to evaluate statistically significant difference of the median Agatston scores. Intraclass correlation coefficients (ICC) with a 95% confidence interval (CI) and Bland-Altman plots of the Agatston scores between two different techniques were assessed. A p value of < 0.05 was considered statistically significant. Agatston scores are given as median values of the five measurements.

RESULTS

Radiation dose and noise values

Reference dose levels at 120 kVp for the small, medium, and large phantom size were 1.57, 2.59, and 3.84 mGy respectively. For the scans with automatic tube voltage selection, tube voltage was reduced to 90 kVp for the small and medium phantom size, while 100 kVp was selected for the large phantom. In comparison with the corresponding reference, radiation dose levels decreased by 22%, 15%, and 12% for the small, medium, and large phantom size, respectively.

Within the dedicated Tin CACS protocol, dose values were 55% lower for both small and medium phantom size and 60% for the large phantom size compared with the reference dose levels at 120 kVp.

Median noise values for the 120 kVp and the images obtained with automatic tube voltage selection increased with increasing phantom diameter for both reconstruction techniques (*Figure 1*). The noise level in all three phantom sizes was highest when using Tin filtration. Moreover, the recommended noise target for calcium scoring CT scans was exceeded for some tube voltages in the medium phantom and for all tube voltages in the large phantom size (*Figure 1*). Despite the high number of noise limit exceeding scans, BAS values were zero for most reconstructions. A BAS > 0 was found only for the large phantom in combination with a tube voltage of 70 kVp or Sn100.

CT number constancy

Considering the large calcification with 200 mg HA/cm³ in the CCI insert for all phantom sizes, CT numbers increased with decreasing tube voltage for the standard reconstruction technique, while these numbers remained virtually constant for the calcium-aware reconstruction technique (*Table 2*). Median HU (min HU–max HU) of the reference (120 kVp + Qr36) was 266 HU (265HU—268HU), 257 HU (257HU—258HU), and 247 HU (246HU—248HU) for the small, medium, and large phantom size, respectively. Compared with the reference, the deviation was up to 64% with the standard reconstruction technique and up to 11% with the calcium-aware reconstruction technique when varying the tube voltage (*Table 2*).
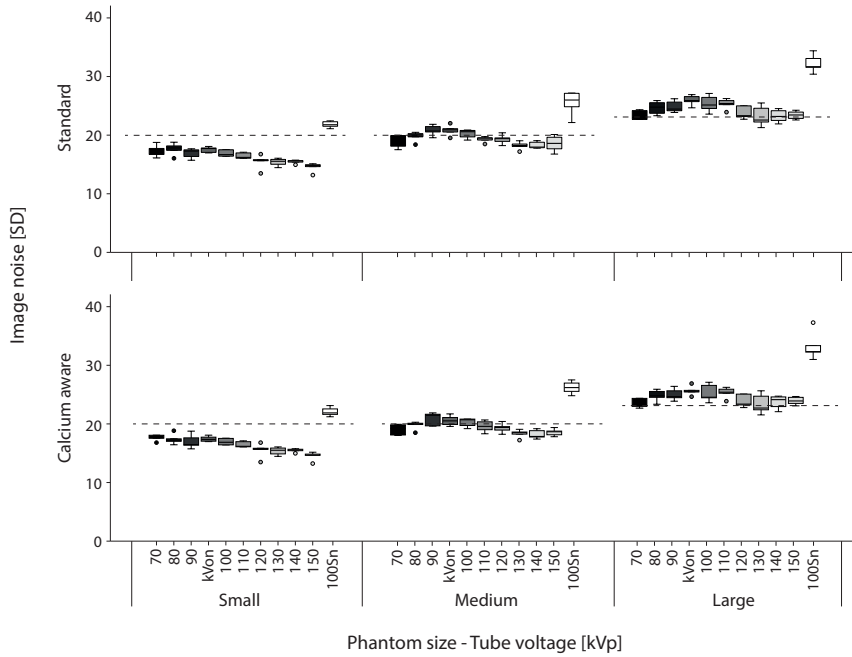
Figure 1 Box-and-whisker plots of the noise measurements of the homogeneous central slice of the CCI insert. Recommended noise targets (in HU) for calcium scoring CT scans defined for different chest sizes were applied to the images as dotted lines: 20 HU for the small and medium chest width, and 23 HU for the large chest width. The automatic tube voltage selection is illustrated by "kVon"

Agatston score

When varying the tube voltage, Agatston scores deviated up to 40% and 8% from the reference for the standard and calcium-aware reconstruction technique, respectively (*Table 3*). The overall spread in median Agatston scores for varying tube voltages decreased for the calcium-aware reconstruction technique for both the CCI and D100 insert (*Figures 2* and *3*). Considering all phantom sizes, the Agatston scores in the CCI insert increased with 14% for the automated tube voltage selection and decreased with 14% within the tin-filtrated scans for the standard reconstruction technique (*Figure 2a*). For the calcium-aware reconstruction technique, Agatston score deviations from the reference were much less: 3.6% at automated tube voltage selection and 2.4% with the tin-filtrated scans (*Figure 2b*). For the D100 insert, we observed similar results; however, the deviations from the reference were larger than in the CCI insert, especially for the varying tube voltage in combination with the standard reconstruction technique (*Figure 3*). Representative images of the D100 insert for the standard reconstruction technique with 120 kVp and

11

Table 2 Deviation of the CT number of calcium at varying tube voltage and phantom size compared with the reference with a tube voltage of 120 kV and the standard reconstruction technique (Qr36)

	Deviation of the CT number of calcium		
	Calcium-aware reconstruction technique		
Tube voltage	Small phantom	Medium phantom	Large phantom
70	-11.0% (-9.9% to-11.1%)	-4.3% (-4.0% to -5.0%)	2.2% (1.8% to 3.5%)
80	-7.0% (-6.7% to -7.6%)	-2.7% (-2.3% to -2.8%)	3.2% (2.6% to 3.8%)
90	-3.6% (-3.4% to -4.6%)	-0.5% (-0.9% to 0.7%)	3.7% (3.6% to 5.4%)
100	-2.4% (-2.0% to -2.9%)	0.6% (0.3% to 1.7%)	4.7% (3.3% to 5.6%)
110	-2.2% (-1.7% to -2.7%)	1.1% (0.8% to 1.3%)	5.2% (4.2% to 5.6%)
120	-1.5% (-0.7% to -2.0%)	2.1% (1.9% to 2.3%)	5.4% (5.1% to 6.0%)
130	-1.4% (-0.6% to -2.3%)	1.9% (0.9% to 2.1%)	5.8% (5.5% to 6.5%)
140	-0.9% (-0.8% to -1.7%)	2.0% (0.8% to 2.9%)	4.7% (4.4% to 6.1%)
150	-1.1% (-0.6% to -1.8%)	1.9% (1.2% to 3.0%)	5.7% (5.2% to 6.5%)
Sn100	-5.3% (-4.3% to -7.0%)	-2.5% (-1.4% to -3.8%)	1.5% (0.2% to 3.7%)

	Deviation of the CT number of calcium		
	Standard reconstruction technique		
Tube voltage	Small phantom	Medium phantom	Large phantom
70	60.5% (59.6% to 61.6%)	61.0% (60.4% to 62.0%)	63.6% (62.0% to 65.6%)
80	40.1% (39.4% to 40.4%)	39.1% (39.0% to 40.1%)	40.8% (40.3% to 41.6%)
90	25.4% (25.3% to 26.3%)	24.4% (24.1% to 25.9%)	25.3% (24.2% to 26.5%)
100	14.9% (14.6% to 15.3%)	13.6% (13.3% to 15.0%)	14.0% (12.5% to 15.9%)
110	6.3% (5.7% to 6.9%)	5.8% (5.6% to 6.2%)	6.4% (5.4% to 6.7%)
120	0.0% (-0.5% to 0.8%)	0.0% (-0.2% to 0.2%)	0.0% (-0.3% to 0.6%)
130	-5.3% (-4.5% to -6.2%)	-5.6% (-5.4% to -6.3%)	-4.6% (-4.0% to -4.7%)
140	-9.4% (-9.3% to 10.1%)	-9.6% (-8.8% to 10.7%)	-9.6% (-8.6% to 10.0%)
150	-13.0% (-12.6% to -13.6%)	-13.2% (-12.4% to -13.8%)	-12.3% (-11.3% to -12.5%)
Sn100	-16.5% (-16.2% to -17.9%)	-13.9% (-13.0% to -14.8%)	-11.1% (-9.0% to -11.8%)

Values given in Median% (Min% to Max%]

<div style="margin-right:4em; float:right;">**11**</div>

the calcium-aware reconstruction technique at reduced tube voltage for all three phantom sizes are shown in *Figure 4*. This figure shows calcifications with an Agatston score of zero for the reference, while the calcium-aware reconstruction technique Agatston scores are non-zero.

There was a very high ICC (0.991) and 95% CI for the automated tube voltage selection with the standard reconstruction technique compared with the reference when considering all calcifications (*Figure 5a*). When considering only the low Agatston scores, both the ICC and 95% CI decreased (*Figure 5b*). There was a very high ICC (0.998) and 95% CI for the automated tube voltage selection and

Table 3 Agatston score deviation at varying tube voltage and phantom size compared with the reference with a tube voltage of 120 kV and the standard reconstruction kernel (Qr36)

| | Agatston score deviation | | |
| | Calcium-aware reconstruction technique | | |
Tube voltage	Small phantom	Medium phantom	Large phantom
70	-7.5% (-1.9% to -10.4%)	-2.5% (2.6% to -7.1%)	1.8% (-3.1% to 5.6%)
80	-7.0% (-3.0% to -11.7%)	-3.8% (1.7% to -8.0%)	-1.6% (5.2% to -7.0%)
90	-5.2% (-1.2% to -8.8%)	-2.0% (2.5% to -7.4%)	-2.2% (4.3% to -6.1%)
100	-3.4% (-1.2% to -7.2%)	-2.8% (1.9% to -6.8%)	-1.5% (4.4% to -5.7%)
110	-5.6% (-0.5% to -8.5%)	-0.8% (3.2% to -6.2%)	2.5% (-3.4% to 6.8%)
120	-1.1% (3.8% to -5.6%)	0.7% (-2.8% to 4.6%)	2.2% (-2.3% to 10.3%)
130	-3.5% (1.1% to -5.7%)	-0.9% (4.9% to -5.5%)	2.4% (-5.2% to 10.7%)
140	-2.6% (2.4% to -5.8%)	-0.1% (-4.1% to 6.8%)	0.6% (-3.5% to 7.9%)
150	-3.1% (3.1% to -5.5%)	0.1% (-2.9% to 6.7%)	1.7% (-3.4% to 10.1%)
Sn100	-4.4% (0.1% to -8.7%)	-2.5% (5.6% to -7.8%)	-0.4% (9.3% to -17.4%)

| | Agatston score deviation | | |
| | Standard reconstruction technique | | |
Tube voltage	Small phantom	Medium phantom	Large phantom
70	39.7% (33.6% to 44.0%)	38.1% (33.0% to 44.3%)	36.7% (31.6% to 43.3%)
80	26.3% (21.9% to 28.8%)	24.6% (20.9% to 28.4%)	24.4% (19.1% to 27.7%)
90	18.1% (14.2% to 22.3%)	16.8% (11.6% to 20.0%)	14.5% (9.9% to 20.4%)
100	13.1% (5.8% to 16.5%)	7.0% (2.7% to 15.8%)	7.7% (1.1% to 13.1%)
110	1.7% (-1.3% to 7.4%)	2.2% (-0.6% to 7.2%)	3.7% (-2.5% to 8.4%)
120	0.0% (-4.4% to 5.3%)	0.0% (-4.6% to 3.9%)	0.0% (-4.7% to 5.6%)
130	-6.2% (-2.5% to -7.5%)	-6.2% (-2.0% to -11.1%)	-7.5% (-0.6% to -11.3%)
140	-10.2% (-5.3% to 12.2%)	-10.5% (-6.6% to -14.4%)	-11.3% (-6.8% to -16.5%)
150	-12.7% (-10.0% to -14.1%)	-11.3% (-8.8% to -16.8%)	-14.5% (-7.2% to -17.1%)
Sn100	-13.7% (-9.7% to -17.2%)	-15.7% (10.5% to -41.0%)	-12.2% (-3.0% to -28.3%)

Values given in Median% [min% to max%]

the calcium-aware reconstruction technique compared with the reference (*Figure 5c*). When considering only the low Agatston scores, both the ICC and 95% CI decreased (*Figure 5d*). However, this decrease was less than observed within the standard reconstruction technique. A Bland-Altman analysis of the data is shown in *Figure 6*. The Bland-Altman plots demonstrate the agreement between the two reconstruction kernels. The negative mean difference within *Figure 6a, b,* and *d* demonstrates that, regardless of reconstruction technique, Agatston scores are higher for automatic tube voltage selection in comparison with 120 kVp. The opposite applies for the calcium-aware reconstruction technique and automatic tube voltage selection (*Figure 6c*).
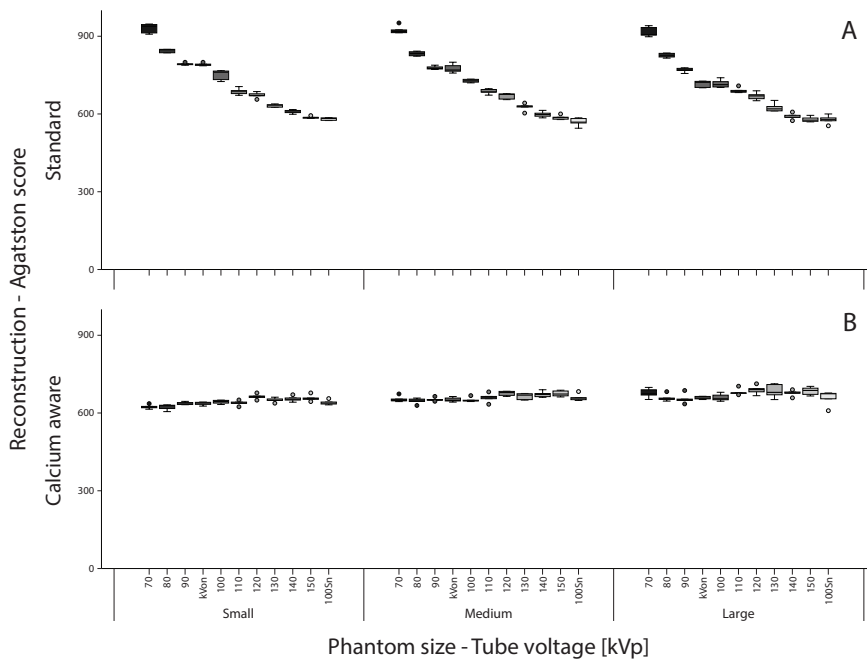
Figure 2 **A** Box-and whisker plots of the Agatston score within the CCI insert with the standard reconstruction technique. **B** Box-and-whisker plots of the Agatston score within the CCI insert with the calcium-aware reconstruction technique. Scores are given per phantom size-tube voltage combination. The automatic tube voltage selection is illustrated by "kVon"

11

For the CCI insert, increasing the phantom diameter from small to large demonstrated no statistically significant decrease ($p = 0.5$) of the median (range) Agatston scores from 671 (656.2—686.5) to 669.9 (651.1—689.4) for the reference (120 kVp and Qr36). A statistically significant increase ($p < 0.05$) in Agatston score from 639 (626.9—642.4) to 657.4 (652–664.5) was observed for the calcium-aware technique with automated tube voltage selection. For the D100 insert and increasing phantom size from small to large, there was a statistically significant decrease ($p < 0.05$) in Agatston score from 29.3 (26.5—31.2) to 25.6 (21.4—27.8) for the reference ($p < 0.05$), and a statistically significant decrease ($p < 0.05$) in Agatston score from 49.0 (47.4—56.7) to 37.1 (34.0—39.4) for the calcium-aware reconstruction technique with automatic tube voltage selection.
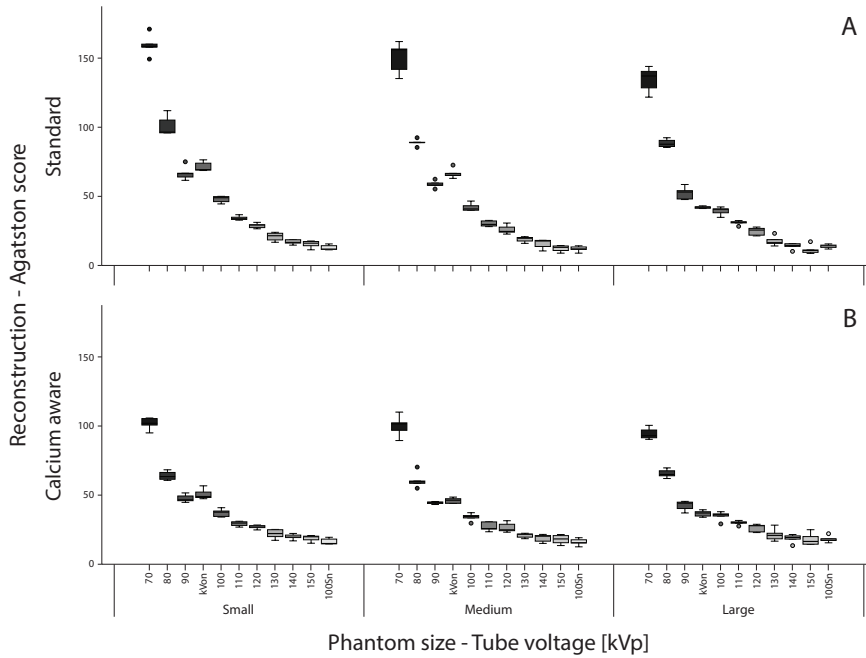
Figure 3 **A** Box-and-whisker plots of the Agatston score within the D100 insert with the standard reconstruction technique. **B** Box-and whisker-plots of the Agatston score within the D100 insert with the calcium-aware reconstruction technique. Scores are given per phantom size-tube voltage combination. The automatic tube voltage selection is illustrated by "kVon"
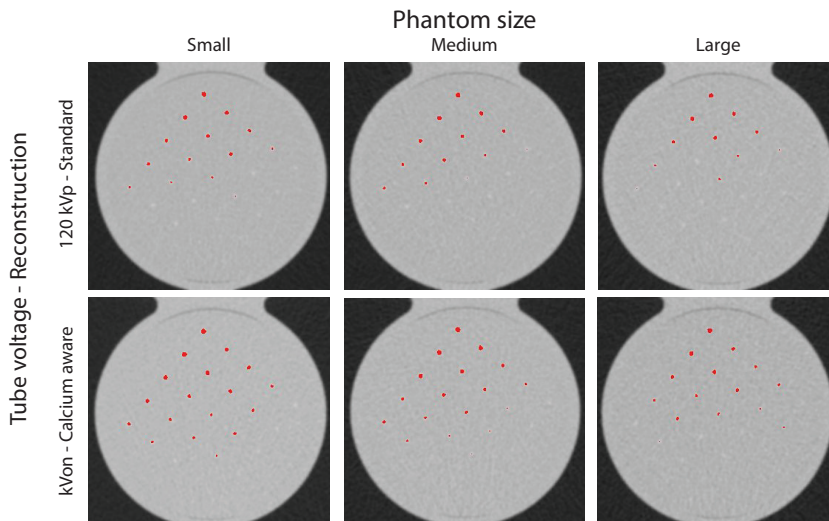
Figure 4 Visualization of calcifications in the D100 insert with all voxels with a CT number above the threshold colored red. From left to right, the phantom size increases. The upper row images were reconstructed with the standard reconstruction technique with a tube voltage of 120 kVp. Lower row images were reconstructed with the calcium-aware reconstruction technique and automated tube voltage selection (90 kVp for the small and medium size phantom and 100 kVp for the large size phantom)
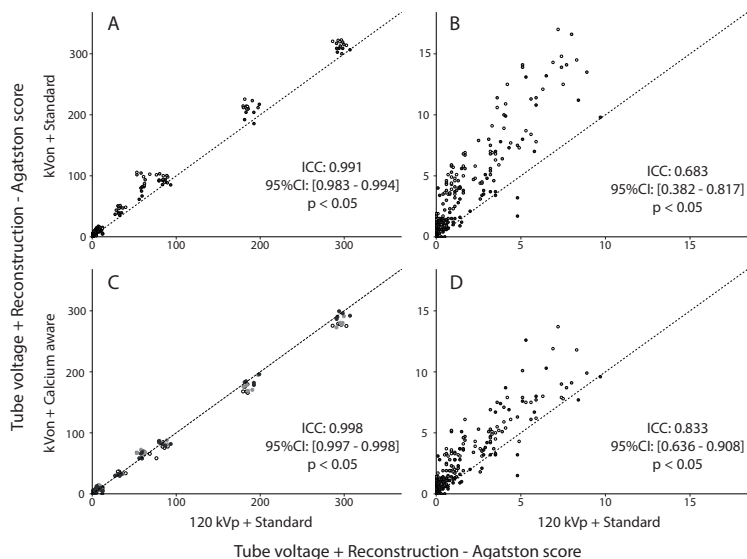
Figure 5 The ICC of the Agatston score for the small, medium, and large phantom for. **A** The standard reconstruction technique with automatic tube voltage selection compared with the standard reconstruction with 120 kVp. **B** Detail of the graph in a representing the low density and small calcifications. **C** The calcium-aware reconstruction technique with automatic tube voltage selection and the standard reconstruction with 120 kVp. **D** Detail of the graph in **c** representing the low density and small calcifications
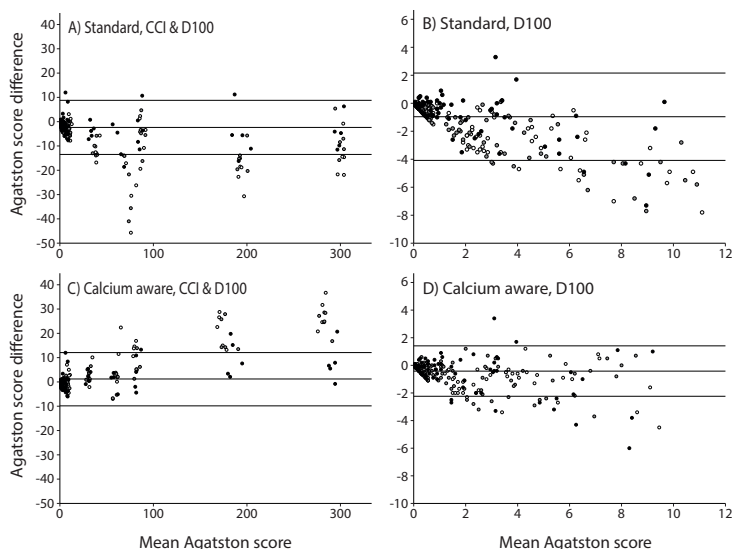
Figure 6 Bland-Altman Plots with mean difference and 95% limits of agreement for the small, medium, and large phantom with the CCI and/or D100 insert. All plots show an Agatston score comparison between the reference at 120 kVp (with standard reconstruction technique) and scans with automatic tube voltage selection (with standard reconstruction technique (**A, B**) and with calcium-aware reconstruction technique (**C, D**))

DISCUSSION

Our results demonstrate that CACS with a calcium-aware image reconstruction technique allows for consistent CT numbers when varying the tube voltage and allows for reduced radiation exposure with automatic reduction of tube voltage.

The Agatston scores with the calcium-aware reconstruction technique deviated up to 8% for the calcifications of the CCI insert across 70 to 150 kVp and Sn100, whereas the Agatston score with the standard reconstruction deviated much more with up to 40%. The latter might be explained by the increase of the photo-electric effect for calcium when scanning with low tube voltage settings. In contrast to the CCI insert, Agatston scores were not stable for the calcifications of the D100 insert when varying the tube voltage. For the calcium-aware reconstruction technique, this might be explained by a sub-optimal identification of the voxels containing calcifications of small diameter and low density and subsequently a sub-optimal correction of the CT numbers.

As seen in *Figure 4*, there were additional calcifications detected when lowering tube voltage. Thus, it might be possible that a patient with a zero Agatston score at 120 kVp might have a non-zero Agatston score at a lower tube voltage, despite the application of the calcium-aware reconstruction technique. This might influence the work-up of patients suspected for coronary artery disease. However, the increase of Agatston score in the D100 insert, as demonstrated in *Figure 4*, is due to true calcified lesions. Instead of improving the calcium-aware reconstruction technique presented in this study to better resemble the Agatston scores at 120 kV, we prefer to reinvent calcium imaging and think it is time to let go the conventional scoring method.[14,15] For example, Groen et al described a correction applied to the 130 HU calcium scoring threshold for the increased CT numbers of calcium when varying tube voltage and applying the standard reconstruction technique.[16]

Our study demonstrated a decrease in Agatston score with increasing phantom size, as previously described for the standard reconstruction technique and the D100 insert.[17] However, our study used both the CCI and the D100 insert and in addition the calcium-aware reconstruction technique. We observed an increase of the Agatston score for the CCI insert when using the calcium-aware reconstruction technique. The increase in Agatston score might be explained by the sub-optimal identification of the voxels containing small and low-density calcifications, while noise increased.

Calcium CT numbers were constant for the calcium-aware reconstruction technique with automated tube voltage selection, irrespective of phantom size. However, Agatston scores varied more than the reference for different patient sizes. The reason for this is twofold. First, the constancy of CT numbers is calculated as the mean of a large ROI enclosing the calibration rod of the CCI phantom, while Agatston scores are calculated for the smaller nine calcifications. Second, despite the use of clinical scan protocols, higher noise levels were shown especially for the lower tube voltages and the automated tube voltage selection (*Figure 1*). Our computation of the Agatston score was validated to the standard vendor-specific software, calculating every single voxel above a threshold of 130 HU for CACS. With higher noise levels, Agatston scores also increase.

Technological developments like tin filtration and automated tube voltage selection allow for a substantial dose reduction. For example, a 100 kVp with tin filtration CACS protocol demonstrated similar Agatston scores as the reference protocol with 120 kVp despite using the standard reconstruction technique.[18] Larger deviations are expected for tube voltages like 70 and 80 kVp (*Table 2*). A great advantage of the currently considered calcium-aware reconstruction technique is that CACS can be obtained more accurately from any acquisition, regardless of applied tube voltage and filtration. This allows for CACS to be considered within cancer screening protocols. The use of a CACS with the aid of tin filtration combined with an early prototype of a calcium-aware reconstruction technique was described in a patient study and considered to be potentially feasible for calcium scoring.[19] However, in this study and our study, an increased image noise for the tin-filtrated scans was observed. The noise levels were above the recommended noise levels by the SCCT in all three phantom sizes, especially for the large phantom size. Possible solutions for sub-optimal identification of calcification when applying tin filtration with increased noise levels are proposed, e.g., a HU threshold correction for CACS or investigation to apply iterative reconstructions.[20] Within our study, we observed BAS of >0 for the tin-filtrated vendor-recommended scans in the large phantom size. Therefore, caution must be taken when applying the tin-filtrated scans in clinical routine, especially when CACS is obtained for calcification of small diameter and low density, as the calcium-aware reconstruction technique is also not able to correct these.

The recommended noise levels were not only exceeded for the tin-filtrated scanning protocols, but also for all tube voltage settings within the large phantom diameter, despite the use of the vendor-recommended scanning protocols. This

warrants further investigation for adjusting the reference tube current value or the adaptation strength of the CARE Dose4D dose curve to achieve the recommended noise target level.[13] However, it seems that the recommended noise target limit comes with a very safe margin. After all, the BAS was zero for all reconstructions in the small- and medium-sized phantoms and for the calcium-aware reconstruction technique with automated tube voltage selection in all phantoms.

There are limitations in this study that need to be considered. This study was phantom-based and despite the effort to represent clinical routine, patient studies are necessary to validate our findings. CTDIvol is an indicator of the CT scanner radiation output. The dose received by a patient depends on this CTDIvol and the individual patient size. It is recommended to use the size-specific dose estimates (SSDE) to reflect estimated doses for the individual patient.[21] Furthermore, it might be of interest to use a non-stationary phantom model instead of a stationary one. This makes it feasible to assess whether or not heart rate variability will influence Agatston scores when using the calcium-aware reconstruction technique.

CONCLUSION

In general, CT numbers remained consistent with comparable calcium scores when the calcium-aware image reconstruction technique was applied with varying tube voltage. Less consistency was observed in small calcifications with low density. Automatic reduction of tube voltage resulted in a dose reduction of up to 22%.
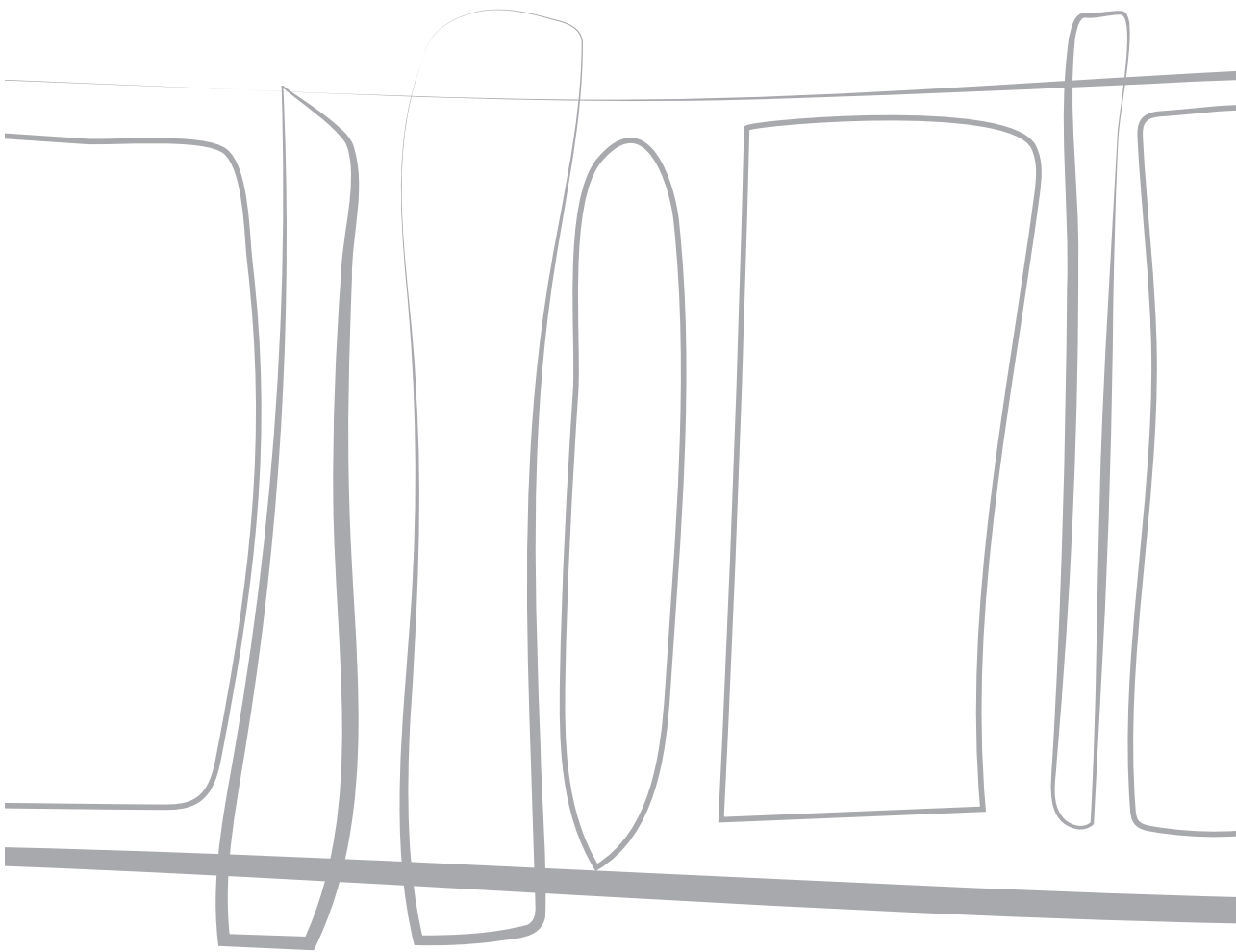
11

REFERENCES

1. Greenland P, Alpert JS, Beller GA, et al. 2010 ACCF/AHA guideline for assessment of cardiovascular risk in asymptomatic adults: Executive summary: A report of the American College of cardiology foundation/American Heart association task force on practice guidelines. Circulation. 2010;122(25):2748-2764. doi:10.1161/CIR.0b013e3182051bab

2. Eurostat (2018) Statistics causes of death in the European Union (EU) 2005 - 2015. https://ec.europa.eu/eurostat/statistics-explained/index.php/Causes_of_death_statistics

3. Elias-Smale SE, Proença RV, Koller MT, et al. Coronary calcium score improves classification of coronary heart disease risk in the elderly: The Rotterdam study. J Am Coll Cardiol. 2010;56(17):1407-1414. doi:10.1016/j.jacc.2010.06.029

4. Willemink MJ, van der Werf NR, Nieman K, Greuter MJW, Koweek LM, Fleischmann D. Coronary artery calcium: A technical argument for a new scoring method. J Cardiovasc Comput Tomogr. 2019;13(6):347-352. doi:10.1016/j.jcct.2018.10.014

5. Hecht HS, Cronin P, Blaha MJ, et al. 2016 SCCT/STR guidelines for coronary artery calcium scoring of noncontrast noncardiac chest CT scans: A report of the Society of Cardiovascular Computed Tomography and Society of Thoracic Radiology. J Cardiovasc Comput Tomogr. 2017;11(1):74-84. doi:10.1016/j.jcct.2016.11.003

6. Hecht H, Blaha MJ, Berman DS, et al. Clinical indications for coronary artery calcium scoring in asymptomatic patients: Expert consensus statement from the Society of Cardiovascular Computed Tomography. J Cardiovasc Comput Tomogr. 2017;11(2):157-168. doi:10.1016/j.jcct.2017.02.010

7. Bijwaard H, Pruppers M, de Waard-Schalkx I. The Influence of Population Aging and Size on the Number of CT Examinations in The Netherlands. Health Phys. 2014;107(1). https://journals.lww.com/health-physics/Fulltext/2014/07000/The_Influence_of_Population_Aging_and_Size_on_the.8.aspx

8. Allmendinger T, Hamann A. Agatston score calcium quantification with arbitrary tube voltage [white paper]. Siemens Healthineers. Published 2019. https://www.siemens-healthineers.com/computed-tomography/clinical-imaging-solutions/cardiovascular-imaging#CLINICAL_USE

9. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. J Am Coll Cardiol. 1990;15(4):827-832. doi:10.1016/0735-1097(90)90282-T

10. Groen JM, Kofoed KF, Zacho M, Vliegenthart R, Willems TP, Greuter MJ. Calcium score of small coronary calcifications on multidetector computed tomography: Results from a static phantom study. Eur J Radiol. 2013;82(2):e58-e63. doi:10.1016/j.ejrad.2012.09.018

11. McCollough CH, Ulzheimer S, Halliburton SS, Shanneik K, White RD, Kalender WA. Coronary Artery Calcium: A Multi-institutional, Multimanufacturer International Standard for Quantification at Cardiac CT. Radiology. 2007;243(2):527-538. doi:10.1148/radiol.2432050808

12. Lamba R, McGahan JP, Corwin MT, et al. CT Hounsfield Numbers of Soft Tissues on Unenhanced Abdominal CT Scans: Variability Between Two Different Manufacturers' MDCT Scanners. Am J Roentgenol. 2014;203(5):1013-1020. doi:10.2214/AJR.12.10037

11

13. Voros S, Rivera JJ, Berman DS, et al. Guideline for minimizing radiation exposure during acquisition of coronary artery calcium scans with the use of multidetector computed tomography: A report by the Society for Atherosclerosis Imaging and Prevention Tomographic Imaging and Prevention Counci. J Cardiovasc Comput Tomogr. 2011;5(2):75-83. doi:10.1016/j.jcct.2011.01.003

14. Blaha MJ, Mortensen MB, Kianoush S, Tota-Maharaj R, Cainzos-Achirica M. Coronary Artery Calcium Scoring: Is It Time for a Change in Methodology? JACC Cardiovasc Imaging. Published online 2017. doi:10.1016/j.jcmg.2017.05.007

15. Nieman K. Evolve or perish for coronary calcium imaging. Eur Heart J Cardiovasc Imaging. 2015;16(4):354-355. doi:10.1093/ehjci/jeu220

16. Groen JM, Dijkstra H, Greuter MJ, Oudkerk M. Threshold adjusted calcium scoring using CT is less susceptible to cardiac motion and more accurate. Med Phys. 2009;36(2):438-446.

17. Willemink MJ, Abramiuc B, den Harder AM, et al. Coronary calcium scores are systematically underestimated at a large chest size: A multivendor phantom study. J Cardiovasc Comput Tomogr. 2015;9(5):415-421. doi:10.1016/j.jcct.2015.03.010

18. Apfaltrer G, Albrecht MH, Schoepf UJ, et al. High-pitch low-voltage CT coronary artery calcium scoring with tin filtration: accuracy and radiation dose reduction. Eur Radiol. Published online February 2018. doi:10.1007/s00330-017-5249-2

19. Tesche C, De Cecco CN, Schoepf UJ, et al. CT coronary calcium scoring with tin filtration using iterative beam-hardening calcium correction reconstruction. Eur J Radiol. 2017;91(December 2016):29-34. doi:10.1016/j.ejrad.2017.03.011

20. Vonder M, Pelgrim GJ, Huijsse SEM, et al. Feasibility of spectral shaping for detection and quantification of coronary calcifications in ultra-low dose CT. Eur Radiol. 2017;27(5):2047-2054. doi:10.1007/s00330-016-4507-z

21. Brink JA, Morin RL. Size-specific Dose Estimation for CT: How Should It Be Used and What Does It Mean? Radiology. 2012;265(3):666-668. doi:10.1148/radiol.12121919

11
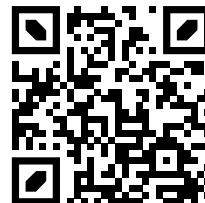
Online material is available via:

CHAPTER 12

Evaluating a calcium-aware kernel for CT CAC-scoring with varying surrounding materials and heart rates: a dynamic phantom study

Niels R. van der Werf, MSc

Ronald Booij, PhD

Bernhard Schmidt, PhD

Thomas G. Flohr, PhD

Tim Leiner, MD PhD

Joël J. de Groen, BSc

Daniel Bos, MD PhD

Ricardo P.J. Budde, MD PhD

Martin J. Willemink, MD PhD

Marcel J.W. Greuter, PhD

ABSTRACT

Objective

The purpose of this study was twofold. First, the influence of a novel calcium-aware (Ca-aware) computed tomography (CT) reconstruction technique on coronary artery calcium (CAC) scores surrounded by a variety of tissues was assessed. Second, the performance of the Ca-aware reconstruction technique on moving CAC was evaluated with a dynamic phantom.

Methods

An artificial coronary artery, containing two CAC of equal size and different densities (196±3, 380±2 mg hydroxyapatite cm^{-3}), was moved in the center compartment of an anthropomorphic thorax phantom at different heart rates. The center compartment was filled with mixtures, which resembled fat, water and soft tissue equivalent CT numbers. Raw data was acquired with a routine clinical CAC protocol, at 120 peak kilovolt (kVp). Subsequently, reduced tube voltage (100 kVp) and tin-filtration (150Sn kVp) acquisitions were performed. Raw data was reconstructed with a standard and a novel Ca-aware reconstruction technique. Agatston scores of all reconstructions were compared with the reference (120 kVp) and standard reconstruction technique, with relevant deviations defined as >10%.

Results

For all heart rates, Agatston scores for CAC submerged in fat were comparable to the reference, for the reduced-kVp acquisition with Ca-aware reconstruction kernel. For water and soft tissue, medium density Agatston scores were again comparable to the reference for all heart rates. Low density Agatston scores showed relevant deviations, up to 15% and 23% for water and soft tissue, respectively.

Conclusion

CT CAC-scoring with varying surrounding materials and heart rates is feasible at patient-specific tube voltages with the novel Ca-aware reconstruction technique.

12

INTRODUCTION

Coronary artery calcifications (CAC), as detected by non-contrast cardiac computed tomography (CT), are a strong predictor for future adverse cardiovascular events.[1–3] With CT, CAC is traditionally quantified according to the Agatston scoring standard.[4] Other CAC scores, such as the mass score, were developed to decrease shortcomings of the Agatston methodology.[5–10] The number of CT CAC assessments in clinical practice has increased substantially, which in turn increased the cumulative radiation exposure to patients undergoing these exams.[11] In order to standardize CT CAC scoring across different types of CT equipment, a setting of 120 kVp in combination with 3 mm slice thickness is recommended.[12] On the other hand, the most efficient way to reduce radiation dose for CT CAC imaging is to decrease the peak tube voltage (kVp) to values below 120 kVp. However, adjusting peak tube voltage will change Agatston scores and has therefore shown to be difficult to implement.[4,10,13]

To address this tradeoff, a novel calcium aware (Ca-aware) reconstruction technique was recently introduced by one of the main CT manufacturers.[14] The aim of this Ca-aware reconstruction kernel is to minimize the previously shown kVp-induced variability of Agatston scores in order to allow for acquisitions at patient-specific lower tube voltages without affecting CAC scores.[15,16] This novel reconstruction kernel is optimized for CAC surrounded by fat, as found in-vivo where the arteries are embedded in the epicardial fat. In the reconstruction process, bone and calcium are identified after which a voltage-dependent lookup table is used to convert CT numbers to values which correspond to a tube potential of 120 kVp. Ideally, this will yield CAC scores obtained at reduced tube voltages that are equivalent to traditional scores that would have been obtained with 120 kVp acquisitions. Changing the resulting tube voltage-dependent Hounsfield unit (HU) values of CAC to their 120 kVp values may enable the use of Agatston calcium scoring methodology independent of tube voltage. For acquisitions with tube voltages below 120 kVp, or with added filtration, this may enable decreased radiation dose while maintaining unchanged CAC scores.

An important confounder of CAC scores is residual motion of coronary arteries during image acquisition.[17–20] Motion artifacts can increase or decrease CAC scores, depending on the density of the calcification.[20] While recent studies have assessed the influence of the novel Ca-aware reconstruction technique on both stationary calcifications and in patients, the effect of this reconstruction technique on CAC scores of moving calcifications of different densities at varying heart rates remains unknown.[21–23]

12

Against this background, we formulated the following two aims. Our first aim was to assess the influence of the Ca-aware reconstruction technique on CAC scores of calcifications with different densities surrounded by a variety of patient equivalent tissues. Second, the influence of the Ca-aware reconstruction technique on moving calcifications was assessed with a dynamic anthropomorphic phantom.

METHODS

An anthropomorphic thorax phantom (QRM-thorax, QRM) containing artificial lungs, a spine and a shell of soft tissue-equivalent material was used (*Figure 1*). An extension ring of fat-equivalent material (QRM-Extension Ring, QRM) was used to increase the phantom dimensions to 400 x 300 mm, similar to the dimensions of an averaged-sized patient.[15]

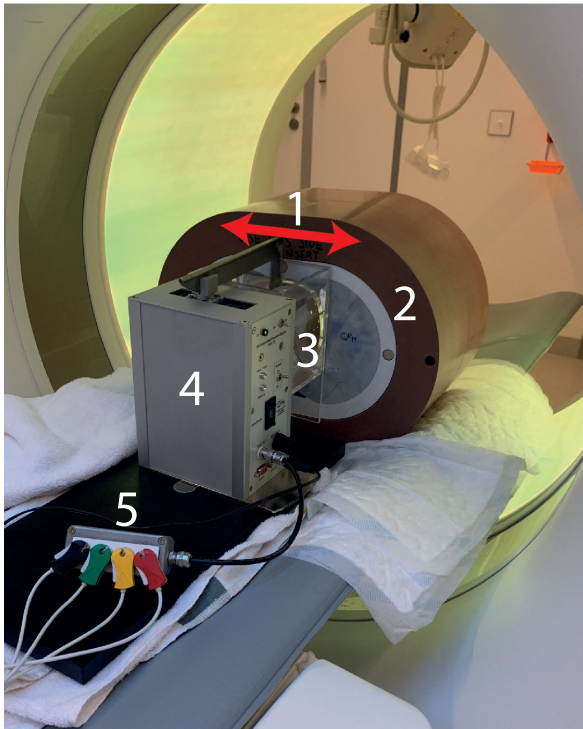*Figure 1* Overview of phantom set-up, with the robotic arm moving a coronary artery along the direction indicated by the red arrow (1), in the anthropomorphic thorax phantom (2), within the fillable compartment (3). Movement Is generated with the Sim2D robot (4), which also provides an electrocardiogram output (5) to ensure data acquisition during linear motion of the artificial coronary artery

12

A fillable compartment was placed within the cylindrical hole in the center of the thorax phantom, in which an artificial coronary artery was linearly translated at constant velocities of 0, 10, 20, and 30 mm/s with the use of a robotic arm (Sim2D, QRM). The artificial coronary artery contained both a low and medium density calcification of 196±3 and 380±2 mg hydroxyapatite (HA) cm^{-3}, respectively. Both calcifications were equal in size: 5.0±0.1 mm in diameter, with a length of 10.0±0.1 mm. The movement was in a horizontal plane, perpendicular to the scan direction. The velocities of the artificial coronary artery corresponded to the average movement of in-vivo coronary arteries during the scan phase at 0, <60, 60-75 and >75 beats per minute (bpm).[18] To ensure that only constant velocities were present during the scan phase, the robotic arm was synchronized with the CT system during acquisition with the use of the electrocardiography trigger output.

Data acquisition and reconstruction

Raw data was acquired with a vendor recommended protocol for CT CAC scoring at 120 kVp on a state-of-the-art CT system (SOMATOM Force, Siemens Healthineers) (*Table 1*). Images were reconstructed with filtered back projection (FBP), using the standard CAC scoring technique (kernel Qr36f), and the Ca-aware reconstruction technique (kernel Sa36). Furthermore, besides the standard 120 kVp acquisition, two other acquisitions were performed. First, data was acquired based on automatic tube voltage selection (CARE kV, Siemens Healthineers) for the water equivalent thickness of the phantom. Second, a dedicated CAC tin-filtration protocol was used. For all protocols, tube current was adjusted according to automatic tube current modulation (CARE Dose4D, Siemens Healthineers) (*Table 1*). The quality reference was set at 80 mAs/rotation, with the dose optimization slider on position 5 (calcium / bone). Due to a limitation in tube current with automatic tube voltage selection for the used phantom size and 100SnkVp, a tube potential of 150Sn kVp was manually selected. To increase sample size and precision, each acquisition was repeated five times for each heart rate. Between each scan, the phantom was manually translated and rotated by approximately 2 mm and 2 degrees, respectively.

In addition to Agatston scores we obtained mass scores by acquiring additional images according to each of the three above-mentioned protocols with a static cardiac calcification insert that included calcium calibration rods (CCI, QRM). These reconstructions were used to calculate the mass calibration factor for each of our protocols, according to the methodology described by McCollough et al.[13]

Table 1 Acquisition and reconstruction parameters for the reference, reduced-kVp and tin-filtration scans

Parameter	Reference	Reduced-kVp	Tin-filtration
Acquisition mode	Sequential	Sequential	Sequential
Ref. tube voltage [kVp]	120	120	100Sn
Ref. tube current product[1] [mAs / rot]	80	80	534
Tube voltage [kVp]	120	100[2]	150Sn[2]
Tube current [mAs / rotation]	100[3]	148[3]	136[3]
Collimation [mm]	2 x 96 x 0.6	2 x 96 x 0.6	2 x 96 x 0.6
Rotation time [s]	0.25	0.25	0.25
Temporal resolution [ms]	66	66	66
Slice thickness [mm]	3	3	3
Slice increment [mm]	1.5[4]	1.5[4]	1.5[4]
Kernel	Qr36f / Sa36f[5]	Qr36f / Sa36f[5]	Qr36f / Sa36f[5]
Reconstruction	FBP	FBP	FBP
Matrix	512x512	512x512	512x512
Field of view [mm]	220	220	220
CTDIvol [mGy]	3.93	3.45	2.47

1 Default quality reference tube current, with dose optimization slider at position 5 (calcium / bone); 2 Automatically selected based on phantom size; 3 Based on water equivalent thickness of used; phantom setup; 4 Standard for calcium scoring with Siemens Healthineers equipment; 5 Ca-aware reconstruction kernel (Sa36f) used to compare results with reference reconstruction kernel (Qr36f) Ref = Reference; rot = rotation; CTDIvol = CT Dose Index Volume; FBP = Filtered Back Projection

12

The fillable compartment, placed in the anthropomorphic phantom, was used to subsequently acquire data with three materials, adjacent to the artificial coronary artery. These materials resembled fat (-100 HU, water-ethanol mixture), water (0 HU) and soft tissue (50 HU, water-iodine contrast agent (Iodixanol) mixture). To account for the tube voltage dependency of our mixtures, for each acquisition (120, 100, 150Sn kVp) the compartment was filled for that specific tube voltage, as indicated by the flowchart in *Figure 2* for the soft tissue acquisitions. Prior to every acquisition, each mixture was manually stirred to prevent curdling of the liquids. In addition, due to the usage of the robotic arm of the dynamic phantom, the mixture was stirred continuously, except during the 0 mm/s acquisitions.
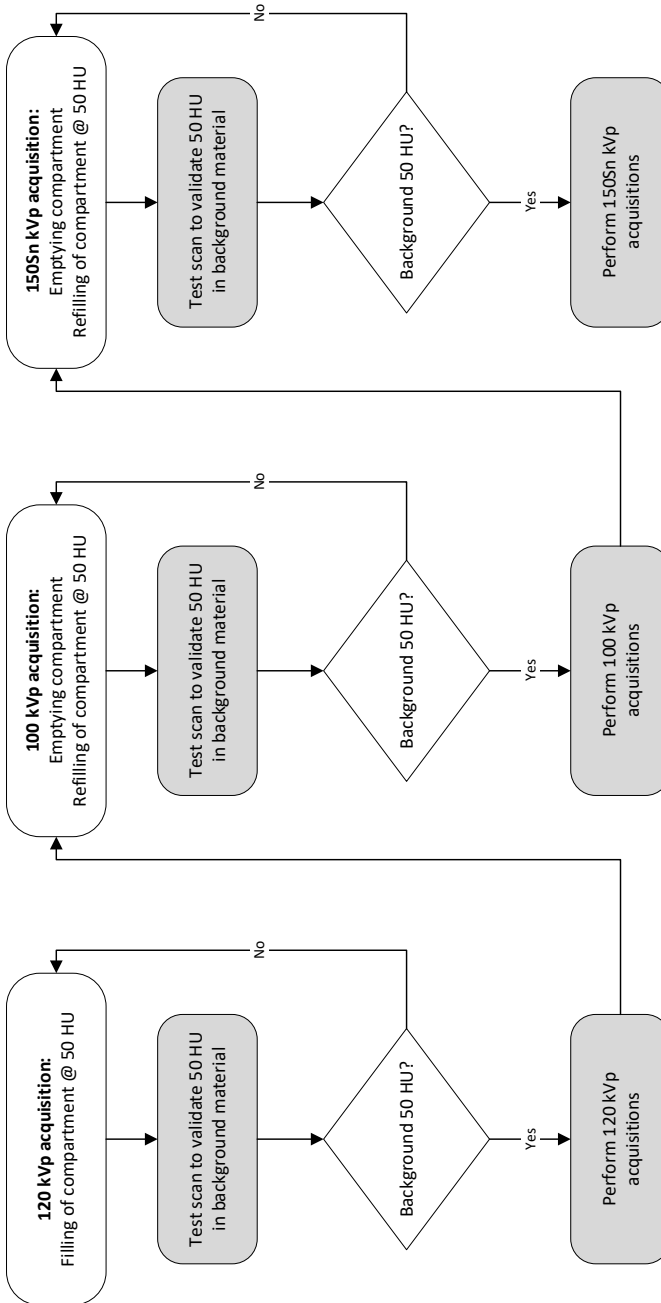
Figure 2 Flowchart which describes the methodology to account for the tube voltage dependency of the used iodine-water mixture to resemble soft tissue

12

Data analysis

CAC scores were determined from the resulting reconstructed images using a previously validated, in-house developed Python script (Python version 3.7).[21] A calcium scoring threshold of 130 HU was used for all reference and Ca-aware reconstruction kernel data. For the reduced-kVp acquisition in combination with the standard kernel, a threshold of 147 HU was used, as described previously.[16] As such an adapted threshold was not available for 150Sn kVp acquisitions, the regular 130 HU threshold was used. For all acquisitions, mean Agatston and mass scores and standard deviation (SD) were calculated from the five repeated measurements for each combination of heart rate, background material, and acquisition protocol. In addition, an Agatston score was calculated for each reconstruction in a uniform background region-of-interest, without any calcium content. This resulted in a background Agatston score (BAS), which is only larger than zero for high image noise levels, as previously described by Booij et al.[21]

For each heart rate, Agatston scores of the reduced-kVp and tin-filtration acquisition were compared to the 120 kVp reference. Differences in Agatston score ≥ 10% were deemed to be relevant. Resulting mass scores were compared to the physical mass of the calcifications. Again, relevant differences were set at ≥ 10%.

RESULTS

Background material and radiation dose

Automatic tube voltage selection, based on the water equivalent thickness of the phantom, resulted in a tube voltage of 100kVp, while 150Sn kVp was manually selected. For these protocols, this resulted in a radiation dose of 3.45 and 2.83 milliGray (mGy), respectively. In comparison with the radiation dose of 3.93 mGy for the reference acquisition, this was a reduction of 12% and 28% for the reduced-kVp and tin-filtration acquisitions, respectively.

On average over all velocities and repetitions, background material mean HU (± SD) and image noise values for fat, water and soft tissue equivalent material, are shown in *Table 2*. Noise levels from the tin-filtration protocol resulted in BAS > 0 for all acquisitions.

Table 2 Background material mean (mean ± SD) and noise (mean ± SD) for all combinations of tube potential, reconstruction kernel and background material, on average for all used velocities and repetitions

Tube potential [kVp]	Kernel	Fat		Water		Soft-tissue	
		Mean	**Noise**	**Mean**	**Noise**	**Mean**	**Noise**
120	Standard	-94.4 ± 0.2	19.9 ± 0.7	0.0 ± 0.4	23.8 ± 0.3	51.3 ± 0.4	25.0 ± 0.4
	Ca-aware	-93.8 ± 0.3	19.9 ± 0.7	0.8 ± 0.4	23.8 ± 0.4	52.3 ± 0.4	25.0 ± 0.4
100	Standard	-99.0 ± 0.3	21.0 ± 1.1	-1.7 ± 0.3	25.3 ± 0.4	51.8 ± 0.3	26.3 ± 0.6
	Ca-aware	-98.5 ± 0.3	21.2 ± 1.0	-1.0 ± 0.3	25.4 ± 0.4	52.3 ± 0.3	26.2 ± 0.5
150Sn	Standard	-84.5 ± 0.4	23.9 ± 0.6	2.4 ± 0.4	27.6 ± 0.6	49.7 ± 0.5	28.5 ± 0.5
	Ca-aware	-84.2 ± 0.3	23.7 ± 0.3	2.9 ± 0.4	27.8 ± 0.6	54.6 ± 0.6	31.4 ± 0.7

Influence of background material on Agatston scores

Representative images for the highest heart rate are shown in *Figure 3* and *Figure 4*, for the low and medium density calcification, respectively. These figures show a reduced detectability for decreased CAC density and increased surrounding material HU. Agatston scores resulting from the different acquisition and reconstruction protocols are shown in *Figure 5* and *Figure 6*, again for the low and medium density calcification, respectively. From these data, it is clear that the standard reconstruction technique leads to clinically relevant differences for all background materials and heart rates for the low-density CAC. With the Ca-aware reconstruction technique, Agatston scores are comparable with the reference. For each velocity, deviations from the reference (FBP with standard kernel) were smaller for the medium density calcification compared to the low-density calcification. Further, almost all tin-filtration acquisitions led to >10% differences in comparison with the reference acquisition for all three background materials.

For all heart rates, Agatston scores for the calcifications submerged in fat-like background material were comparable to the reference, for the reduced-kVp acquisition with Ca-aware reconstruction kernel. For water and soft tissue, medium density Agatston scores were again comparable to the reference for all heart rates. However, low density Agatston scores showed clinically relevant deviations for all heart rates, when the calcification was submerged in water or soft-tissue like material.

12
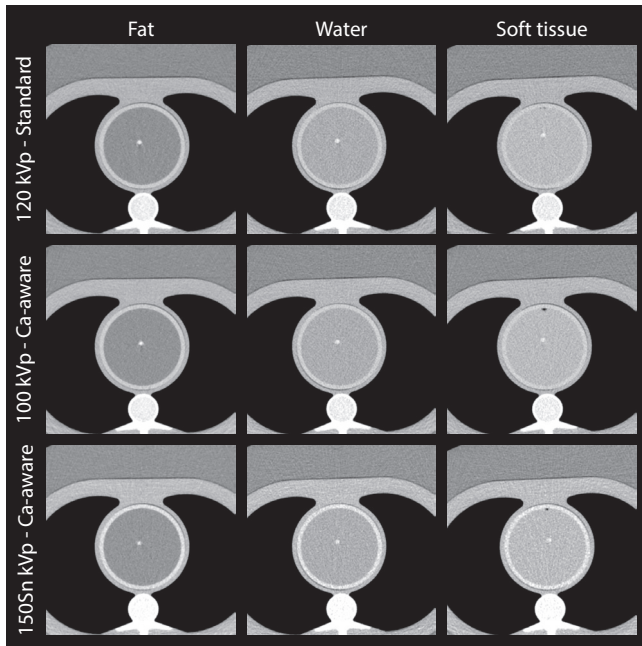
Figure 3 Overview of representative images of the low density calcifications for different combinations of acquisition and reconstruction parameters, and background material, for >75 bpm
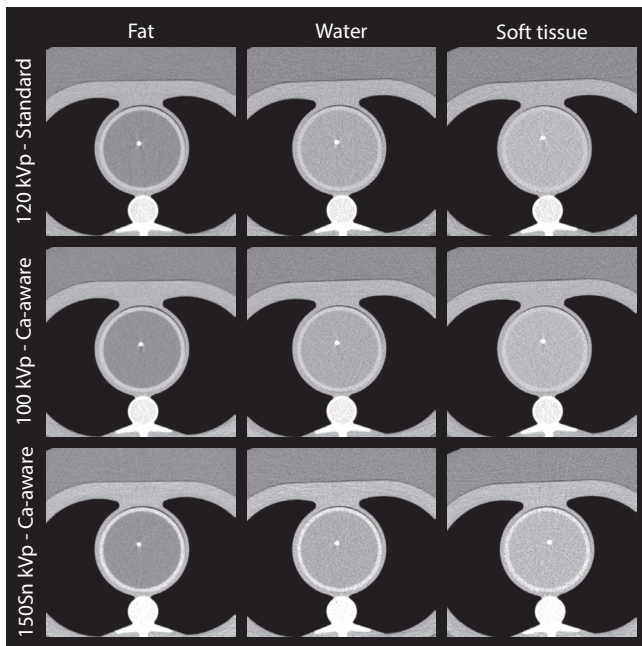
Figure 4 Overview of representative images of the medium density calcifications for different combinations of acquisition and reconstruction parameters, and background material, for >75 bpm
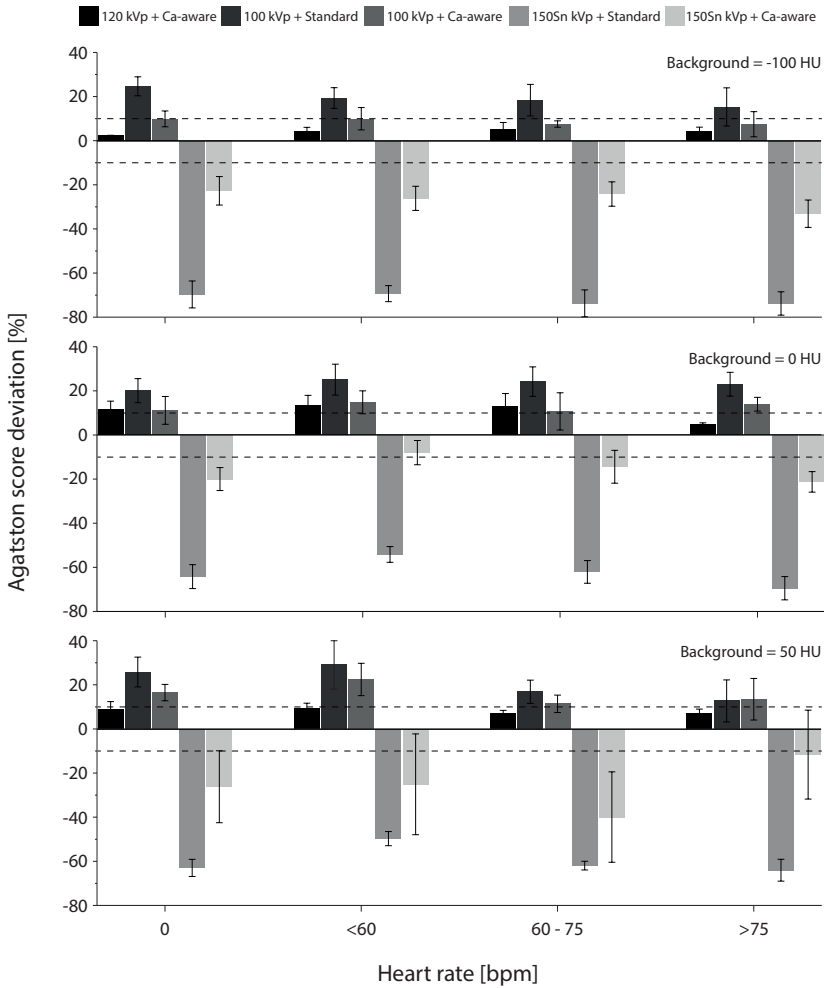
Figure 5 Deviations in low density CAC Agatston score from the reference (for each heart rate: 120 kVp + standard reconstruction kernel) for different heart rates and combinations of tube voltage [kVp] and reconstruction. Results are shown for three background materials: fat (-100 HU, top), water (0 HU, middle) and soft tissue (50 HU, bottom). Clinically relevant differences, at ≥ ±10%, are indicated with dashed lines
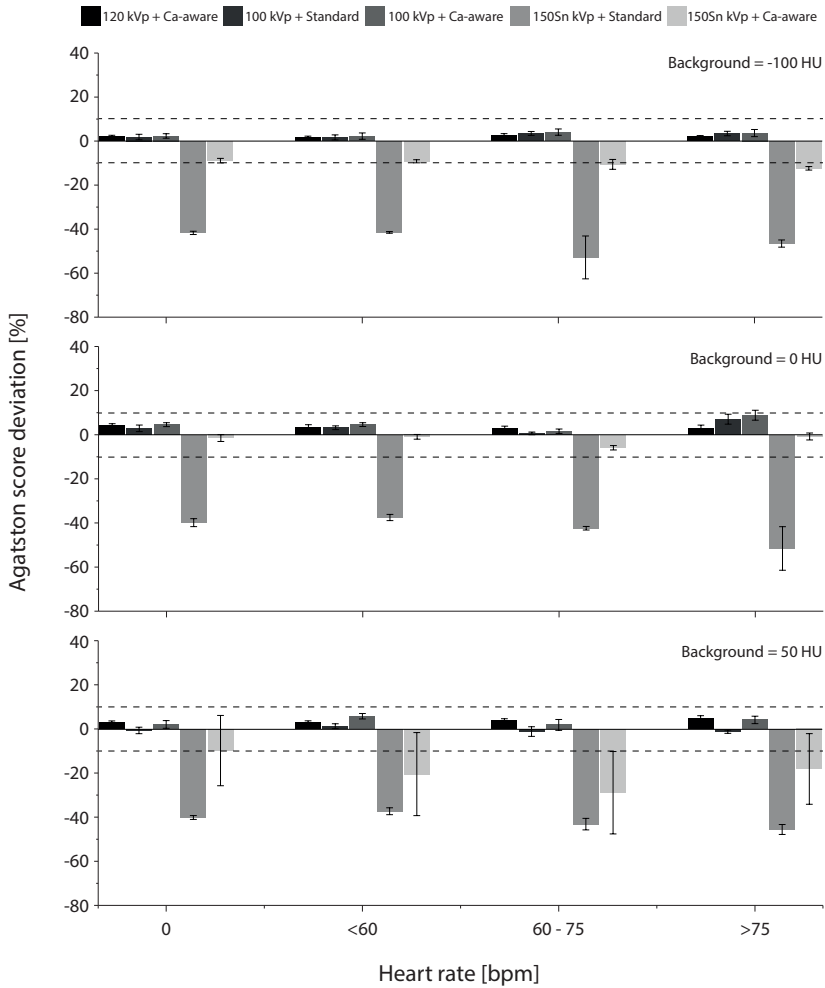
Figure 6 Deviations in medium density CAC Agatston score from the reference (for each heart rate: 120 kVp + standard reconstruction kernel) for different heart rates and combinations of tube voltage [kVp] and reconstruction kernel. Results are shown for three background materials: fat (-100 HU, top), water (0 HU, middle) and soft tissue (50 HU, bottom). Clinically relevant differences, at ≥ ±10%,are indicated with dashed lines

Comparison with physical mass

Physical mass underestimated low-density calcifications by approximately 50% for all reference acquisition and reconstruction settings (*Table 3* and *Table 4*). These underestimations of the physical mass changed to -50%, -41%, and -34% for the reduced-kVp on average for all heart rates, for fat, water and soft tissue adjacent material, respectively. For the tin-filtration protocol, the underestimation of HU values ranged -65%, -53% and -31%. Medium density calcification mass scores underestimated physical mass by -29%, -18% and -30%, again on average for all heart rates, for fat, water and soft tissue circumjacent material, respectively. These values changed to -31%, -20% and -9% for the kVp-reduced protocol, and -36%, -19% and -10% for the tin-filtration protocol.

*Table 3* Percent difference with physical mass of the low density calcification in mean ± SD for all acquisition, reconstruction, background material and heart rate parameters. To calculate the percentage difference, the numerator was the Mass score, and the denominator was the physical mass (38 mg)

Background material	Heart rate	120 kVp		100 kVp		150Sn kVp	
		Standard	Ca-aware	Standard	Ca-aware	Standard	Ca-aware
Fat	0	-47.9 ± 2.2	-47.8 ± 2.3	-45.1 ± 2.0	-46.2 ± 1.8	-73.7 ± 5.4	-63.0 ± 4.1
	<60	-48.9 ± 1.7	-48.8 ± 1.5	-46.0 ± 3.0	-47.2 ± 2.9	-72.8 ± 2.5	-63.0 ± 2.7
	60 - 75	-53.5 ± 0.6	-53.5 ± 0.6	-51.1 ± 1.6	-52.6 ± 1.6	-77.2 ± 8.8	-64.9 ± 3.5
	>75	-55.8 ± 1.8	-55.8 ± 1.8	-54.7 ± 3.5	-56.0 ± 3.5	-78.7 ± 5.3	-69.8 ± 3.8
Water	0	-39.5 ± 2.9	-38.7 ± 2.9	-38.0 ± 2.3	-39.0 ± 2.3	-62.3 ± 3.7	-49.9 ± 2.6
	<60	-42.7 ± 2.3	-42.2 ± 2.4	-37.7 ± 1.1	-38.8 ± 1.0	-59.5 ± 4.1	-47.5 ± 3.7
	60 - 75	-41.4 ± 2.5	-40.7 ± 2.1	-39.1 ± 1.6	-40.7 ± 1.5	-65.3 ± 7.4	-53.4 ± 6.7
	>75	-47.4 ± 2.0	-46.8 ± 1.8	-42.9 ± 0.7	-44.2 ± 1.3	-74.3 ± 8.5	-57.7 ± 3.0
Soft tissue	0	-62.3 ± 3.7	-49.9 ± 2.6	-33.5 ± 2.3	-32.4 ± 2.0	-28.5 ± 2.6	-29.8 ± 2.9
	<60	-59.5 ± 4.1	-47.5 ± 3.7	-39.3 ± 6.4	-37.5 ± 5.9	-26.3 ± 4.0	-27.6 ± 4.1
	60 - 75	-74.3 ± 8.5	-57.7 ± 3.0	-35.3 ± 4.5	-33.4 ± 4.2	-30.5 ± 3.6	-33.6 ± 2.6
	>75	-74.3 ± 8.5	-57.7 ± 3.0	-35.3 ± 4.5	-33.4 ± 4.2	-30.5 ± 3.6	-33.6 ± 2.6

12

Table 4 Percent difference with physical mass of the medium density calcification in mean ± SD for all acquisition, reconstruction, background material and heart rate parameters. To calculate the percentage difference, the numerator was the Mass score, and the denominator was the physical mass (74 mg)

Background material	Heart rate	120 kVp		100 kVp		150Sn kVp	
		Standard	Ca-aware	Standard	Ca-aware	Standard	Ca-aware
Fat	0	-26.7 ± 1.2	-27.3 ± 1.1	-28.2 ± 0.9	-29.1 ± 1.0	-36.1 ± 1.9	-34.0 ± 1.5
	<60	-27.4 ± 0.4	-28.1 ± 0.3	-29.4 ± 1.1	-30.3 ± 1.0	-36.4 ± 1.1	-34.1 ± 1.2
	60 - 75	-29.3 ± 1.4	-29.9 ± 1.4	-30.7 ± 1.3	-31.7 ± 1.4	-49.9 ± 19.5	-37.5 ± 2.6
	>75	-34.0 ± 0.8	-34.5 ± 0.8	-33.1 ± 1.6	-34.2 ± 1.5	-42.8 ± 2.7	-40.0 ± 2.2
Water	0	-17.4 ± 1.7	-17.4 ± 1.7	-18.3 ± 1.9	-19.1 ± 2.1	-21.1 ± 1.3	-16.3 ± 0.7
	<60	-17.7 ± 1.6	-17.6 ± 1.7	-18.5 ± 1.3	-19.2 ± 1.4	-22.0 ± 1.0	-18.4 ± 1.2
	60 - 75	-18.7 ± 1.0	-18.7 ± 1.1	-21.2 ± 1.1	-22.1 ± 1.1	-26.2 ± 2.7	-21.6 ± 2.8
	>75	-20.1 ± 2.1	-20.1 ± 2.0	-18.8 ± 1.9	-19.9 ± 1.9	-38.9 ± 23.5	-21.1 ± 2.4
Soft tissue	0	-21.1 ± 1.3	-16.3 ± 0.7	-10.9 ± 0.9	-10.6 ± 0.9	-11.8 ± 0.8	-12.6 ± 0.7
	<60	-22.0 ± 1.0	-18.4 ± 1.2	-12.6 ± 1.1	-12.2 ± 1.2	-11.4 ± 0.5	-12.1 ± 0.5
	60 - 75	-38.9 ± 23.5	-21.1 ± 2.4	-7.0 ± 1.9	-5.6 ± 2.1	-5.5 ± 2.9	-7.3 ± 2.2
	>75	-38.9 ± 23.5	-21.1 ± 2.4	-7.0 ± 1.9	-5.6 ± 2.1	-5.5 ± 2.9	-7.3 ± 2.2

DISCUSSION

The main finding of this study is that the Ca-aware reconstruction kernel performs well for a patient-specific tube voltage acquisition protocol (12% radiation dose reduction), for medium density CAC, irrespective of CAC adjacent material or heart rate. However, in the presence of low density CAC substantial deviations in Agatston scores were observed, when calcifications were surrounded by water (up to 15%) or soft-tissue (up to 22%) equivalent material, irrespective of heart rate. Furthermore, the tin-filtration protocol also led to substantial deviations in Agatston scores for low density calcifications, for most combinations of heart rates and surrounding tissue. Furthermore, noise levels for this protocol were high, leading to BAS > 0. Finally, mass scores as assessed by CT underestimated the true physical mass.

To the best of our knowledge, this study is the first to systematically assess the performance of a novel Ca-aware reconstruction kernel for different CAC surrounding materials, CAC densities, and heart rates. In general, and especially for the low-density calcification, reduced-kVp acquisitions resulted in increased Agatston scores. This is expected, as the energy-dependence of CT numbers of the surrounding material (fat / water / soft-tissue) is different from the energy-

dependence of the CT number of calcium, as previously described by Jakobs et al.[24] Because of the phenomenon, the detectability of calcium is increased, especially at the margins of CAC where voxels might be just below the calcium scoring threshold for 120kVp. In turn, as more CAC is detected, more voxels are taken into account by the Ca-aware reconstruction kernel for its recalculation to 120 kVp HU values.[14]

Our results are in line with a phantom study by Booij et al., who demonstrated that the consistency of CT numbers was reduced for low density CAC when comparing CT numbers from reduced tube voltage acquisitions reconstructed with the Ca-aware kernel, and CT numbers from routine protocols.[21] In their study, however, a base-material correction factor was provided by the CT manufacturer.[14] This correction factor was necessary to account for the tube voltage dependency of all materials other than water. The usage of this correction factor hampers direct comparison with our results, as this artificial step was necessary due to the nature of the used phantom in their study, which might have influenced resulting Agatston scores.

In addition, two patient studies have been carried out by Vingiani et al.[22,23] Both studies showed the feasibility of the Ca-aware reconstruction kernel, in combination with patient specific tube voltages, where one study considered spectral beam shaping with tin-filtration. Comparison of results is hampered by the fact that for their first study 100 kVp with tin-filtration was applied, whereas in our study a tube voltage of 150 kVp with tin-filtration was manually selected for the phantom.[22] In the other study by Vingiani et al. forty-three patients were imaged with both 120 kVp and an individualized tube voltage.[23] A high concordance in Agatston scores between both scans was found. Since the density of these CAC, and the HU of the CAC surrounding material is unknown, it is not known if these results are in-line or contradictory to our results.

In line with previous studies, we found that CT generally underestimates the physical mass of the low density calcifications by approximately 50% for all reference acquisition and reconstruction settings.[20]

Our study has some limitations that merit consideration. First, this was an in-vitro study, with artificial CAC containing coronary arteries and artificial background material. Nevertheless, the coronary arteries were translated in an anthropomorphic chest phantom at velocities which were observed in in-vivo studies.[18] Also, the mass of the calcifications was in the range which is observed in patients.[25] Second, movement of the artificial coronary arteries was only linear and

in the horizontal plane, while in-vivo complex movements in three dimensions are observed. As the actual scan phase of a CAC scan is only (104 ms, based on the total detector coverage and rotation time, we approximate that the addition of 3D movement would only result in minor changes in our results. Third, the materials used to simulate in-vivo CAC circumjacent tissue are artificial. However, the linear attenuation coefficient of the materials is only used for the Ca-aware reconstruction kernel. The exact chemical composition is therefore irrelevant for the current analysis, as we changed the ratio of our mixtures for each tube voltage to ensure stable background material HU. Fourth, tube current limitations of the dedicated CAC tin-filtration protocol led to the usage of semi-automatic tube voltage selection for our tin-filtration acquisitions and a manual selection of 150Sn kVp. However, calcium contrast is inherently low for this hardened X-ray spectrum, resulting in reduced CAC detectability and quantification results. Although available to be manually selected by CT radiographers, this 150Sn protocol is not recommended by the CT manufacturer for CAC scoring for large patients. In addition, a low dose value was selected for the 150Sn kVp protocol which resulted in increased noise levels and may have affected the determination of the Agatston score of the calcifications and may have led to Agatston scores for a non-CAC containing ROI (BAS > 0). This means that the resulting Agatston scores for these acquisitions might be overestimated. However, for all combinations of surrounding material, CAC density and heart rate, clinically relevant decrease in Agatston score was shown. This decrease should therefore be even larger, when noise levels were lower.

CT CAC-scoring with varying surrounding materials and heart rates is feasible at patient-specific tube voltages with the novel Ca-aware reconstruction technique.
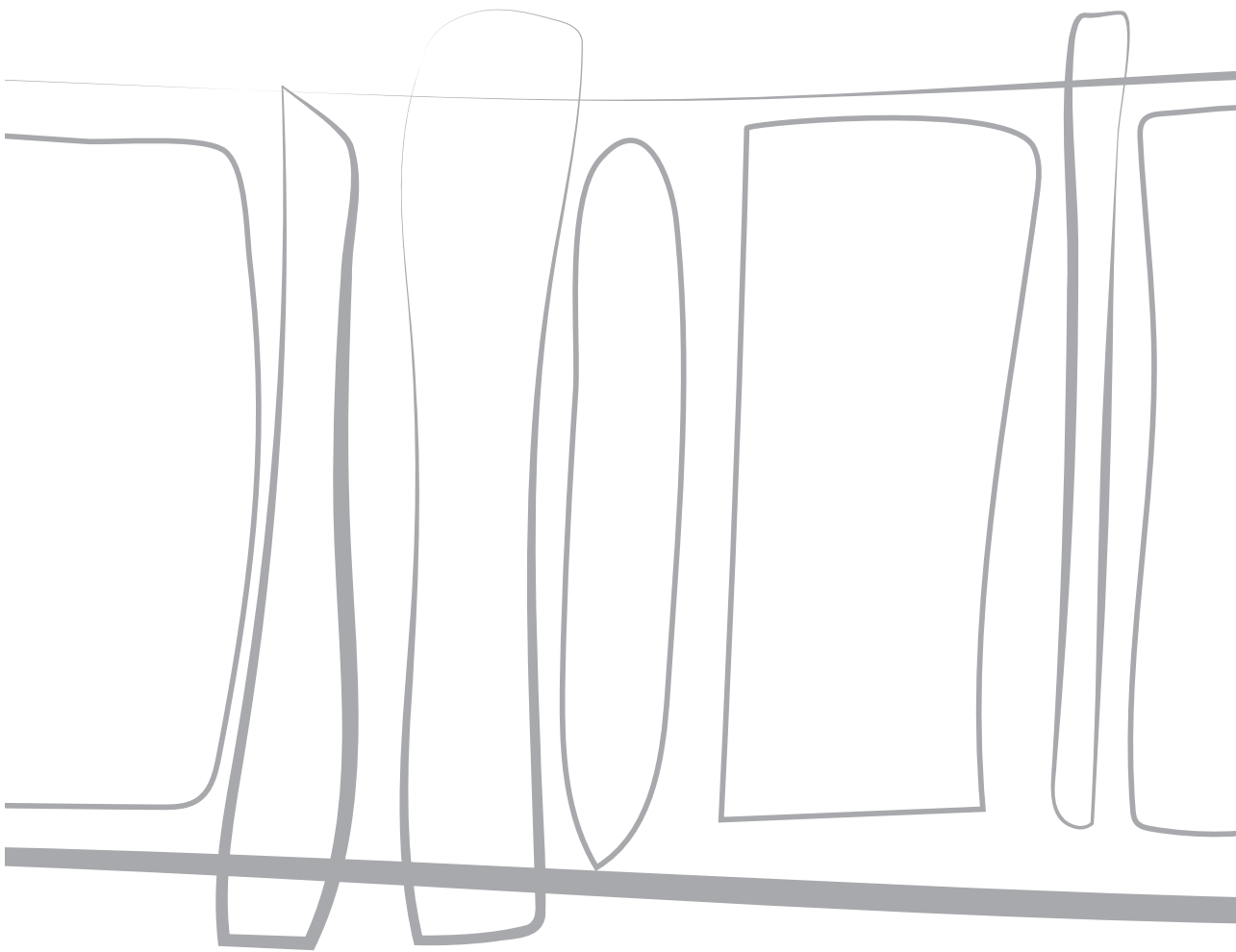
12

REFERENCES

1. Detrano R, Guerci AD, Carr JJ, et al. Coronary calcium as a predictor of coronary events in four racial or ethnic groups. N Engl J Med. 2008;358(13):1336-1345. doi:10.1056/NEJMoa072100

2. Shaw LJ, Raggi P, Schisterman E, Berman DS, Callister TQ. Prognostic value of cardiac risk factors and coronary artery calcium screening for all-cause mortality. Radiology. 2003;228(3):826-833. doi:10.1148/radiol.2283021006

3. Keelan PC, Bielak LF, Ashai K, et al. Long-term prognostic value of coronary calcification detected by electron-beam computed tomography in patients undergoing coronary angiography. Circulation. 2001;104(4):412-417. doi:10.1161/hc2901.093112

4. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. J Am Coll Cardiol. 1990;15(4):827-832. doi:10.1016/0735-1097(90)90282-T

5. Hong C, Bae KT, Pilgram TK. Coronary Artery Calcium: Accuracy and Reproducibility of Measurements with Multi–Detector Row CT—Assessment of Effects of Different Thresholds and Quantification Methods. Radiology. 2003;227(3):795-801. doi:10.1148/radiol.2273020369

6. Detrano RC, Anderson M, Nelson J, et al. Coronary calcium measurements: effect of CT scanner type and calcium measure on rescan reproducibility--MESA study. Radiology. 2005;236(2):477-484. doi:10.1148/radiol.2362040513

7. Hoffmann U, Siebert U, Bull-Stewart A, et al. Evidence for lower variability of coronary artery calcium mineral mass measurements by multi-detector computed tomography in a community-based cohort--consequences for progression studies. Eur J Radiol. 2006;57(3):396-402. doi:10.1016/j.ejrad.2005.12.027

8. Rutten A, Isgum I, Prokop M. Coronary calcification: effect of small variation of scan starting position on Agatston, volume, and mass scores. Radiology. 2008;246(1):90-98. doi:10.1148/radiol.2461070006

9. Groen JM, Greuter MJ, Schmidt B, Suess C, Vliegenthart R, Oudkerk M. The influence of heart rate, slice thickness, and calcification density on calcium scores using 64-slice multidetector computed tomography: A systematic phantom study. Invest Radiol. 2007;42(12):848-855. doi:10.1097/RLI.0b013e318154c549

10. Willemink MJ, van der Werf NR, Nieman K, Greuter MJW, Koweek LM, Fleischmann D. Coronary artery calcium: A technical argument for a new scoring method. J Cardiovasc Comput Tomogr. 2019;13(6):347-352. doi:10.1016/j.jcct.2018.10.014

11. Mahabadi AA, Mohlenkamp S, Lehmann N, et al. CAC Score Improves Coronary and CV Risk Assessment Above Statin Indication by ESC and AHA/ACC Primary Prevention Guidelines. JACC Cardiovasc Imaging. 2017;10(2):143-153. doi:10.1016/j.jcmg.2016.03.022

12. Hecht HS, Blaha MJ, Kazerooni EA, et al. CAC-DRS: Coronary Artery Calcium Data and Reporting System. An expert consensus document of the Society of Cardiovascular Computed Tomography (SCCT). J Cardiovasc Comput Tomogr. 2018;12(3):185-191. doi:10.1016/j.jcct.2018.03.008

12

13. McCollough CH, Ulzheimer S, Halliburton SS, Shanneik K, White RD, Kalender WA. Coronary Artery Calcium: A Multi-institutional, Multimanufacturer International Standard for Quantification at Cardiac CT. Radiology. 2007;243(2):527-538. doi:10.1148/radiol.2432050808

14. Allmendinger T, Hamann A. Agatston score calcium quantification with arbitrary tube voltage [white paper]. Siemens Healthineers. Published 2019. https://www.siemens-healthineers.com/computed-tomography/clinical-imaging-solutions/cardiovascular-imaging#CLINICAL_USE

15. Willemink MJ, Abramiuc B, den Harder AM, et al. Coronary calcium scores are systematically underestimated at a large chest size: A multivendor phantom study. J Cardiovasc Comput Tomogr. 2015;9(5):415-421. doi:10.1016/j.jcct.2015.03.010

16. Marwan M, Mettin C, Pflederer T, et al. Very low-dose coronary artery calcium scanning with high-pitch spiral acquisition mode: Comparison between 120-kV and 100-kV tube voltage protocols. J Cardiovasc Comput Tomogr. 2013;7(1):32-38. doi:10.1016/j.jcct.2012.11.004

17. Achenbach S, Ropers D, Holle J, Muschiol G, Daniel WG, Moshage W. In-plane coronary arterial motion velocity: measurement with electron-beam CT. Radiology. 2000;216(2):457-463. doi:10.1148/radiology.216.2.r00au19457

18. Husmann L, Leschka S, Desbiolles L, et al. Coronary artery motion and cardiac phases: dependency on heart rate -- implications for CT image reconstruction. Radiology. 2007;245(2):567-576. doi:10.1148/radiol.2451061791

19. Rutten A, Krul SPJ, Meijs MFL, De Vos AM, Cramer MJM, Prokop M. Variability of coronary calcium scores throughout the cardiac cycle: Implications for the appropriate use of electrocardiogram-dose modulation with retrospectively gated computed tomography. Invest Radiol. 2008;43(3):187-194. doi:10.1097/RLI.0b013e31815cdd56

20. van der Werf NR, Willemink MJ, Willems TP, Vliegenthart R, Greuter MJW, Leiner T. Influence of heart rate on coronary calcium scores: a multi-manufacturer phantom study. Int J Cardiovasc Imaging. 2017;34(6):959-966. doi:10.1007/s10554-017-1293-x

21. Booij R, van der Werf NR, Budde RPJ, Bos D, van Straten M. Dose reduction for CT coronary calcium scoring with a calcium-aware image reconstruction technique: a phantom study. Eur Radiol. 2020;30(6):3346-3355. doi:10.1007/s00330-020-06709-9

22. Vingiani V, Abadia AF, Schoepf UJ, et al. Low-kV coronary artery calcium scoring with tin filtration using a kV-independent reconstruction algorithm. J Cardiovasc Comput Tomogr. 2020;14(3):246-250. doi:10.1016/j.jcct.2019.11.006

23. Vingiani V, Abadia AF, Schoepf UJ, et al. Individualized coronary calcium scoring at any tube voltage using a kV-independent reconstruction algorithm. Eur Radiol. Published online May 2020. doi:10.1007/s00330-020-06951-1

24. Jakobs TF, Wintersperger BJ, Herzog P, et al. Ultra-low-dose coronary artery calcium screening using multislice CT with retrospective ECG gating. Eur Radiol. 2003;13(8):1923-1930. doi:10.1007/s00330-003-1895-7

25. Hong C, Pilgram TK, Zhu F, Bae KT. Is coronary artery calcium mass related to Agatston score? Acad Radiol. 2004;11(3):286-292. doi:10.1016/s1076-6332(03)00714-1

12

12

Online material is available via:
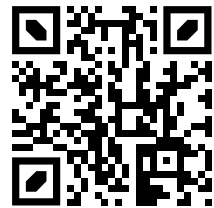
CHAPTER 13

Systematic assessment of coronary calcium detectability and quantification on four generations of CT reconstruction techniques: a patient and phantom study

Magdalena Dobrolinska*, MD

Gijs D. van Praagh*, MSc

Luuk J. Oostveen*, BSc

Keris Poelhekken, BSc

Marcel J.W. Greuter, PhD

Dominik Fleischmann, MD

Martin J. Willemink, MD PhD

Frank de Lange, PhD

Riemer H.J.A. Slart, MD PhD

Tim Leiner, MD PhD

Niels R. van der Werf, MSc

*Authors contributed equally to this work

ABSTRACT

Background

In computed tomography, coronary artery calcium (CAC) scores are influenced by image reconstruction. For a newly introduced deep learning-based reconstruction (DLR), the effect on CAC scoring in relation to other algorithms is unknown. The aim of this study was to evaluate the effect of four generations of image reconstruction techniques (filtered back projection (FBP), hybrid iterative reconstruction (HIR), model-based iterative reconstruction (MBIR), and DLR) on CAC detectability, quantification, and risk classification.

Methods

First, CAC detectability was assessed with a dedicated static phantom containing 100 small calcifications varying in size and density. Second, CAC quantification was assessed with a dynamic coronary phantom with velocities equivalent to heart rates of 60-75 bpm. Both phantoms were scanned and reconstructed with four techniques. Last, scans of fifty patients were included and the Agatston calcium score was calculated for all four reconstruction techniques. FBP was used as a reference.

Results

In the phantom study, all reconstruction techniques resulted in less detected small calcifications for both, static and dynamic phantom. In the patient study, the cardiovascular risk classification resulted, for all reconstruction techniques, in excellent agreement with the reference, although MBIR resulted in significantly higher Agatston scores and 6% reclassification rate.

Conclusion

Agatston score agreement between FBP, HIR, and DLR was excellent, with a low-risk reclassification rate.

13

INTRODUCTION

Coronary artery calcium (CAC) is important for cardiovascular risk determination in asymptomatic individuals.[1] CAC is visualized with cardiac computed tomography (CT) and quantified using the Agatston score.[2] Furthermore, an Agatston score of zero is proven to be a strong negative predictor of future cardiovascular events.[3] This, in turn, indicates the importance of accurate detection and subsequent quantification of small calcified lesions.

One important factor influencing CAC quantification is the type of image reconstruction used in CT.[4] Over the last decade advanced reconstruction techniques such as hybrid iterative reconstruction (HIR) and model-based iterative reconstruction (MBIR) became available for CT.[5] These reconstruction algorithms reduce image noise, and therefore allow for a decrease in radiation dose while maintaining image quality equal to traditional filtered back projection (FBP).[6,7] Previous studies have shown a good agreement in Agatston scores between FBP and HIR and MBIR.[8–10] However, it was also shown that HIR resulted in decreased Agatston scores for small and/or low density lesions.[9] Similarly, MBIR resulted in decreased detection of small calcifications.[8]

Recently, one of the main CT manufacturers introduced a new deep learning-based reconstruction (DLR) technique. DLR improves image quality by applying a deep learning network trained on pairs of high-dose, advanced MBIR and HIR images and prevents image quality degradation and 'plastic-like' appearance of the image.[11,12] As previously shown with low dose acquisitions, DLR outperforms MBIR in terms of noise reduction which may potentially allow for further radiation dose reduction beyond current levels.[11,13] However, the influence of this novel image reconstruction technique on CAC detection and quantification is unknown.

As previously noted, the detection of CAC, resulting subsequently in zero or non-zero Agatston scores, is of utmost importance for correct risk stratification. Because small or low-density CAC can resemble image noise and HIR, MBIR, and DLR all decrease image noise, these CT reconstruction techniques may impact the detection of very small or low-density CAC. This is even more important for acquisitions at a reduced radiation dose.[14] As previously shown, risk classification was underestimated up to 50% for CAC scores from IR images acquired at reduced radiation dose.[4] Consequently, the Society of Cardiovascular Computed Tomography recommends further evaluation of reconstruction techniques before clinical implementation.[15]

Therefore, we designed a phantom study in which we aimed to investigate the influence of four reconstruction methods (FBP, HIR, MBIR, and DLR) on static and dynamic CAC detectability and quantification for standard and reduced radiation dose.

13

Subsequently, we verified the effect of all four image reconstruction techniques on CAC quantification and risk classification in a patient study.

METHODS

Phantom study

Phantom

In this study, a small patient (300x200 mm) anthropomorphic thorax phantom (Thorax, QRM, Möhrendorf, Germany) was used (*Figure 1*).[16] To simulate large patient dimensions, an extension ring (Extension ring, QRM, Möhrendorf, Germany) of fat tissue equivalent material was used to increase the outer dimensions of the phantom to 400x300 mm.

CAC detectability was assessed with a cylindrical insert (D100, QRM, Möhrendorf, Germany). This insert contained one hundred small cylindrical calcifications differing in size (0.5 to 2.0 mm in diameter and length) and density (90 to 540 mg hydroxyapatite (HA)/cm3).[17]

CAC quantification was assessed with the use of a dynamic artificial coronary artery, which was translated by a computer-controlled lever (Sim2D, QRM, Möhrendorf, Germany) in a water-filled compartment in the thorax phantom (*Figure 1*).
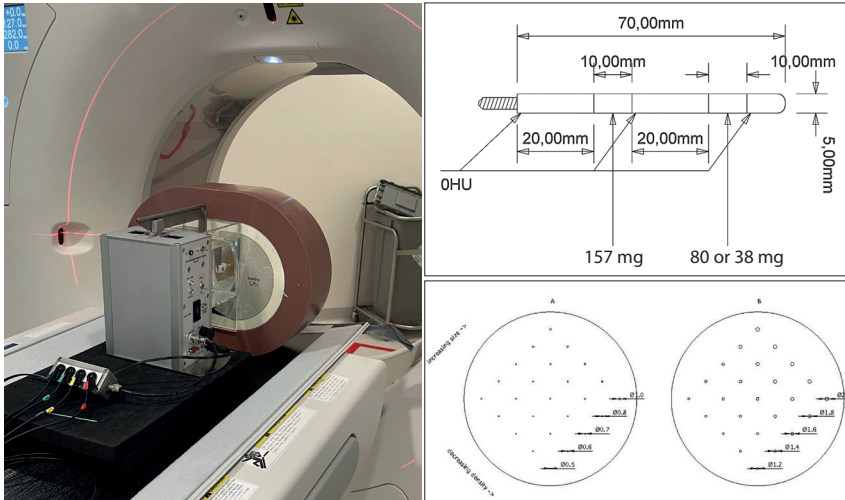

Figure 1 Overview of the phantom setup with the anthropomorphic thorax phantom and motion controller (left), and a schematic overview of the used calcium inserts for dynamic scan (upper right) and for static scan (bottom right).

During acquisition, the artery remained static or moved at a constant velocity of 20 mm/s in the horizontal plane during the scan phase, simulating a heart rate of 0 or 60-75 bpm, respectively.[18,19] Two arteries were used containing three cylindrical calcifications with equal dimensions (diameter: 5 mm, length: 10 mm), but varying densities of 196±3, 408±2 and 800±2 mgHA/cm3, designated as low, medium, and high density, respectively (*Figure 1*).

Data acquisition

Both phantom sizes were scanned on a state-of-the-art 320 slice CT system (Aquilion One PRISM edition, Canon Medical Systems, Otawara, Japan) with routinely used clinical CAC protocols (*Table 1*). Automatic tube current selection (SureExposure 3D, Canon Medical Systems, Otawara, Japan) was used to select appropriate radiation dose levels for the small and large phantom size. The reference level was based on setting the automatic tube current modulation to a standard deviation (SD) of 27.76 at 3 mm, with 40 and 300 mA as the minimum and maximum tube current, respectively. Next, tube current was reduced to 75%, 50%, and 25% of the clinical radiation dose. Raw data was acquired at 120 kVp. Besides raw data reconstruction with FBP, three other reconstruction methods were used: HIR (adaptive iterative dose reduction 3D; AIDR 3D enhanced), MBIR (forward projected model based iterative reconstruction solution; FIRST standard), and DLR (advanced intelligent clear-IQ engine; AiCE standard) (*Table 1*). Each protocol was repeated ten times for the D100 and five times for the dynamic phantom. A larger number of repetitions was used for the D100 phantom, as the small size of the calcifications (≤ 2mm) was highly impacted by partial volume effects due to the 3mm slice thickness. Between each scan the phantom was manually repositioned (approximately 2 mm translational and 2 degrees rotational) to assess interscan variability.

CAC detection and Agatston score calculations on the phantom scans were performed using a validated fully automated quantification method with vendor specific CAC scoring parameters.[20] A standard CAC scoring threshold of 130 Hounsfield units (HU) was used.[2]

For each scan, a background Agatston score (BAS) was calculated, as described previously by Booij et al.[21] For scans with a nonzero BAS, it was unknown if a CAC was detected or if the score was based on noise only. For CAC detection purposes, a scan with a nonzero BAS was therefore defined as non-diagnostic and was omitted from further analysis.

13

Table 1 CAC acquisition and reconstruction parameters for phantom and patient study

	Phantom study	**Patient study**
Acquisition mode	Axial	Axial
ECG-triggering	Prospective	Prospective
Peak tube potential [kVp]	120 / 100	120
Reference image noise [HU]	27.76	27.76
Rotation time [s]	0.275	0.275
Field of view [mm]	220 x 220	Patient specific
Matrix size [pixels]	512 x 512	512 x 512
Slice thickness / increment [mm]	3.0 / 3.0	3.0 / 3.0
Reconstruction kernel	FC12	FC12
Reconstruction algorithm	FBP / AIDR 3D standard / AIDR 3D enhanced / FIRST standard / AiCE standard	FBP / AIDR 3D enhanced / FIRST standard / AiCE standard

ECG = electrocardiogram; bpm = beats per minute; FBP = filtered back projection.

Patient study

A patient study was performed to assess differences in Agatston scores resulting from the application of different reconstruction algorithms. This retrospective study was approved by the local ethics committee (CMO 2016-3045, Project 20045), who waived the requirement for patient informed consent after de-identification of all patient information from the study data. Raw data was acquired on the same CT system as used for the phantom scans, in a consecutive cohort of 50 patients between July and October 2020 (*Table 2*). All patients were scanned at 120 kVp. Raw data was reconstructed using the same four reconstruction methods as for the phantom studies: FBP, HIR, MBIR, and DLR.

Agatston scores in patient scans were determined using a dedicated workstation (Vitrea 7.11; Vital Images Inc.).

Table 2 Patients' characteristics of the 50 patients included in the study

Patients' characteristics	
Median age (range) [years]	60 (41 – 77)
Female	32 (64%)
Heart rate (range) [bpm]	60 (57 – 68)*
Median Agatston score (range)	61 (0 – 2935)

*For 2 patients, heart rate could not be retrieved retrospectively

Statistical analysis

Agatston scores resulting from the default clinical protocol (120 kVp, 100% dose, FBP) were used as the reference for both the phantom and patient study. Scores from other acquisition and reconstruction settings were compared with this reference. For the phantom study, the comparison was performed within the same repetition. For the different combinations of radiation dose, and reconstruction method, deviations of more than 10% in Agatston score from the reference were considered clinically relevant.[22] Categorical variables and number of detected calcifications were presented as percentages. Depending in the distribution of the data, continuous variables were presented as means with standard deviation (±SD) or medians with interquartile region (IQR, 1st – 3rd).

Patient Agatston scores resulting from the different reconstruction techniques were compared with the reference score (120 kVp, FBP) using Bonferroni corrected Wilcoxon signed-rank tests. Next, patients were divided into five risk groups (0 Agatston score – 0; 0.1 to 10 Agatston score – 1; 10.1 to 100 Agatston score – 2; 100.1 to 400 Agatston score – 3; >400 Agatston score – 4) and the agreement in risk classification between the different reconstruction methods was compared based on a Cohen weighted linear κ with 95% confidence intervals (95% CI). The cardiac risk classification was determined for each patient and each reconstruction technique.[23] The agreement between FBP Agatston score and HIR, MBIR, and DLR Agatston score was analysed with Bland-Altman plots. A false-positive result was defined as a calcification not detected on the reference scan, a false-negative result was defined as calcification detected on the reference scan but not on the HIR, MBIR, or DLR scan. P values smaller than 0.05 were considered statistically significant. SPSS version 25 (IBM Corp., Armonk, NY, USA) was used for statistical analyses.

RESULTS

Phantom study

Full dose settings resulted in 80 and 300 mA for the small and large phantom, respectively. Tube currents were reduced to the nearest available setting to obtain 75%, 50% and 25% of the full dose setting. The resulting volume CT dose indexes (CTDIvol) for 100% dose setting were 1.2 mGy (120 kVp) for the small phantom and 4.4 mGy (120 kVp) for the large phantom.

CAC detectability

For all used reconstruction algorithms, the CT numbers for a calcification with a density of 300 mgHA/cm³ and varying sizes within the small phantom are depicted in *Figure 2*. This figure shows a difference in the HU peak reached by each of the reconstruction methods, whereby the CAC scoring threshold of 130HU is not reached for the smallest calcification by MBIR and DLR.
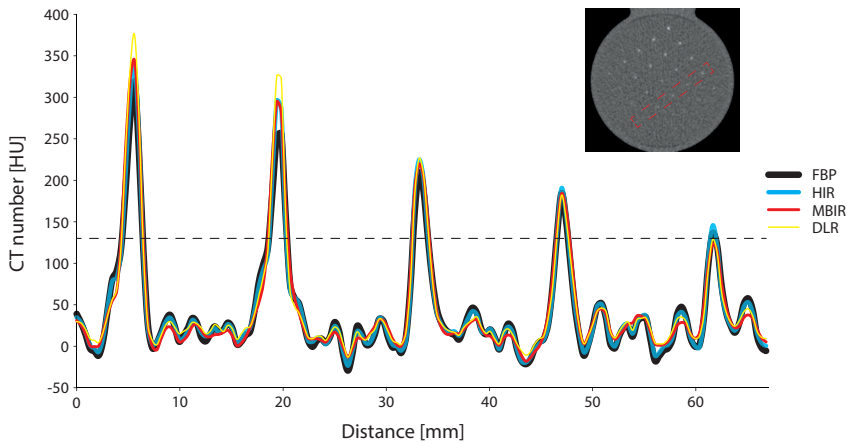
Figure 2 A profile plot through calcifications with 300 mgHA/cm3 of 2, 1.8, 1.6, 1.4 and 1.2 mm in diameter, respectively, as indicated in the red box in the right-upper image. These plots summarize the difference between the four reconstruction methods (FBP, HIR, MBIR, DLR). The conventional CAC scoring threshold of 130 HU is indicated with a dotted line. As depicted on the plot, calcifications of the lowest diameter reconstructed with MBIR and DLR, do not reach the 130 HU threshold

FBP – filtered back projection, HIR – hybrid iterative reconstruction, MBIR – model-based iterative reconstruction, DLR – deep learning-based reconstruction

For all repeated scans, the reference protocol resulted in a CAC detection of 150 and 87 calcifications out of 1000 for the small and large phantom, respectively. Relative results for the other reconstruction algorithms and dose levels are shown in *Figure 3* and *supplementary Figure 1*.

For the small phantom at full dose, MBIR, and DLR resulted in 4%, and 1% less detected calcifications, while 8% more calcifications were detected with HIR. For the large phantom at full dose, 2%, 22%, and 9% less calcifications were detected for HIR, MBIR, and DLR, respectively.
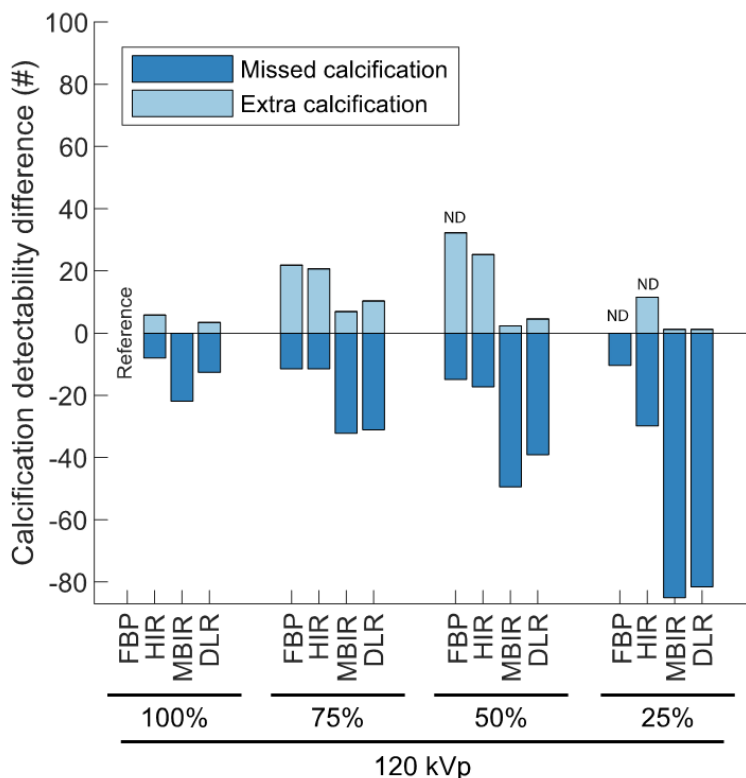
Figure 3 Difference in total number of detected calcifications of the static (D100) insert in the large thorax phantom for all combinations of tube currents (in percentage of reference) and reconstruction methods compared with the reference (120 kVp, 100% dose, FBP). For each repetition, a calcification was defined as 'missed' when the calcification was detected with the reference protocol but was not detected with varying acquisition and/or reconstructions parameters. The opposite was defined as an 'extra calcification'. All repetitions with BAS>0 were defined as nondiagnostic (ND) image quality and were therefore omitted from the analysis.

FBP – filtered back projection, HIR – hybrid iterative reconstruction, MBIR – model-based iterative reconstruction, DLR – deep learning-based reconstruction, # - number

For the small phantom, 75% dose with 120 kVp resulted in 7%, 2%, 55%, and 59% less detected calcifications for FBP, HIR, MBIR, and DLR, respectively. In the large phantom the reduction was even larger, with 10%, 18%, 84%, and 80% less detected calcifications, respectively. The number of missed calcifications was even more pronounced for 50% and 25% dose (*Figure 3* and *supplementary Figure 1*).

253

CAC quantification

For the small static phantom, median (IQR) Agatston scores were 96 (95 – 108), 350 (344 – 363), and 413 (403 – 427) for the low-, medium-, and high-density CAC in the reference protocol. At 60-75 bpm, these Agatston scores changed to 87 (82 – 88), 379 (368 – 419), and 474 (464 – 513) (*Supplementary* Figure S2). This resulted in the overall change of Agatston score by -22%, 9%, and 25% for low, medium, and high-density calcifications, respectively (*Supplementary* Figure S2).

For the large static phantom, Agatston scores were 74 (70 – 82), 303 (301 – 306), and 381 (379 – 388) for the low-, medium-, and high-density CAC (*Figure 4*). These Agatston scores changed at 60-75 bpm to 48 (42 – 67), 355 (348 – 361), and 503 (469 – 515). Briefly, for the large phantom Agatston scores increased compared to the static situation by 10.4% (-49% to 115.2%), 200% (103.2% to 346%), and 189.5% (120.3% to 400.6%) for the low-, medium-, and high- density calcifications, respectively (*Figure 4*).

As compared to reference Agatston scores, deviations in Agatston scores for data reconstructed with the other reconstruction methods, were non-relevant (<10%) (*Figure 4* and *Supplementary Figure 2*). For 120 kVp with 50% radiation dose, the majority of reconstruction methods resulted in small non-relevant deviations in Agatston score, as depicted on *Figure 4*.

Patient study

The age range of the 50 patients was 41 – 77 years with a median age of 60 years and 32 (64%) patients were female. Median dose length product for the calcium scoring acquisitions was 60.2 mGycm (full range: 30.8 – 73.6 mGycm) corresponding to an estimated effective dose of 1.56 (0.8-1.91) mSv using a conversion factor of 0.026 mSv. mGycm^{-1}.[24]

CAC quantification and detectability

The median (IQR) Agatston scores were 61 (5.5 – 435.0), 63 (8.5 – 412.0), 81.5 (9.25 – 435.0), and 72.5 (9.25 – 401.0), for FBP, HIR, MBIR, and DLR, respectively. Only MBIR Agatston scores were significantly different from FBP (p<0.001). Within all reconstruction methods, only for MBIR one false-positive calcification was detected. Additionally, differences in Agatston score between FBP and HIR, MBIR, and DLR, increased with increasing Agatston scores (*Figure 5*). The difference between the four reconstruction methods in calcium detection is depicted on *Figure 6*.
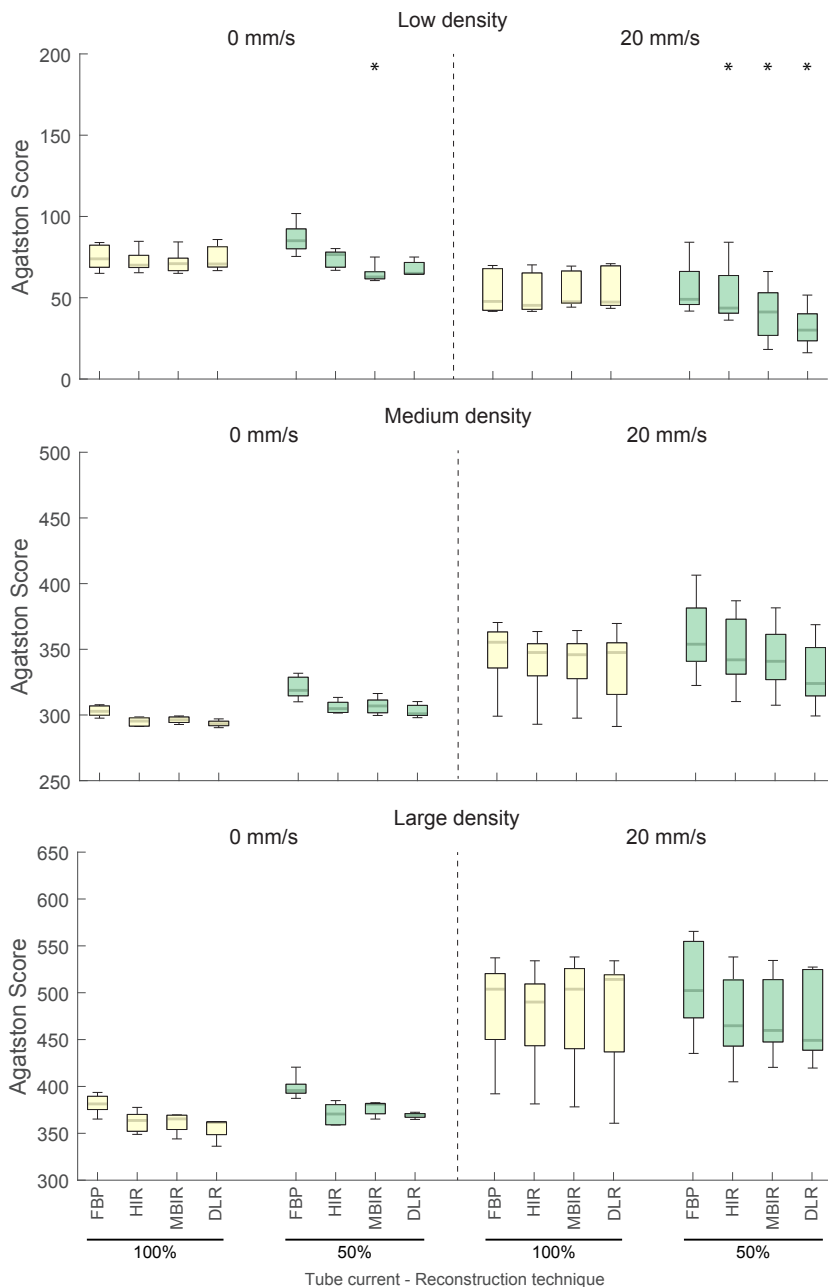
Figure 4 Large phantom static and dynamic Agatston scores for the low (top), medium (middle) and high (bottom) density calcifications, for all used tube current (in percentage of reference), and reconstruction methods. Asterix (*) marks a protocol that results in a clinically relevant (>10% change) median Agatston score change compared with the reference protocol.
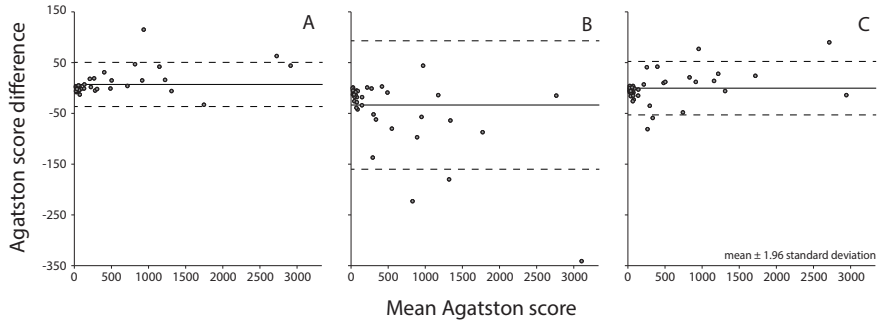
Figure 5 Bland-Altman plots of the difference between HIR (A), MBIR (B) and DLR (C) and FBP for all fifty patients. Agatston score difference was calculated as FBP Agatston score minus IR Agatston score. The solid line resembles a mean difference, the dashed lines resemble standard deviation.
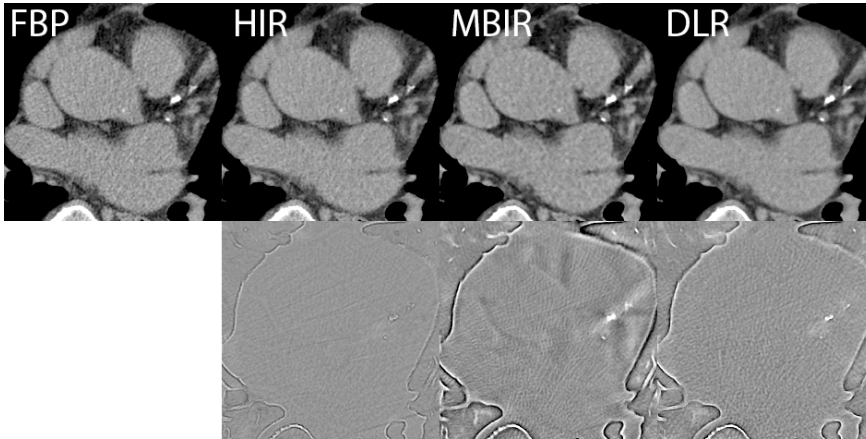
Figure 6 CAC detection for FBP, HIR, MBIR and DLR for one of the patients. Window width (WW) and window level (WL) setting were constant for the upper row: WW = 380, WL = 40. The bottom row shows subtraction images between FBP and HIR, MBIR and DLR. For MBIR, overall CAC quantification increases with respect to FBP.

FBP – filtered back projection, HIR – hybrid iterative reconstruction, MBIR – model-based iterative reconstruction, DLR – deep learning-based reconstruction

Risk classification

Overall, the agreement between cardiovascular risk classification based on FBP compared to HIR, MBIR, and DLR was excellent (κ = 0.97, 95% CI: 0.94 – 1.0; κ = 0.96, 95% CI: 0.92 – 1.0; κ = 0.97, 95% CI: 0.94 – 1.0) (*Table 3 a-c)*. However, based on MBIR, three patients (1.5%) were included in a higher risk category as

compared to FBP. Within these patients, one was reclassified from zero to a non-zero Agatston score. For HIR as well as for DLR, reclassification occurred in two cases (*Table 3 a-c*). In both reconstruction methods one case to a lower category and one to a higher category.

Table 3 a-c. The agreement between patient risk classification based on FBP and risk classification based on MBIR, HIR, and DLR respectively

a)

		Risk classification based on FBP					
		0	1	2	3	4	
	0	10	0	0	0	0	10 (20 %)
	1	1	1	0	0	0	2 (4 %)
Risk classification based on MBIR	2	0	1	16	0	0	17 (34 %)
	3	0	0	1	7	0	8 (16 %)
	4	0	0	0	0	13	13 (26 %)
		11(22 %)	2(4 %)	17(34 %)	7(14 %)	13(26 %)	50

b)

		Risk classification based on FBP					
		0	1	2	3	4	
	0	11	0	0	0	0	11 (22 %)
	1	0	1	0	0	0	1 (2 %)
Risk classification based on HIR	2	0	1	17	0	0	18 (36 %)
	3	0	0	0	7	1	8 (16 %)
	4	0	0	0	0	12	12 (24 %)
		11(22 %)	2(4 %)	17(34 %)	7(14 %)	13(26 %)	50

c)

		Risk classification based on FBP					
		0	1	2	3	4	
	0	11	0	0	0	0	11 (22 %)
	1	0	1	0	0	0	1 (2 %)
Risk classification based on DLR	2	0	1	17	0	0	18 (36 %)
	3	0	0	0	7	1	8 (16 %)
	4	0	0	0	0	12	12 (24 %)
		11(22 %)	2(4 %)	17(34 %)	7(14 %)	13(26 %)	50

Risk groups are defined as follows: 0 Agatston score – 0; 0.1 to 10 Agatston score – 1; 10.1 to 100 Agatston score – 2; 100.1 to 400 Agatston score – 3; >400 Agatston score – 4.

DISCUSSION

The main finding of the phantom part in the present study is that detection of small calcifications at routine (100%) radiation dose is reduced up to 22% depending on the used reconstruction algorithm. Furthermore, this trend was even more pronounced on reduced radiation dose scans. For CAC quantification, our dynamic phantom study showed no clinically relevant differences in Agatston score based on reconstruction algorithm for the routine radiation dose protocol. The patient study showed excellent agreement between FBP and HIR, MBIR, and DLR, with only a small number of risk reclassifications, although MBIR resulted in significantly higher Agatston scores.

To the best knowledge of the authors, this study is the first to systematically assess the influence of all reconstruction techniques available for one vendor on CAC detection and quantification. Compared to FBP all reconstruction methods reduced CAC detection, except in the case of the small chest phantom at full dose level. Both IR techniques as well as DLR reduced image noise.[11] The, in general, reduced CAC detectability in comparison with FBP for these reconstruction techniques might therefore be explained by erroneous identification of CAC containing voxels as noise. Furthermore, as we presented in this study, decreased detectability may result from a reduced peak of HU value of small calcification. This behavior will, of course, be more pronounced at reduced tube current and increased patient size, where noise levels are increased. As a result, HIR, MBIR, and DLR may miss small calcifications and improperly classify patients into the zero Agatston score risk group. However, based on our patient study, none of the patients was incorrectly assigned to the zero Agatston score group.

Independent of the reconstruction method, for medium and large density calcifications, the Agatston score increased with velocity, while for small density calcification, Agatston score decreased. This finding is in line with previous results of van der Werf et al. and Groen et al. and might be explained by motion blurring.[19,25] Due to motion blurring, the number of voxels above 130 HU increases in medium and large density calcifications, which increases the Agatston score. In low density calcifications, in turn, the number of voxels above 130 HU decreases, which decreases the Agatston score.

As we know from the CONFIRM registry, small calcifications visually detected on CCTA scans in patients previously assigned to the zero Agatston score risk group, increased risk of major adverse cardiac events.[26] Therefore, detectability of small

13

calcifications plays a crucial role in further patient management. Importantly, when reduced tube currents were used, detectability of small calcifications decreased, especially for MBIR and DLR. Our hypothesis is that this can be explained by the need for increased noise suppression by these reconstruction algorithms. Therefore, based on these static phantom results we assume that patients might be misclassified into the zero Agatston score risk group when a reduced radiation dose protocol is used. Future patient studies with more small calcifications should verify this.

Additionally, at routine tube current level, the current study did not show relevant differences between reconstruction methods in terms of Agatston scores. However, when the tube current was decreased to 50%, Agatston score of low density calcifications acquired from the large dynamic phantom deviated from the standard measurement.[2] Therefore, as also underlined in SCCT guidelines, caution should be taken in terms of radiation dose reduction by decreasing tube current, especially in combination with iterative reconstruction methods.[15]

The patient study showed that only the Agatston score measured from MBIR differed from the reference Agatston score based on FBP. When considering patients with a zero Agatston score as defined by the reference method, MBIR classified one patient as a nonzero Agatston score, thereby increasing the risk classification. However, similar results were presented before, with 17% of cases reclassified into higher risk group, including 8% of patient misclassified as non-zero Agatston scores.[8] One explanation for this behavior might by the impact of the edge enhancement algorithm, whereby more pronounced CAC edges increase overall Agatston scores. Nevertheless, overall agreement in risk classification was excellent. Importantly, for our patient group, none of the patients was reclassified as a false negative. Currently, both American and European guidelines use CAC scoring as an additional tool not only for patient risk classification, but also for guiding statin and aspirin therapy.[27] Therefore, the lack of CAC measurement reproducibility and its dependency on different reconstruction methods, may affect patient management and outcome.[23] Based on patients results from our study and using FBP as reference, the most accurate calcium scoring was achieved when HIR or DLR was used, in terms of correct patient risk classification.

This study has several limitations. First, while our systematic analysis included both a static and dynamic phantom as well as a patient study, we only included a small number of patients. Moreover, only twelve patients (24%) presented with Agatston score between 0 and 10, which is the most susceptible group in terms of

13

calcium detectability. Despite the relatively small sample size of the patient study, we feel that the size was large enough for validation purposes of our phantom study. Second, we acquired data from one vendor. Therefore, a multivendor study analyzing the influence of different reconstruction methods on calcium detectability, quantification and risk stratification is certainly needed. Third, all patients were scanned with the standard protocol. Therefore, the effect of decreased radiation dose could not be evaluated in patients. Fourth, the D100 phantom is a static phantom. Thus, we were not able to acquire dynamic detectability phantom data. However, due to the decrease in detectability, even in a static situation, care should be taken when using non-FBP reconstructions for detecting CAC with this CT system.

In conclusion, based on our patient results, HIR and DLR reconstructed scans resulted in similar Agatston scores with excellent agreement and low-risk reclassification rate compared with routine reconstructed scans (FBP). These results suggest that these reconstruction methods might be applied for CAC scoring. However, based on our phantom study, caution should be taken when patients have Agatston scores between 0 and 10, as detectability of small calcifications varies with the used reconstruction algorithm, especially with MBIR and DLR. More clinical studies with a large amount of low Agatston score calcifications are needed to verify this. Moreover, decreased radiation dose impaired Agatston scoring of small calcifications which may lead to improper patient risk classification.
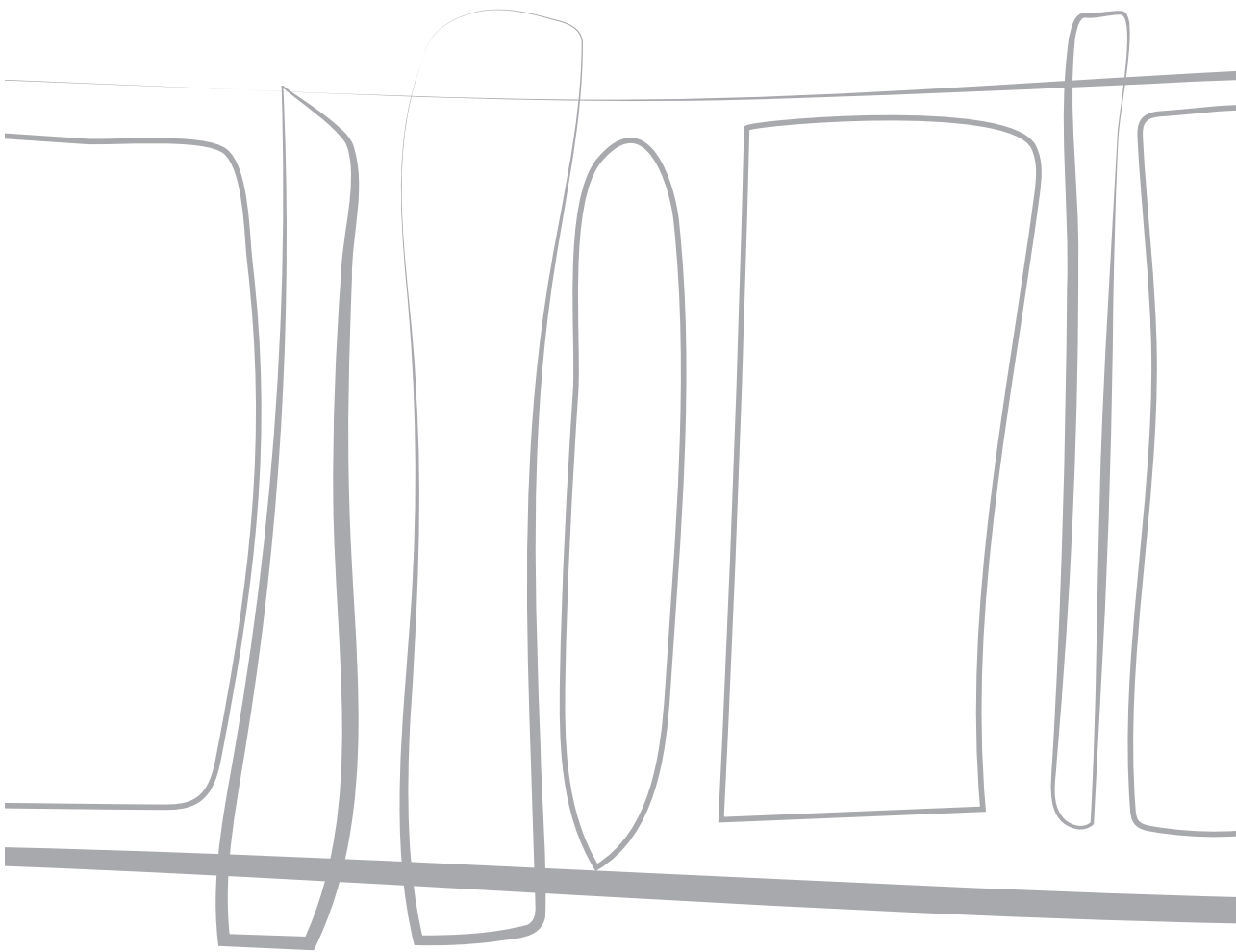
13

REFERENCES

1. Knuuti J, Wijns W, Achenbach S, et al. 2019 ESC guidelines for the diagnosis and management of chronic coronary syndromes. European Heart Journal. 2020;41(3):407-477. doi:10.1093/eurheartj/ehz425

2. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. Journal of the American College of Cardiology. 1990;15(4):827-832. doi:10.1016/0735-1097(90)90282-T

3. Blaha MJ, Cainzos-Achirica M, Greenland P, et al. Role of Coronary Artery Calcium Score of Zero and Other Negative Risk Markers for Cardiovascular Disease: The Multi-Ethnic Study of Atherosclerosis (MESA). Circulation. 2016;133(9):849-858. doi:10.1161/CIRCULATIONAHA.115.018524

4. Kurata A, Dharampal A, Dedic A, et al. Impact of iterative reconstruction on CT coronary calcium quantification. European Radiology. 2013;23(12):3246-3252. doi:10.1007/s00330-013-3022-8

5. Willemink MJ, Noël PB. The evolution of image reconstruction for CT-from filtered back projection to artificial intelligence. European radiology. 2019;29(5):2185-2195. doi:10.1007/S00330-018-5810-7

6. Xu F, Mueller K. Real-time 3D computed tomographic reconstruction using commodity graphics hardware. Physics in Medicine and Biology. Published online 2007. doi:10.1088/0031-9155/52/12/006

7. Willemink MJ, de Jong P a, Leiner T, et al. Iterative reconstruction techniques for computed tomography Part 1: Technical principles. European radiology. Published online January 2013. doi:10.1007/s00330-012-2765-y

8. Choi AD, Leifer ES, Yu JH, et al. Reduced radiation dose with model based iterative reconstruction coronary artery calcium scoring. European Journal of Radiology. 2019;111:1-5. doi:10.1016/j.ejrad.2018.12.010

9. Tatsugami F, Higaki T, Fukumoto W, et al. Radiation dose reduction for coronary artery calcium scoring at 320-detector CT with adaptive iterative dose reduction 3D. International Journal of Cardiovascular Imaging. 2015;31(5):1045-1052. doi:10.1007/s10554-015-0637-7

10. Tang YC, Liu YC, Hsu MY, Tsai HY, Chen CM. Adaptive Iterative Dose Reduction 3D Integrated with Automatic Tube Current Modulation for CT Coronary Artery Calcium Quantification: Comparison to Traditional Filtered Back Projection in an Anthropomorphic Phantom and Patients. Academic Radiology. Published online 2018. doi:10.1016/j.acra.2017.12.018

11. Akagi M, Nakamura Y, Higaki T, et al. Deep learning reconstruction improves image quality of abdominal ultra-high-resolution CT. European Radiology. Published online 2019. doi:10.1007/s00330-019-06170-3

12. Geyer LL, Schoepf UJ, Meinel FG, et al. State of the Art: Iterative CT Reconstruction Techniques. Radiology. 2015;276(2):339-357. doi:10.1148/RADIOL.2015132766

13. Higaki T, Nakamura Y, Zhou J, et al. Deep Learning Reconstruction at CT: Phantom Study of the Image Characteristics. Academic Radiology. Published online 2020. doi:10.1016/j.acra.2019.09.008

13

14. Hecht HS, Cronin P, Blaha MJ, et al. 2016 SCCT/STR guidelines for coronary artery calcium scoring of noncontrast noncardiac chest CT scans: A report of the Society of Cardiovascular Computed Tomography and Society of Thoracic Radiology. Journal of Cardiovascular Computed Tomography. 2017;11(1):74-84. doi:10.1016/j.jcct.2016.11.003

15. Hecht H, Blaha MJ, Berman DS, et al. Clinical indications for coronary artery calcium scoring in asymptomatic patients: Expert consensus statement from the Society of Cardiovascular Computed Tomography. Journal of Cardiovascular Computed Tomography. 2017;11(2):157-168. doi:10.1016/j.jcct.2017.02.010

16. McCollough CH, Ulzheimer S, Halliburton SS, Shanneik K, White RD, Kalender WA. Coronary Artery Calcium: A Multi-institutional, Multimanufacturer International Standard for Quantification at Cardiac CT. Radiology. 2007;243(2):527-538. doi:10.1148/radiol.2432050808

17. Groen JM, Kofoed KF, Zacho M, Vliegenthart R, Willems TP, Greuter MJW. Calcium score of small coronary calcifications on multidetector computed tomography: Results from a static phantom study. European Journal of Radiology. 2013;82(2):e58-e63. doi:10.1016/j.ejrad.2012.09.018

18. Husmann L, Leschka S, Desbiolles L, et al. Coronary Artery Motion and Cardiac Phases: Dependency on Heart Rate—Implications for CT Image Reconstruction. Radiology. 2007;245(2):567-576.

19. van der Werf NR, Willemink MJ, Willems TP, Vliegenthart R, Greuter MJW, Leiner T. Influence of heart rate on coronary calcium scores: a multi-manufacturer phantom study. International Journal of Cardiovascular Imaging. 2017;34(6):959-966. doi:10.1007/s10554-017-1293-x

20. van Praagh GD, van der Werf NR, Wang J, et al. Fully Automated Quantification Method (FQM) of Coronary Calcium in an Anthropomorphic Phantom. Medical Physics. Published online 2021:in press. doi:10.1002/mp.14912

21. Booij R, van der Werf NR, Budde RPJ, Bos D, van Straten M. Dose reduction for CT coronary calcium scoring with a calcium-aware image reconstruction technique: a phantom study. European Radiology. 2020;30(6):3346-3355. doi:10.1007/s00330-020-06709-9

22. Werf NR van der, Booij R, Schmidt B, et al. Evaluating a calcium-aware kernel for CT CAC scoring with varying surrounding materials and heart rates : a dynamic phantom study. European Radiology. 2021;31(12):9211-9220. doi:10.1007/s00330-021-08076-5

23. Detrano R, Guerci AD, Carr JJ, et al. Coronary calcium as a predictor of coronary events in four racial or ethnic groups. The New England journal of medicine. 2008;358(13):1336-1345. doi:10.1056/NEJMoa072100

24. Lee CH, Goo JM, Lee HJ, et al. Radiation dose modulation techniques in the multidetector CT era: from basics to practice. Radiographics : a review publication of the Radiological Society of North America, Inc. 2008;28(5):1451-1459. doi:10.1148/RG.285075075

25. Groen JM, Greuter MJ, Schmidt B, Suess C, Vliegenthart R, Oudkerk M. The Influence of Heart Rate, Slice Thickness, and Calcification Density on Calcium Scores Using 64-Slice Multidetector Computed Tomography: A Systematic Phantom Study. Investigative Radiology. 2007;42(12).

26. Han D, Klein E, Friedman J, et al. Prognostic significance of subtle coronary calcification in patients with zero coronary artery calcium score: From the CONFIRM registry. Atherosclerosis. 2020;309(June):33-38. doi:10.1016/j.atherosclerosis.2020.07.011

13

27. Osei AD, Mirbolouk M, Berman D, et al. Prognostic value of coronary artery calcium score, area, and density among individuals on statin therapy vs. non-users: The coronary artery calcium consortium. Atherosclerosis. 2021;316:79-83. doi:10.1016/J.ATHEROSCLEROSIS.2020.10.009

13

Online material is available via:
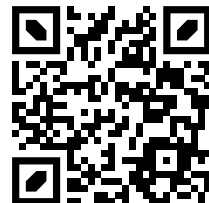
CHAPTER 14

Coronary artery calcium scoring: toward a new standard

Gijs D. van Praagh, MSc

Jia Wang, PhD

Niels R. van der Werf, MSc

Marcel J.W. Greuter, PhD

Domenico Matrodicasa, MD

Koen Nieman, MD PhD

Robbert W. van Hamersvelt, MD PhD

Luuk J. Oostveen, BSc

Frank de Lange, PhD

Riemer H.J.A. Slart, MD PhD

Tim Leiner, MD PhD

Dominik Fleischmann, MD

Martin J. Willemink, MD PhD

ABSTRACT

Objectives

Although the Agatston score is a commonly used quantification method, re-scan reproducibility is suboptimal, and different CT scanners result in different scores. In 2007, McCollough et al. proposed a standard for coronary artery calcium (CAC) quantification. Advancements in CT technology over the last decade, however, allow for improved acquisition and reconstruction methods. This study aims to investigate the feasibility of a reproducible reduced dose alternative of the standardized approach for CAC-quantification on state-of-the-art CT systems from four major vendors.

Materials and Methods

An anthropomorphic phantom containing nine calcifications and two extension rings were used. Images were acquired with four state-of-the-art CT systems using routine protocols and a variety of tube voltages (80-120 kV), tube currents (100 to 25% dose levels), slice thicknesses (3/2.5 and 1/1.25-mm), and reconstruction techniques (filtered back projection and iterative reconstruction (IR)). Every protocol was scanned five times after repositioning the phantom to assess reproducibility. Calcifications were quantified as Agatston scores.

Results

Reducing tube voltage to 100 kV, dose to 75%, and slice thickness to 1- or 1.25-mm combined with higher IR-levels resulted in an on average 36% lower intrascanner variability (IQR) compared with the standard 120 kV protocol. Interscanner variability per phantom size decreased by 34% on average. With the standard protocol, on average, 6.2±0.4 calcifications were detected, while 7.0±0.4 were detected with the proposed protocol. Pairwise comparisons of Agatston scores between scanners within the same phantom size demonstrated three significantly different comparisons at the standard protocol (P<0.05), whereas no significantly different comparisons arose at the proposed protocol (P>0.05).

Conclusions

On state-of-the-art CT systems of four different vendors, a 25% reduced dose, thin-slice calcium scoring protocol led to improved intra- and interscanner

reproducibility and increased detectability of small and low-density calcifications in this phantom. The protocol should be extensively validated before clinical use, but it could potentially improve clinical interscanner/interinstitutional reproducibility and enable more consistent risk assessment and treatment strategies.

14

INTRODUCTION

Quantification of coronary artery calcifications (CAC) with the Agatston score using cardiac CT is a strong predictive marker for future cardiovascular events in asymptomatic individuals at low-to-intermediate risk.[1,2] Even though CT-technology has improved tremendously over the last decade, the CAC-quantification standard – using 120 kV, 3- or 2.5-mm slices, and filtered back projection (FBP) – has not been renewed since its introduction in 2007 by McCollough et al.[3] Larger CT detector coverage, more efficient detector technology, higher spatial resolution, shorter gantry rotation time, increased X-ray tube power, and improved reconstruction algorithms are nowadays available to reduce the image noise of cardiac CT and thus allow for radiation dose reduction.[4] A large body of evidence has shown that the reproducibility of currently used CAC acquisition and quantification methods is limited as it can result in reclassifications.[4-7] Hence, the current quantification standard for CAC scoring should be updated.

Rutten et al. demonstrated in patients that small variations of scan starting position affect the reproducibility of Agatston scores with a potential risk of reclassification in 9% of the individuals.[8] In addition, Willemink et al. found a substantial inter-vendor variability of the Agatston score, which resulted in a modest cardiovascular risk reclassification in up to 6.5% of ex-vivo hearts.[5] CAC reproducibility can be improved by reducing slice thickness and tube voltage.[4,5,9] The current, standard slice thickness of 3- or 2.5-mm results in a suboptimal spatial resolution and a possible underestimation of small or less dense calcifications due to partial volume effects. The use of thinner slices reduces these effects, improves reproducibility, and enhances detectability of (micro)calcifications, where one microcalcification is typically 0.5-15 mm.[4,9,10] This may be an opportunity for risk factor modification and thus prevention of cardiovascular events, as non-zero CAC scores are associated with higher risk for future events, while zero scores are strong negative predictors.[4,9-11] Reducing tube voltage results in two effects: it will reduce radiation dose when other settings remain unchanged; concurrently it will increase CT attenuation, which can potentially improve the detectability and quantification of less dense calcifications. Iterative reconstruction (IR) can be used to reduce radiation dose without compromising image noise compared to routine dose FBP.[13,14] Multiple studies have shown that radiation dose reduction over 50% is possible in CAC-imaging with IR.[13] Modifying the CAC acquisition method will result in more accurate estimation of coronary calcified lesions as well as reduced radiation, hence important for consistent risk assessment and treatment.

This study aims to investigate the feasibility and reproducibility of a reduced dose alternative of the standardized approach for CAC-quantification on state-of-the-art CT systems from four major vendors.

MATERIALS AND METHODS

Phantom

For all experiments, a commercially available static anthropomorphic thoracic phantom (Thorax, QRM, Möhrendorf, Germany) was used. At the position of the heart, a cardiac calcification insert (CCI, QRM, Möhrendorf, Germany) was placed, containing nine small cylindrical calcifications with different sizes and densities (diameters: 1, 3, and 5 mm; densities: 200, 400, and 800 mg/cm³ calcium hydroxyapatite (CaHA)), and two large calibration rods (one water-equivalent and one 200 mg/cm³ calcium hydroxyapatite). The thoracic phantom simulates a small-sized patient (anterior-posterior x lateral: 200x300 mm). Two additional tissue-equivalent oval rings (Extension rings, QRM, Möhrendorf, Germany) were used to simulate medium- and large-sized patients (250x350 mm and 300x400 mm, respectively; *Figure 1*).

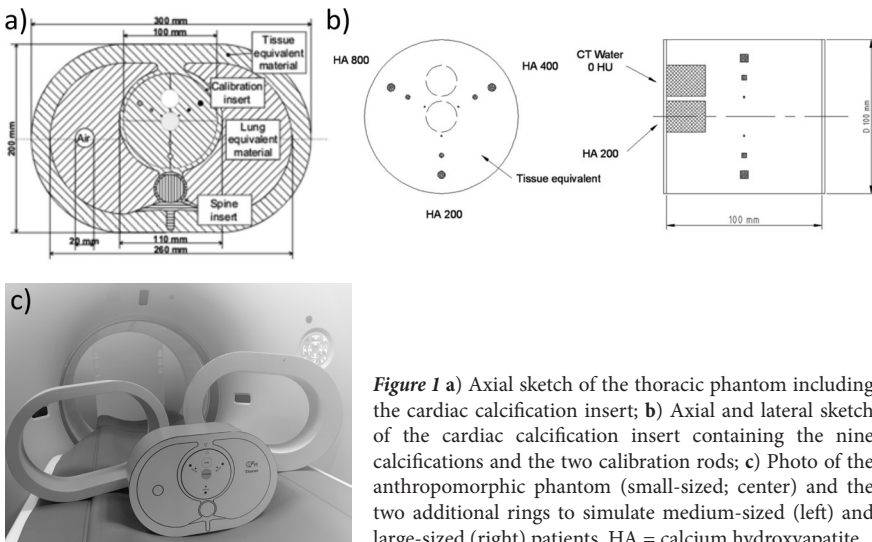*Figure 1 a*) Axial sketch of the thoracic phantom including the cardiac calcification insert; **b**) Axial and lateral sketch of the cardiac calcification insert containing the nine calcifications and the two calibration rods; **c**) Photo of the anthropomorphic phantom (small-sized; center) and the two additional rings to simulate medium-sized (left) and large-sized (right) patients. HA = calcium hydroxyapatite.

Image Acquisition and Reconstruction

All phantom sizes were scanned using state-of-the-art CT systems from four different vendors (scanner A: SOMATOM Force, Siemens Healthineers, Erlangen, Germany; scanner B: Revolution, GE Healthcare, Milwaukee, Wis, USA; scanner C: iCT, Philips Healthcare, Best, the Netherlands; scanner D: Aquilion One PRISM Edition, Canon Medical Systems, Otawara, Japan). Acquisition and reconstruction parameters are listed in *Tables 1* and *2*. A reference tube voltage of 120 kV and a decreased tube voltage of 100 kV were used. An exploratory scan session on each CT system showed that tube power limitations arose for 80 kV, therefore, 100 kV was the most optimal reduced tube potential for all four vendors (supplementary *Table 1*). Tube current was modulated with anatomic based automatic exposure control on scanners A, B, and D. For scanner C, this was not available. Volumetric CT dose indices ($CTDI_{vol}$) were kept similar between CT systems. Reference dose-levels (100%) were chosen as $CTDI_{vol}$ values of 1.5, 3.3, and 7.0 mGy for small, medium, and large phantom size, respectively, which is in line with clinically used protocols.[15-17] The following parameters were varied on the scanners: radiation dose, slice thickness, and IR-level (*Table 2*). Radiation dose was incrementally reduced by lowering $CTDI_{vol}$, resulting in four dose levels: 100%, 75%, 50%, and 25%. Multiple reconstructions were made per scan with varying slice thickness and increment: 3.0 and 1.0 mm on scanners A, C, and D, and 2.5 and 1.25 mm on scanner B. Besides FBP, three settings of IR algorithms were used on every scanner. Every protocol was scanned five times with small repositioning (approximately 5 mm translational and 2 degrees rotational) of the phantom after each scan to measure and correct for interscan variability.[18]

Quantification of Coronary Calcifications

The calcifications inside the phantom were quantified as Agatston scores using the validated fully automated quantification method (FQM).[19] The standard 130 Hounsfield unit (HU) threshold was used for 120 kV scans. Due to increased x-ray beam attenuation, the threshold was increased to 147 HU for 100 kV scans as proposed by Nakazato et al., and effects in similar Agatston results with decreased tube voltage, decreased slice thickness, and use of IR shown by Hou et al.[20,21] Additionally, the output of FQM gave the following information about each individual calcification: volume, mass, maximum area, and mean Agatston weightfactor (based on maximum HU value per area).[22] Volume was quantified according to Callister et al.[23] Volume, area, and weightfactor were only used for clarification purposes of the changes in Agatston scores due to acquisition or reconstruction adjustments.

Table 1 Computed tomography acquisition and reconstruction parameters

		Scanner A	**Scanner B**	**Scanner C**	**Scanner D**
Acquisition mode		**Axial**	**Axial**	**Axial**	**Axial**
ECG-triggering		Prospective	Prospective	Prospective	Prospective
ECG-phase (% of R-R interval)		70%	75%	78%	75%
Simulated heart rate (bpm)		60	60	60	60
Rotation time (s)*		0.25	0.28	0.27	0.275
Detector collimation (mm)		160 x 0.6	140 x 0.625	128 x 0.625	200 x 0.5
Field of view (mm)		200 x 200	200 x 200	200 x 200	200 x 200
Matrix size (pixels)		512 x 512	512 x 512	512 x 512	512 x 512
Reconstruction kernel		Qr36d	Standard	XCA	FC12
Tube current modulation		CareDose 4D	SmartmA	-	Sure Exposure
IR algorithms		ADMIRE (1-5)	ASiR-V (0-100%)	iDose[4] (1-7)	AIDR 3D (mild-strong)
Reconstruction settings	Routine**	2	50%	FBP	FBP
	Low	2	30%	1	Mild
	Intermediate	3	50%	3	Standard
	High	4	70%	5	Strong

ECG = electrocardiogram; bpm = beats per minute; IR = iterative reconstruction; FBP = filtered back projection. ADMIRE = Advanced Modeled Iterative Reconstruction, Siemens Healthineers, Erlangen, Germany; ASiR-V = Adaptive Statistical Iterative Reconstruction-V, GE Healthcare, Milwaukee, Wis, USA; iDose[4], Philips Healthcare, Best, the Netherlands; AIDR 3D = Adaptive Iterative Dose Reduction 3D, Canon Medical Systems, Otawara, Japan
* Scanner B: to reach the highest possible dose with 100 kV for the large-sized phantom, the rotation time was prolonged to 0.35 s.
**Reconstruction settings of the standard protocol were based on the institution of the CT system to simulate clinical situations.

Table 2 Variable computed tomography acquisition and reconstruction parameters. Five repetitions were done for every protocol with small repositioning of the phantom on the table. The standard protocol was used for every session, repetition, and phantom size.

	kV	**Dose***	**Slice thickness / increment (mm)**	**Reconstruction level**
Standard protocol**	120		3 / 3 or 2.5 / 2.5	Routine
Variable parameters	100	100%, 75%, 50%, and 25%***	3 / 3 or 2.5 / 2.5	Routine
			1 / 1 or 1.25 / 1.25	FBP, low, intermediate, and high

IR = iterative reconstruction; FBP = filtered back projection.
* Full dose level (100%) is based on clinically used protocols with CTDI$_{vol}$ value of 1.5, 3.3, and 7.0 mGy for the small, medium and large phantom, respectively.
**Standard settings were based on the institution of the CT system to simulate clinical situations.
*** Scanner B: At the small phantom, 40% was the lowest possible dose-level; at the large phantom, 89% was the highest possible dose-level when prolonging rotation time to 0.35 s. Scanner C: At the small phantom, 33% was the lowest possible dose-level.

14

Image Noise

The output of FQM included the number of detectable calcifications and noise-level. Noise-level was defined as the standard deviation of pixel values in HU within a circular region of interest of 1.5 cm^2 placed in the center slice of the water-equivalent rod. McCollough et al. described a noise target measured in the water-equivalent rod of 20-HU for small- and medium-sized phantoms and 23-HU for the large-sized phantom. Noise-levels below these targets were not selected to keep radiation dose-levels reasonable, thus not too high.[3] After review of the experimental data, we chose an upper threshold as well of 30 HU for the small- and medium-sized phantom and 35 HU for the large-sized phantom to keep image noise reasonable and reduce the possibility of false positives.

Statistical Analysis

The interquartile range (IQR) of the total Agatston scores from all phantom sizes and all CT systems scanned with the standard protocol were used as a reference range for total Agatston scores to keep scores similar to the current protocol. We compared the different scan protocols based on intrascanner and interscanner reproducibility, calcification detectability, and image noise as described above. Intrascanner reproducibility was defined as the calcium score IQR of the five repetitions within that protocol. Interscanner reproducibility was defined as IQR of the sum of all calcium scores from that protocol acquired on all four CT systems. Change of variability was thus calculated as:

$$\text{Variability change} = \left(-1 + \frac{\text{IQR}_{\text{new protocol}}}{\text{IQR}_{\text{current protocol}}} \right) \cdot 100\%$$

The FQM used a mask based on the physical locations of all calcifications in the phantom. A calcification was defined detected when connected components above the 130 or 147 HU threshold arose within the mask. On that basis, a protocol for each scanner was proposed and compared with the currently used standard protocol. To show trends, results in the text are shown in median (IQR) for all phantom sizes and scanners combined unless indicated otherwise. More detailed comparisons between scores are displayed in the figures. Only pairwise comparisons were done for the total Agatston scores of the standard and proposed protocol. After testing for normality, differences were evaluated with Friedman's tests and post-hoc Dunn's tests with Bonferroni correction. P-values <0.05 were considered statistically significant. SPSS version 25 (IBM Corp., Armonk, NY, USA) was used for statistical analyses.

RESULTS

Intrascanner reproducibility

Agatston score variability of repeated acquisitions on the same scanner slightly increased after decreasing the tube potential (see *Figure 2* and *supplementary Figures 1* and *2*), i.e., IQR of 100 kV, 100% dose, thick slices, and routine reconstruction protocol changed with 7% (-26% to 140%). Reducing slice thickness in addition to the decreased tube potential lowered intrascanner variability with IQR changes of -54% (-76% to -44%). Comparing this 100 kV, thin-slice, 100% dose protocol with tube current-reductions to 75%, 50%, and 25% resulted in Agatston score IQR changes of 22% (-39% to 141%), 0 (-45% to 86%), and 99% (40% to 519%), respectively. Similar trends were seen in volume scores (see *supplementary results*).

Calcification detectability

Table 3 shows the number of detected calcifications per protocol and phantom size of all scanners combined. Out of the nine calcification inserts, the number of detected calcifications with the standard protocol was on average 6.2±0.4. This was similar with standard slice thickness, 100 kV, 100% dose, and routine reconstruction. Additional reduction of the slice thickness increased the number of detected calcifications with on average one calcification. Protocols with 100 kV, thin sections, routine reconstruction, and reduced radiation dose resulted in a similar number of detected calcifications. Increasing IR-levels showed a trend towards a slight decrease in detected number of calcifications. The higher number of calcifications detected was due to small calcifications of 800 or 400 mg/cm^3 CaHA.

Radiation dose and image noise

Noise-levels differed between scanners, but noise-levels of the standard protocols were equal or below the lower noise limit on all scanners (see *Figure 3* and *supplementary Figures 3* and *4*). Thin-sliced protocols at 100 kV with 50 or 25% dose resulted in noise-levels exceeding the upper threshold on scanner C, even with highest IR-level for medium- and large-sized phantoms, respectively. Thin-sliced protocols with 75% dose stayed within the noise thresholds when intermediate or high IR was used for all scanners, except scanner C. Noise-levels exceeded the upper threshold at scanner C when 100 kV, 75% dose, thin slices, and intermediate IR was used at the medium- and large-sized phantom, respectively.

14
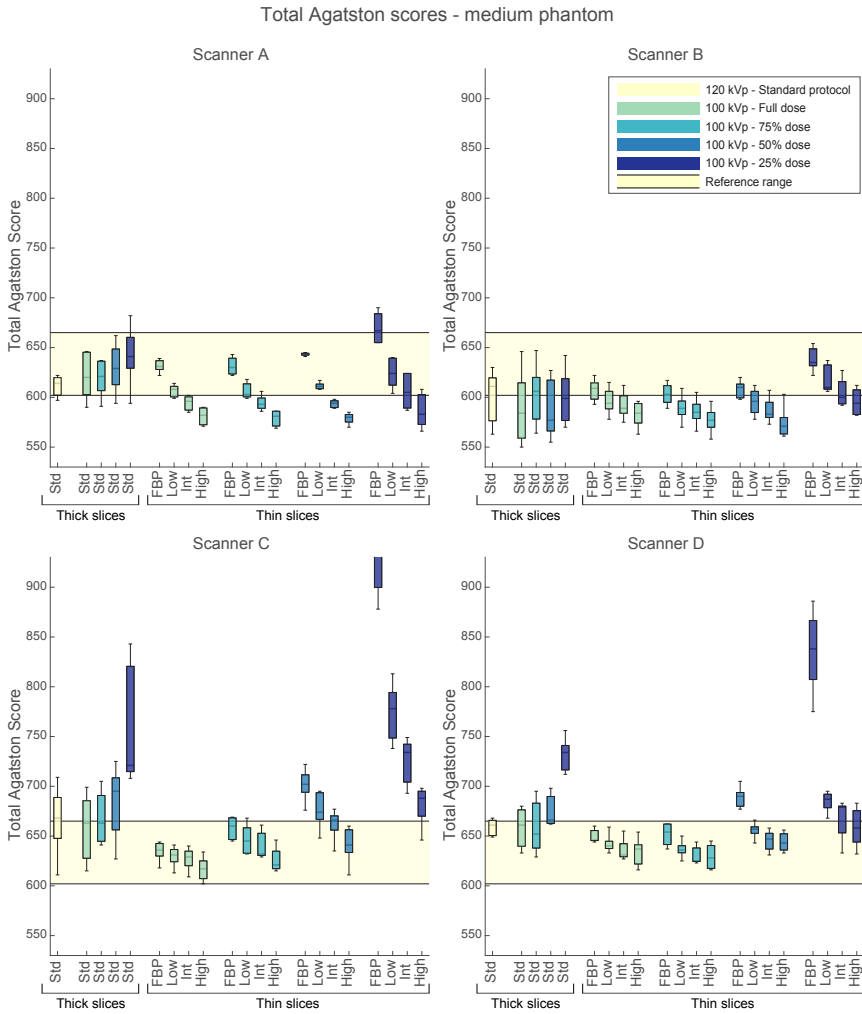
Figure 2 Total Agatston scores of all scanners and protocols scanned with the medium phantom. Reconstruction technique/IR-levels and slice thicknesses are given on the x-axis (thick = 3 or 2.5 mm; thin = 1 or 1.25 mm). Results are grouped in dose-levels. The light-yellow range is the reference range, which is the IQR of all Agatston scores from all scanners and all phantom sizes scanned with the standard protocol. Routine reconstruction (used at thick slices) for scanner A is corresponding to low IR-level; for scanner B corresponding to intermediate IR-level; for scanner C and D corresponding to FBP. HU thresholds were 130 HU for 120 kV scans and 147 HU for 100 kV scans. IR = iterative reconstruction; IQR = interquartile range; FBP = filtered back projection; HU = Hounsfield units. Scanner A: SOMATOM Force, Siemens Healthineers; Scanner B: Revolution, GE Healthcare; Scanner C: iCT, Philips Healthcare; Scanner D: Aquilion One PRISM Edition, Canon Medical Systems

14

Table 3 Number of detected calcifications per protocol. Numbers are presented in average ± standard deviation of all scanners combined

kV	Dose	Slice thickness (mm)	Reconstruction settings	Number of calcifications		
				Small	Medium	Large
120	100%	3.0 / 2.5	Std	6.4±0.5	6.2±0.4	6.1±0.2
	100%			6.2±0.4	6.1±0.3	6.1±0.2
	75%	3.0 / 2.5	Std	6.1±0.3	6.2±0.4	6.1±0.3
	50%			6.1±0.3	6.2±0.4	6.3±0.4
	25%			6.3±0.5	6.3±0.5	6.5±0.8
			FBP	7.4±0.5	7.4±0.5	7.4±0.7
	100%	1.0 / 1.25	Low	7.3±0.4	7.2±0.4	7.0±0.6
			Int	7.3±0.4	7.2±0.4	6.9±0.5
			High	7.1±0.3	7.1±0.3	6.8±0.5
			FBP	7.6±0.5	7.5±0.6	7.4±0.6
100	75%	1.0 / 1.25	Low	7.2±0.4	7.4±0.6	7.2±0.6
			Int	7.1±0.3	7.2±0.5	7.0±0.5
			High	7.1±0.2	7.0±0.2	6.9±0.6
			FBP	7.5±0.5	7.5±0.6	7.6±0.6
	50%	1.0 / 1.25	Low	7.2±0.5	7.2±0.6	7.4±0.6
			Int	7.2±0.5	7.0±0.5	7.3±0.7
			High	7.0±0.5	6.8±0.4	6.9±0.7
			FBP	7.5±0.6	7.5±0.8	7.1±0.8
	25%	1.0 / 1.25	Low	7.3±0.6	7.1±0.8	6.8±0.8
			Int	6.9±0.6	7.0±0.8	6.8±0.7
			High	6.8±0.5	7.0±0.8	6.8±0.7

Proposed protocol

Based on the above results, we proposed the following protocols: for scanners A and B: 100 kV, 75% dose, thin slices, intermediate IR; for scanners C and D: 100 kV, 75% dose, thin slices, high IR (*Table 4*). *Figure 4* shows axial images of the phantom obtained with the standard and proposed protocol on all scanners.

Compared with the standard protocol, these protocols resulted in an improved intrascanner variability (IQR) of -12%, -68%, and -73% for scanners A, B, and C, respectively, and a slight increase of 8% for scanner D. Also, improved intrascanner volume-variability was found for all scanners and improved mass-variability for scanners C and D (see *supplementary results*).

14
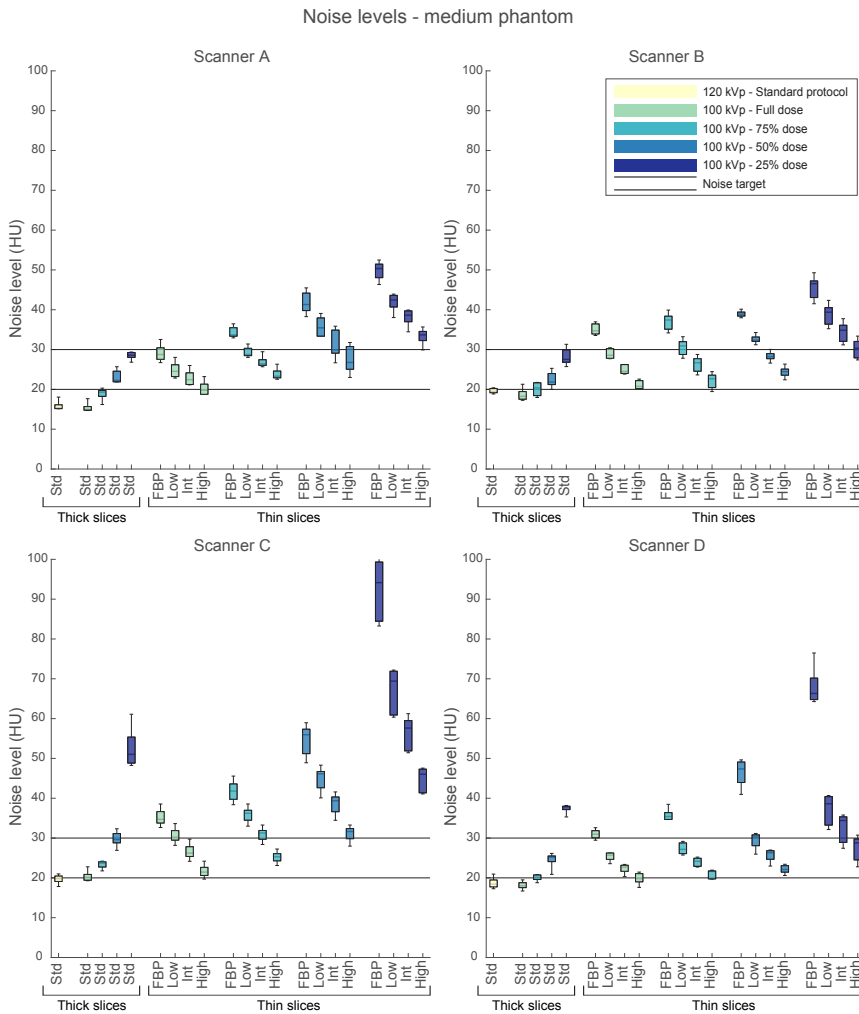
Figure 3 Noise-levels of all scanners and protocols scanned with the medium phantom. IR-levels and slice thicknesses are given on the x-axis (thick = 3 or 2.5 mm; thin = 1 or 1.25 mm). Results are grouped in dose-levels. Noise-levels are calculated in a circular ROI of 1.5 cm² in the center slice of a water-equivalent rod as standard deviation of CT values. The two continuous lines show the target range. IR = iterative reconstruction; ROI = region of interest.

With these proposed protocols interscanner variability changed with -55%, -35%, and -13% for small-, medium-, and large-sized phantoms, respectively. Interscanner variability improved for all volumes and mass scores as well, except for the mass scores in the large-sized phantom (see *supplementary results*).

Table 4 The proposed acquisition and reconstruction protocol per scanner

Scanner	kV All scanners	CTDI$_{vol}$ All scanners	Slice thickness / increment (mm) A, C, and D	B	Iterative reconstruction algorithm and level A	B	C	D
Proposed protocol	100	1.1 / 2.5 / 5.3*	1 / 1	1.25 / 1.25	ADMIRE 3	ASiR-V 50%	iDose 5	AIDR-3D strong

CTDI$_{vol}$ = volumetric computed tomography dose index; ADMIRE = Advanced Modeled Iterative Reconstruction; ASiR-V = Adaptive Statistical Iterative Reconstruction-V; AIDR 3D = Adaptive Iterative Dose Reduction 3D

*CTDI$_{vol}$ values are given for small-, medium-, and large-sized patients, respectively
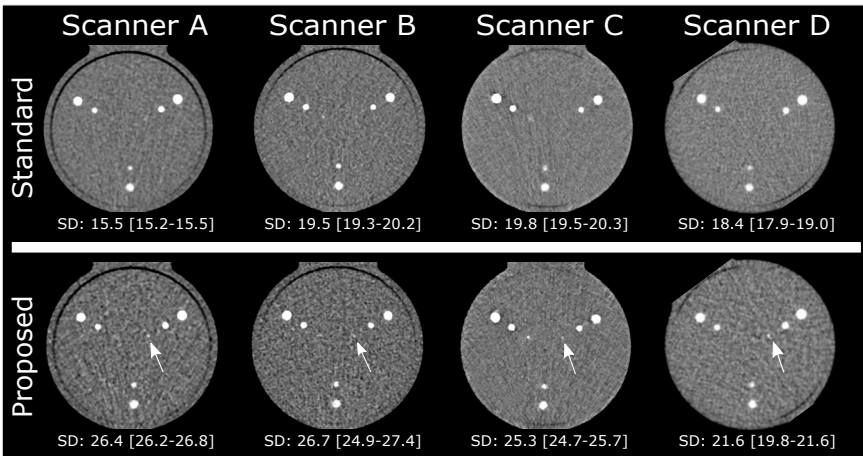
Figure 4 Axial views of the center slice of the calcifications in the CCI in the medium phantom. Upper row shows standard protocol for all scanners; lower row shows the proposed protocol for all scanners. Below every screenshot, the median [IQR] noise-level (standard deviation (SD) of HU in the water-equivalent rod) is given. White arrows show the calcifications that were detected in the proposed protocol, whereas they were not detected in the standard protocol at that scanner. CCI = cardiac calcification insert; IQR = interquartile range; HU = Hounsfield units.

Pairwise comparisons of Agatston scores between scanners within the same phantom size showed that the standard protocol resulted in significantly different scores between scanners: scanner B vs. C for the small phantom (558 [556-574] vs. 702 [661-730], p<0.01), scanner B vs. C for the large phantom (588 [567-603] vs. 653 [635-685], p=0.02), and scanner B vs. D for the small phantom (558 [556-574] vs. 720 [703-733], p=0.02). All other comparisons were not significantly different

(p>0.05). In contrast, Agatston scores were not significantly different between scanners when using the proposed new protocols (p>0.05).

Noise-levels for small-, medium-, and large-sized phantoms with the proposed protocols ranged from 23.7 (23.3-24.2) to 27.9 (27.5-29.3) for scanner A; 26.1 (26.0-27.1) to 29.1 (28.3-29.4) for scanner B; 21.3 (21.1-21.6) to 35.1 (33.5-36.9) for scanner C; and 20.8 (20.4-22.4) to 23.3 (21.9-24.5) for scanner D.

The number of detectable calcifications with the proposed protocols was 7.1±0.2, 7.1±0.4, and 7.0±0.5 for the small-, medium-, and large-sized phantoms, respectively.

Per-calcification analysis

Median Agatston scores of 800 mg/cm^3 CaHA calcifications changed with the proposed protocol by -15% and -14% for the large- and medium-sized calcifications, respectively (see *Figure 5* and *supplementary Figures 5* and *6*). Median Agatston scores slightly decreased on average for the large-sized 400 mg/cm^3 CaHA calcification (-2%), whereas scores increased for the medium-sized 400 mg/cm^3 CaHA calcification (11%). Median Agatston scores of 200 mg/cm^3 CaHA calcifications increased on average with 39% and 43% for the large- and medium-sized calcifications, respectively.

The proposed protocol resulted in calcification volumes closer to the physical volume for all calcifications (see *supplementary results* and *supplementary Figure 7*). Mass scores remained similar compared with the standard protocol (*supplementary Figure 8*). Maximum areas and weightfactors remained similar or slightly increased (*supplementary Figures 9* and *10*). Weightfactors especially increased in 200 mg/cm^3 CaHA calcifications.

Intrascanner variability of the Agatston score decreased for most calcifications. Medians of IQR-changes per-calcification were -68%, 5%, and -30% for large-sized 800, 400, and 200 mg/cm^3 CaHA calcifications, and -81%, -37%, and 39% for medium-sized 800, 400, and 200 mg/cm^3 CaHA calcifications, respectively.

Furthermore, interscanner variability of Agatston scores per-calcification decreased, resulting in IQR changes of -63%, -29%, and -34% for large-sized 800, 400, and 200 mg/cm^3 CaHA calcifications, and -60%, -44%, and -8% for medium-sized 800, 400, and 200 mg/cm^3 CaHA calcifications, respectively (*Figure 5*).

14
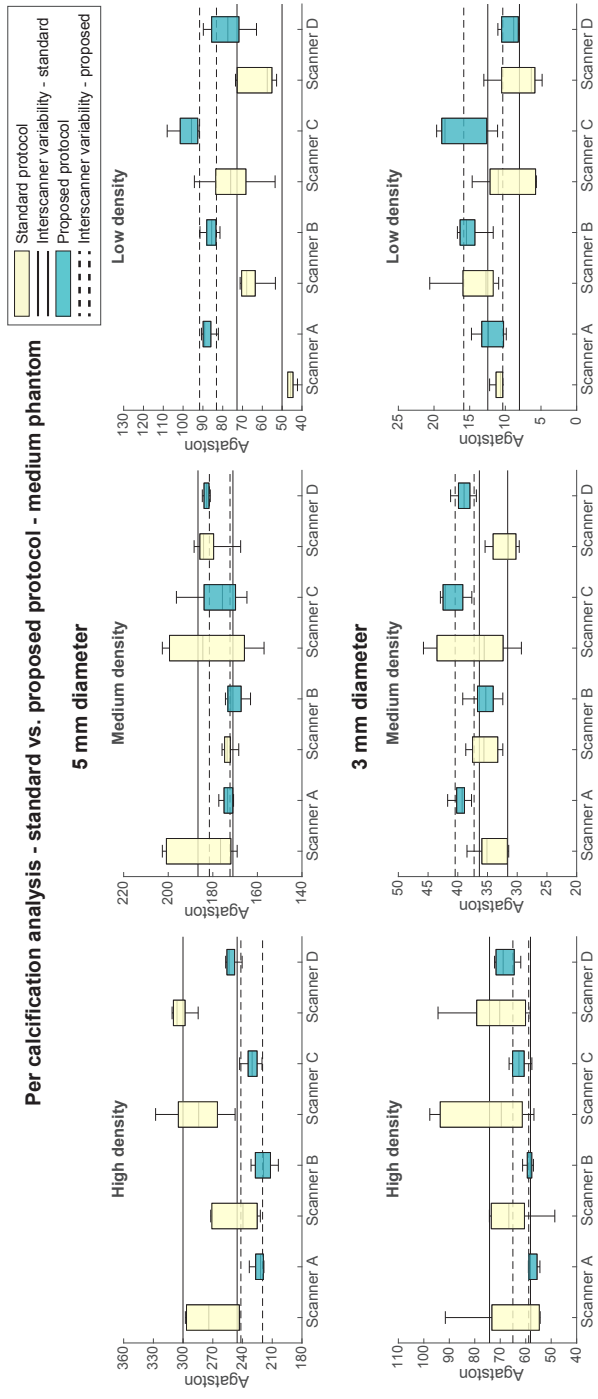
Figure 5 Per-calcification analysis. The Agatston scores per-calcification of the routinely used standard protocol and the proposed protocol at the medium phantom. The continuous lines show the IQR of the Agatston scores for the standard protocol. The dashed lines show the IQR of the Agatston scores for the proposed protocol. The 1 mm diameter calcifications are not shown, as they are not always detectable with every protocol. IQR = interquartile range

14

DISCUSSION

Our multivendor phantom study showed that an updated reduced dose, thin-slice CT acquisition protocol for CAC scoring resulted in improved intra- and interscanner variability and detectability of small and low-density calcifications. The proposed protocol decreased volume-variability of calcifications and thus decreased influence of the partial volume effect, resulting in volumes closer to their physical volume. Also, one additional small calcification was detected and an increase in area and weightfactor, especially for low-density calcifications, showed improved visibility. We aimed to update the 2007 CAC-quantification standard by evaluating multiple protocols with the same phantom setup. For four state-of-the-art CT systems, we established an updated CAC-quantification protocol by combining lower tube voltage (100 kV), reduced radiation dose (CTDI$_{vol}$ 1.1, 2.5, and 5.3 mGy for small-, medium-, and large-sized patients, respectively), thinner slices (1- or 1.25-mm) and higher IR-levels (ADMIRE 3, ASiR-V 50%, iDose[4] 5, or AIDR-3D strong).

Multiple important factors call for an update of the current standard.[4,11] First, in the last decade, CT technology has remarkably improved.[4] Second, recent research from Blaha et al. has shown that the shape and distribution of CAC are important contributors to cardiovascular risk stratification.[24] Third, Criqui et al. found an inversely proportional association between CAC density and future cardiovascular events, therefore making it important to accurately quantify low-density calcifications.[25] Fourth, there is a substantial intra- and interscanner variability with the current CAC-quantification protocols, possibly causing different treatment approaches for the same patient on different scanners or in different hospitals.[5,8] Last, Han et al. recently demonstrated that a small but non-negligible number of patients with zero CAC score actually did have coronary artery calcifications, missed by the current protocol.[12] The presence of these plaques was associated with higher risk for major adverse cardiovascular events.

Multiple phantom and patient studies have been conducted to evaluate radiation dose reduction for CAC CT scans by lowering kV and/or mAs with or without IR.[13,15,26–28] Also, multiple studies evaluated the effect of slice thickness on CAC scores.[9,18,27-30] No studies have tested a variation of all four parameters (tube voltage, tube current, slice thickness, and reconstruction technique) on current state-of-the-art CT systems of multiple vendors. Vonder et al. varied tube voltages, dose-levels, and ADMIRE-levels with a dedicated phantom containing 100 small calcifications on a single CT system with conventional 3-mm slice thickness.[33]

14

They found possible dose-reductions of up to 60.6% with 100 kV and IR-1. The lower tube current and IR-level compared with our study can be explained by the 3-mm slices, as thinner slices result in more noise. Groen et al. varied tube voltage and slice thickness on one CT scanner with the same phantom.[34] They found an optimal protocol of 100 kV and 3-mm slices that resulted in similar Agatston scores and CAC detectability compared to traditional electron beam tomography, thus, not intending to improve detectability or reproducibility.

Mantini et al. showed in patients that thinner slice reconstructions led to significant upward risk reclassifications due to higher CAC scores. Two possible explanations were given: 1) increased detection efficiency of small calcifications; 2) increased noise-level resulting in more false positives.[29] Our results support the first explanation. Although the excess risk of non-zero CAC scores by microcalcifications is not yet precisely known, the study of Han et al. suggests that these calcifications missed by the current protocol should not be neglected.[12,35,36] We expect that the updated protocol may result in improved identification of these false zero scores as detection of smaller calcifications and low-density calcifications is expected to improve. This is confirmed by our study since on average we found one additional small calcification with the proposed protocol compared with the current protocol. The most effective way to assess how this would translate to the clinic would be a study where patients are scanned with both the current protocol and the proposed protocol. Furthermore, we showed that IR minimizes increased noise-levels of thinner sections.

A large body of evidence shows that it is safe to implement reduced radiation dose by lower tube potential and/or current, and increased IR-levels.[13,27,37] Also, the positive effects of thinner slices in CAC acquisition has been thoroughly investigated and coincide with our findings: decreased influence of the partial volume effect, improved reproducibility of CAC scores, and increased detectability of small calcifications.[9,24,27-30] However, the effect of combining all four parameters in scanning patients should be further investigated.

Earlier studies have used the increased threshold before or a similar threshold, which resulted in comparable Agatston scores between the current 120 kVp protocol and a 100 kVp, lower dose protocol with 3 mm slice thickness, or with 0.5 mm slice thickness and iterative reconstruction.[20,21,38,39] Therefore, we decided to use the same threshold and also found similar Agatston scores to the current protocol.

14

The supplementary section provides more specific results of volume, mass, area, and weightfactor per calcification. Similar results between volumes and Agatston scores were seen, indicating that the decreased intrascanner variability is due to the decreased partial volume effects. Also, volumes of all calcifications came closer to the physical volume with the proposed protocol compared with the current standard protocol, thus resulting in a more accurate representation of the calcifications. Besides that, an increase in area and weightfactor is seen in especially low-density calcifications, resulting in a maximum area closer to the physical area for small calcifications and indicating better detectability of these calcifications. This is likely due to the decreased tube voltage, hence increased attenuation.

Our study has limitations. A static phantom was used; hence the effect of motion was not addressed. Cardiac motion can cause artefacts, which have a non-negligible effect on CAC scores. Van der Werf et al. showed in a dynamic phantom that heart rates have a substantial effect on Agatston and mass scores.[40] Future research should investigate the magnitude of this potential problem with our proposed protocol. Also, the CCI only contains nine calcifications of three different sizes and densities. To investigate the exact improved detectability of our protocol, another dedicated phantom with more calcifications should be used. While we expect the reclassification rates to be lower with the updated protocol, clinical studies should be performed to investigate whether reclassifications still occur due to the proposed protocol. Last, this protocol should be tested with new emerging techniques like dedicated kV-independent kernels.[41]

In conclusion, current CT acquisition protocols for CAC-quantification may be updated to a protocol with 100 kV, 75% radiation dose, 1- or 1.25-mm slice thickness, and higher IR-levels. On state-of-the-art CT systems of four different vendors, this protocol led to improved intra- and interscanner reproducibility and increased detectability of small and low-density calcifications. It is important to emphasize that before clinical use, the viability of this protocol should be validated in dynamic phantom and clinical studies. However, due to the improved reproducibility of the Agatston score with this protocol, this could potentially improve clinical interscanner/interinstitutional reproducibility, which would result in more consistent risk assessment and treatment strategies; and it may potentially facilitate in better risk stratification and an improved scoring method.

14

REFERENCES

1. Divakaran S, Cheezum MK, Hulten EA, et al. Use of cardiac CT and calcium scoring for detecting coronary plaque: Implications on prognosis and patient management. Br J Radiol. 2015;88(1046). doi:10.1259/bjr.20140594

2. Hecht H, Blaha MJ, Berman DS, et al. Clinical indications for coronary artery calcium scoring in asymptomatic patients: Expert consensus statement from the Society of Cardiovascular Computed Tomography. J Cardiovasc Comput Tomogr. 2017;11(2):157-168. doi:10.1016/j.jcct.2017.02.010

3. McCollough CH, Ulzheimer S, Halliburton SS, Shanneik K, White RD, Kalender WA. Coronary Artery Calcium: A Multi-institutional, Multimanufacturer International Standard for Quantification at Cardiac CT. Radiology. 2007;243(2):527-538. doi:10.1148/radiol.2432050808

4. Willemink MJ, van der Werf NR, Nieman K, Greuter MJW, Koweek LM, Fleischmann D. Coronary artery calcium: A technical argument for a new scoring method. J Cardiovasc Comput Tomogr. 2019;13(6):347-352. doi:10.1016/j.jcct.2018.10.014

5. Willemink MJ, Vliegenthart R, Takx RAP, et al. Coronary Artery Calcification Scoring with State-of-the-Art CT Scanners from Different Vendors Has Substantial Effect on Risk Classification. Radiology. 2014;273(3):695-702. doi:10.1148/radiol.14140066

6. Greuter MJW, Groen JM, Nicolai LJ, Dijkstra H, Oudkerk M. A model for quantitative correction of coronary calcium scores on multidetector, dual source, and electron beam computed tomography for influences of linear motion, calcification density, and temporal resolution: A cardiac phantom study. Med Phys. 2009;36(11):5079-5088. doi:10.1118/1.3213536

7. Ulzheimer S, Kalender WA. Assessment of calcium scoring performance in cardiac computed tomography. Eur Radiol. 2003;13(3):484-497. doi:10.1007/s00330-002-1746-y

8. Rutten A, Isgum I, Prokop M. Coronary calcification: effect of small variation of scan starting position on Agatston, volume, and mass scores. Radiology. 2008;246(1):90-98. doi:10.1148/radiol.2461070006

9. Aslam A, Khokhar US, Chaudhry A, et al. Assessment of isotropic calcium using 0.5-mm reconstructions from 320-row CT data sets identifies more patients with non-zero Agatston score and more subclinical atherosclerosis than standard 3.0-mm coronary artery calcium scan and CT angiography. J Cardiovasc Comput Tomogr. 2014;8(1):58-66. doi:10.1016/j.jcct.2013.12.007

10. Otsuka F, Sakakura K, Yahagi K, Joner M, Virmani R. Has our understanding of calcification in human coronary atherosclerosis progressed? Arterioscler Thromb Vasc Biol. 2014;34(4):724-736. doi:10.1161/ATVBAHA.113.302642

11. Blaha MJ, Mortensen MB, Kianoush S, Tota-Maharaj R, Cainzos-Achirica M. Coronary Artery Calcium Scoring. JACC Cardiovasc Imaging. 2017;10(8):923-937. doi:10.1016/j.jcmg.2017.05.007

12. Han D, Klein E, Friedman J, et al. Prognostic significance of subtle coronary calcification in patients with zero coronary artery calcium score: From the CONFIRM registry. Atherosclerosis. 2020;309(June):33-38. doi:10.1016/j.atherosclerosis.2020.07.011

13. Vonder M, van der Werf NR, Leiner T, et al. The impact of dose reduction on the quantification of coronary artery calcifications and risk categorization: A systematic review. J Cardiovasc Comput Tomogr. 2018;12(5):352-363. doi:10.1016/j.jcct.2018.06.001

14. Leng S, Hruska CB, McCollough CH. Use of ionizing radiation in screening examinations for coronary artery calcium and cancers of the lung, colon, and breast. Semin Roentgenol. 2015;50(2):148-160. doi:10.1053/j.ro.2014.10.012

14

15. Vingiani V, Abadia AF, Schoepf UJ, et al. Individualized coronary calcium scoring at any tube voltage using a kV-independent reconstruction algorithm. Eur Radiol. Published online May 2020. doi:10.1007/s00330-020-06951-1

16. Tesche C, De Cecco C, Vliegenthart R, et al. Accuracy and Radiation Dose Reduction Using Low-Voltage Computed Tomography Coronary Artery Calcium Scoring With Tin Filtration. Am J Cardiol. 2017;119(4):675-680. doi:10.1016/j.amjcard.2016.10.051

17. Hecht HS, De Siqueira MEM, Cham M, et al. Low- vs. Standard-dose coronary artery calcium scanning. Eur Heart J Cardiovasc Imaging. 2015;16(4):358-363. doi:10.1093/ehjci/jeu218

18. Groen JM, Greuter MJW, Vliegenthart R, et al. Calcium scoring using 64-slice MDCT, dual source CT and EBT: A comparative phantom study. Int J Cardiovasc Imaging. 2008;24(5):547-556. doi:10.1007/s10554-007-9282-0

19. van Praagh GD, van der Werf NR, Wang J, et al. Fully Automated Quantification Method (FQM) of Coronary Calcium in an Anthropomorphic Phantom. Med Phys. Published online 2021:in press. doi:10.1002/mp.14912

20. Nakazato R, Dey D, Gutstein A, et al. Coronary artery calcium scoring using a reduced tube voltage and radiation dose protocol with dual-source computed tomography. J Cardiovasc Comput Tomogr. 2009;3(6):394-400. doi:10.1016/j.jcct.2009.10.002

21. Hou KY, Tsujioka K, Yang CC. Optimization of HU threshold for coronary artery calcium scans reconstructed at 0.5-mm slice thickness using iterative reconstruction. J Appl Clin Med Phys. 2020;21(2):111-120. doi:10.1002/acm2.12806

22. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. J Am Coll Cardiol. 1990;15(4):827-832. doi:10.1016/0735-1097(90)90282-T

23. Callister TQ, Cooil B, Raya SP, Lippolis NJ, Russo DJ, Raggi P. Coronary artery disease: improved reproducibility of calcium scoring with an electron-beam CT volumetric method. Radiology. 1998;208(3):807-814. doi:10.1148/radiology.208.3.9722864

24. Blaha MJ, Budoff MJ, Tota-Maharaj R, et al. Improving the CAC Score by Addition of Regional Measures of Calcium Distribution: Multi-Ethnic Study of Atherosclerosis. JACC Cardiovasc Imaging. 2016;9(12):1407-1416. doi:10.1016/j.jcmg.2016.03.001

25. Criqui MH, Denenberg JO, Ix JH, et al. Calcium Density of Coronary Artery Plaque and Risk of Incident Cardiovascular Events. Jama. 2014;311(3):271. doi:10.1001/jama.2013.282535

26. Choi AD, Leifer ES, Yu JH, et al. Reduced radiation dose with model based iterative reconstruction coronary artery calcium scoring. Eur J Radiol. 2019;111:1-5. doi:10.1016/j.ejrad.2018.12.010

27. Caruso D, De Santis D, Biondi T, et al. Half-dose coronary artery calcium scoring: Impact of iterative reconstruction. J Thorac Imaging. 2019;34(1):18-25. doi:10.1097/RTI.0000000000000340

28. Blobel J, Mews J, Schuijf JD, Overlaet W, Dipl-Ing Desiree Schuijf J, Overlaet W. Determining the Radiation Dose Reduction Potential for Coronary Calcium Scanning With Computed Tomography the Adaptive Iterative Dose Reduction Algorithm for Image Reconstruction. Invest Radiol. 2013;48(12):857-862.

29. Mantini C, Maffei E, Toia P, et al. Influence of image reconstruction parameters on cardiovascular risk reclassification by Computed Tomography Coronary Artery Calcium Score. Eur J Radiol. 2018;101:1-7. doi:10.1016/j.ejrad.2018.01.005

14

30. Horiguchi J, Matsuura N, Yamamoto H, et al. Variability of repeated coronary artery calcium measurements by 1.25-mm- and 2.5-mm-thickness images on prospective electrocardiograph-triggered 64-slice CT. Eur Radiol. 2008;18(2):209-216. doi:10.1007/s00330-007-0734-7

31. Muhlenbruch G, Thomas C, Wildberger JE, et al. Effect of varying slice thickness on coronary calcium scoring with multislice computed tomography in vitro and in vivo. Invest Radiol. 2005;40(11):695-699.

32. Vliegenthart R, Song B, Hofman A, Witteman JCM, Oudkerk M. Coronary Calcification at Electron-Beam CT: Effect of Section Thickness on Calcium Scoring in Vitro and in Vivo. Radiology. 2003;229(2):520-525. doi:10.1148/radiol.2292021305

33. Vonder M, Pelgrim GJ, Meyer M, Henzler T, Oudkerk M, Vliegenthart R. Dose reduction techniques in coronary calcium scoring: The effect of iterative reconstruction combined with low tube voltage on calcium scores in a thoracic phantom. Eur J Radiol. 2017;93(January):229-235. doi:10.1016/j.ejrad.2017.06.001

34. Groen JM, Kofoed KF, Zacho M, Vliegenthart R, Willems TP, Greuter MJW. Calcium score of small coronary calcifications on multidetector computed tomography: Results from a static phantom study. Eur J Radiol. 2013;82(2):e58-e63. doi:10.1016/j.ejrad.2012.09.018

35. Blaha M, Budoff MJ, Shaw LJ, et al. Absence of Coronary Artery Calcification and All-Cause Mortality. JACC Cardiovasc Imaging. 2009;2(6):692-700. doi:10.1016/j.jcmg.2009.03.009

36. Sarwar A, Shaw LJ, Shapiro MD, et al. Diagnostic and Prognostic Value of Absence of Coronary Artery Calcification. JACC Cardiovasc Imaging. 2009;2(6):675-688. doi:10.1016/j.jcmg.2008.12.031

37. Willemink MJ, Takx RAP, De Jong PA, et al. The impact of CT radiation dose reduction and iterative reconstruction algorithms from four different vendors on coronary calcium scoring. Eur Radiol. 2014;24(9):2201-2212. doi:10.1007/s00330-014-3217-7

38. Marwan M, Mettin C, Pflederer T, et al. Very low-dose coronary artery calcium scanning with high-pitch spiral acquisition mode: Comparison between 120-kV and 100-kV tube voltage protocols. J Cardiovasc Comput Tomogr. 2013;7(1):32-38. doi:10.1016/j.jcct.2012.11.004

39. Gräni C, Vontobel J, Benz DC, et al. Ultra-low-dose coronary artery calcium scoring using novel scoring thresholds for low tube voltage protocols—a pilot study. Eur Hear J - Cardiovasc Imaging. 2018;(February):1-10. doi:10.1093/ehjci/jey019

40. van der Werf NR, Willemink MJ, Willems TP, Vliegenthart R, Greuter MJW, Leiner T. Influence of heart rate on coronary calcium scores: a multi-manufacturer phantom study. Int J Cardiovasc Imaging. 2017;34(6):959-966. doi:10.1007/s10554-017-1293-x

41. Tao S, Sheedy E, Bruesewitz M, et al. Technical Note: kV-independent coronary calcium scoring: A phantom evaluation of score accuracy and potential radiation dose reduction. Med Phys. 2021;48(3):1307-1314. doi:10.1002/mp.14663

14

Online material is available via:

PART III
Post-processing parameters
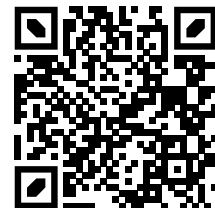
CHAPTER 15

Vendor independent coronary calcium scoring improves individual risk assessment - the Multi-Ethnic Study of Atherosclerosis (MESA)

Niels R. van der Werf*, MSc

Magdalena M. Dobrolinska*, MD

Marcel J.W. Greuter, PhD

Martin J. Willemink, MD PhD

Dominik Fleischmann, MD

Daniel Bos, MD PhD

Riemer H.J.A. Slart, MD PhD

Matthew J. Budoff, MD

Tim Leiner, MD PhD

Shared first authors

ABSTRACT

Background

The substantial variation in Agatston scores (AS) acquired with different computed tomography (CT) scanners may influence patient risk classification. Accuracy of AS, and thus coronary heart disease (CHD) event prediction for patients imaged on different CT systems may be improved by a vendor-neutral AS (vnAS).

This study sought to develop a calibration tool for state-of-the-art CT systems resulting in a vnAS, and to assess the impact of vnAS on CHD event prediction.

Methods

The vnAS calibration tool was derived by imaging two anthropomorphic calcium containing phantoms on seven different CT and one EBT system. To assess the effect of vnAS on event prediction we analyzed data from 3,181 participants from the Multi-Ethnic Study on Atherosclerosis (MESA). Chi-square analysis was used to compare CHD event rates between low (vnAS < 100) and high calcium groups (vnAS ≥ 100). Multivariable Cox proportional hazard regression models were used to assess the incremental value of vnAS.

Results

For all CT systems, a high degree of correlation with EBT AS was shown (R^2 > 0.932). Of the MESA participants originally in the low calcium group (n = 781), 85 (11%) participants were reclassified to a higher risk category based on the recalculated AS, or vnAS. The CHD event rate of 15% for reclassified participants was significantly higher compared to participants in the low calcium group (7%, p = 0.008) with a CHD hazard ratio of 3.39 (95% CI 1.82 – 6.35, p = 0.001).

Conclusion

We developed a calibration tool, which enables to calculate vnAS. MESA participants who were reclassified to a higher calcium category based on vnAS, showed increased CHD event rates indicating improved risk categorization of vnAS. This indicated improved risk categorization with vnAS. Moreover, vnAS also improved statin therapy classification for patients at intermediate ASCVD risk, with reduced NNT. Therefore, our vendor-neutral CAC assessment may improve patient risk classification and subsequent patient management.

15

INTRODUCTION

Computed tomography (CT) coronary artery calcium (CAC) measurements improve risk stratification for both coronary heart disease (CHD) and atherosclerotic cardiovascular disease (ASCVD) in asymptomatic individuals beyond traditional risk factors.[1–9] Due to its risk reclassification potential, CAC is an increasingly important tool to support the decision to initiate or defer statin therapy in individuals at low-to-intermediate risk.[10] CAC assessment is therefore recommended in both American and European guidelines, with statin therapy initiation for individuals at intermediate ASCVD risk with a CAC score of ≥ 100, and intensive statin therapy for individuals with a CAC score ≥ 300.[11–13] Moreover, follow-up CAC assessment is recommended as a way of monitoring atherosclerotic disease.[11,14]

Clinically, CAC assessment is performed according to the Agatston method.[15] This method was developed in 1990 using now almost obsolete electron beam tomography (EBT). Importantly, CAC based on EBT scans occurred to add incremental value to commonly known cardiovascular risk factors, including the Framingham Risk score.[1,16]

Nowadays, multi-detector CT (MDCT) systems are used in clinical practice. For these systems, however, it has been shown that CAC detection and quantification differs substantially not only between different CT manufacturers, but also between different CT systems from the same manufacturer, with differences in median Agatston score (AS) of up to 44%.[17] These variations in AS may translate into risk misclassification and improper treatment in up to 6.5% of asymptomatic individuals.[17] Despite these variations, there are presently no tools or methods to convert MDCT-derived AS, as obtained with modern CT systems, to EBT-derived AS.

We therefore sought to: 1) develop a calibration tool for current state-of-the-art CT systems, allowing for calculation of a vendor-neutral AS (vnAS) which would be equal to the AS an individual would have had if scanned on an EBT system; 2) assess whether the vnAS improves event prediction for both CHD and ASCVD for participants of the Multi-Ethnic Study of Atherosclerosis (MESA); and 3) determine the influence of vnAS on initiation or deferral of statin therapy.

15

METHODS

Calibration tool development

Phantom

The vnAS calibration tool was derived from scans of two static anthropomorphic CAC containing inserts, which were consecutively placed at the center of an anthropomorphic chest phantom (QRM thorax, PTW, Freiburg, Germany). Both phantom inserts contained cylindrical calcifications composed of hydroxyapatite (HA) with a variety of sizes and densities to cover the large variety of in-vivo CAC. The first insert contained 100 small cylindrical calcifications (D100, PTW, Freiburg, Germany) varying in HA density (90 to 540 mg/cm^3) and volume (0.1 to 6.3 mm^3). The second insert contained 9 cylindrical calcifications (CCI, PTW, Freiburg, Germany) with three HA densities (200, 400, and 800 mg/cm^3) and three sizes (0.8, 21.2, and 98.2 mm^3).[18]

The chest phantom itself was composed of a shell of tissue-equivalent material, a spine insert, and artificial lungs. To increase phantom size, a fat equivalent extension ring (QRM extension ring, PTW, Freiburg, Germany) was used. With this extension ring, the phantom size was increased from 300 x 200 cm^2 to 400 x 300 cm^2, to simulate a large patient size (body mass index (BMI) $\geq$ 25 kg/m^2).[18]

Data acquisition

Chest CT examinations were performed using routine clinical CAC protocols on five state-of-the-art MDCT systems: IQon and iCT (Philips Healthcare, Best, The Netherlands), Force (Siemens Healthineers, Forchheim, Germany), Revolution CT (GE Healthcare, Waukesha, WI), and Aquilion One (Canon Medical Systems Corporation, Otawara, Japan), designated CT-A to CT-E, respectively (*Table 1*). In addition, we acquired data on two older MDCT systems which were used to scan participants of the MESA: a Lightspeed Plus (GE Healthcare, Waukesha, WI) and a Volume Zoom (Siemens Healthcare, Forchheim, Germany) CT system, CT-F and CT-G, respectively. For these older systems, MESA CAC protocols were used (*Table 1*).[19] Reference standard acquisitions were obtained on an EBT C-150 system (Imatron, San Francisco, CA, United States of America).

For all systems (7 CT and 1 EBT), raw data was reconstructed using filtered back projection (FBP). For each combination of insert and phantom size, five acquisitions and reconstructions were made, with a manual translation (approximately 2 mm) and rotation (approximately 2 degrees) of the phantom between each acquisition to correct for inter-scan variability.

15

Table 1 Acquisition and reconstruction parameters for the reference standard electron beam tomography (EBT) and CT systems

Parameter	Reference-Standard	CT-A	CT-B	CT-C	CT-D	CT-E	CT-F	CT-G
Manufacturer	Imatron	Philips	Philips	Siemens	GE	Canon	GE	Siemens
CT system	C-150	Iqon	Brilliance iCT	SOMATOM Force	Revolution	Aquilion one	Lightspeed	Volume Zoom
Acquisition mode	Axial	Axial	Axial	Axial	Axial	Axial	Axial	Axial
Tube voltage [kVp]	130	120	120	120	120	120	120	140
Tube current time product [mAs]	Small: 63 Large: 63	Small: 40 Large: 48	Small: 50 Large: 60	Small: 44 Large: 194	Small: 30 Large: 161	SD = 55	Small: 106 Large: 132	Small: 50 Large: 63
AEC	Off	Off	Off	Off	Off	On	Off	Off
Collimation [mm]	-	64x0.625	128x0.625	160x0.6	224x0.625	280x0.5	16x0.625	4 x 2.5
Field of View [mm]	250	250	250	250	250	250	220	250
Rotation time [s]	-	0.27	0.27	0.25	0.28	0.275	0.5	0.5
Slice thickness [mm]	3.0	3.0	3.0	3.0	2.5	3.0	2.5	2.5
Increment [mm]	3.0	3.0	3.0	1.5	2.5	3.0	2.5	2.5
Reconstruction kernel	Sharp R62	XCA	XCA	Qr36d	Standard	FC12	Standard	B35f
Matrix size [pixels]	512x512	512x512	512x512	512x512	512x512	512x512	512x512	512x512
Reconstruction	FBP	iDose – level 0 (FBP)[1]	FBP	FBP	FBP	FBP	FBP	FBP

AEC = Automatic exposure control, SD = Standard deviation (image noise), FBP = Filtered back projection
[1] For CT-A, iDose level 0 is the iterative reconstruction technique which matches FBP the closest

15

vnAS calculator specification and processing

AS were calculated for all CT reconstructions with vendor-specific CAC scoring parameters using a previously validated Python script.[20] For the EBT, vendor-specific CAC scoring software (AccuImage, AccuImage, Diagnostic Corporation, San Francisco, CA) was used. Next, for each combination of phantom size and CT system, linear regression was used to convert the obtained AS to a vnAS for all repeated measurements. Regression model fit was assessed using both R^2 and the ANOVA methodology. The resulting regression model was used in a calculator, where the 3 input parameters (CT system, patient size, and AS) were used to calculate the vnAS (*Figure 1*). The vnAS calculator will be made publicly available.
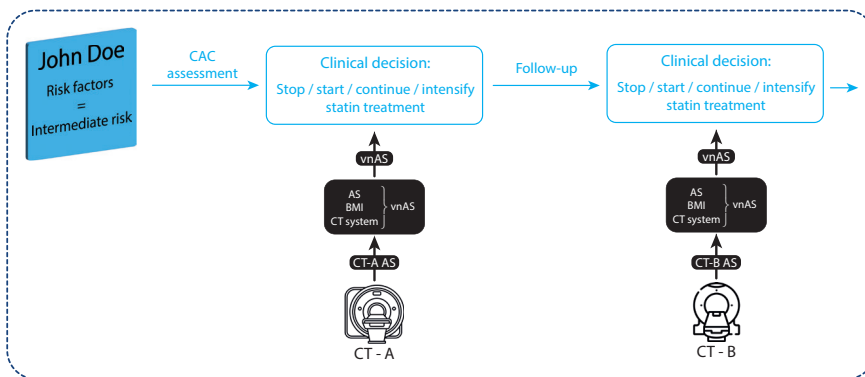
Figure 1 Illustration of the use of the vendor-neutral Agatston score (vnAS) calculator, which converts a MDCT specific Agatston score (AS) (CT-A AS or CT-B AS) into a vnAS based on patient and CT specific parameters. The vnAS can subsequently be used for clinical decisions on statin treatment for patients at intermediate atherosclerotic cardiovascular disease risk, both initially and in follow-up assessments

Validation study: MESA population

Data sharing

Data from the MESA was requested via the MESA website (https://www.mesa-nhlbi.org) following a described data request procedure. MESA data was shared based on a signed data sharing agreement.

Study population

MESA is a multi-center, cohort (n = 6,814) study involving 6 centers (Forsyth County, NC; Bronx and Northern Manhattan, NY; Baltimore City and County, MD; St. Paul, MN; Chicago, IL and Los Angeles County, CA).[2] The rationale for the multi-center prospective, observational cohort study was to determine the underlying factors of cardiovascular diseases, and the study design and methods have been previously described in detail elsewhere.[2] In brief, investigators included 6,814 asymptomatic individuals between 45 and 84 years of age. At baseline, demographic data was collected from each participant and two immediate subsequent CT scans were performed. All study participants provided written informed consent at baseline.

Findings presented in the current analysis represent both baseline data as collected between July 2000 and September 2002, and follow-up information about CHD and ASCVD events from follow-up contacts every 9-12 months from 2000 through December 2018. CHD events were defined as myocardial infarction, resuscitated cardiac arrest, fatal CHD, or revascularization. ASCVD events were defined as nonfatal or fatal myocardial infarction, resuscitated cardiac arrest, probable angina, definite angina followed by revascularization, non-fatal or fatal stroke, other atherosclerotic death, or other cardiovascular death.[6] As only CT-F and CT-G were used within MESA, participants who were not scanned on either CT-F or CT-G were excluded. Also, those for whom follow-up or risk factor data was missing were excluded, which resulted in 3,181 (1,077 and 2,152 on CT-F and CT-G, respectively) included individuals for the current analysis on CHD and ASCVD event prediction in cohort 1 (*Figure 2*). To assess the influence of vnAS on the initiation or deferral of statin therapy, a sub-cohort of participants who were at intermediate cardiovascular risk was formed.[11,13] For this cohort 2, 889 individuals were included (*Figure 2*).[11,13] Intermediate cardiovascular risk was defined as individuals with an ASCVD risk between 7.5-19.9 %, without diabetes, LDL-C 70-189 mg/dl, and without previous ASCVD events.[12,13,21]

vnAS calculation

For all included participants, vnAS were calculated based on the patient size specific regression models from either CT-F or CT-G, as appropriate. A BMI cutoff value of 25 kg/m^2 was used to differentiate between average and large sized participants.

15
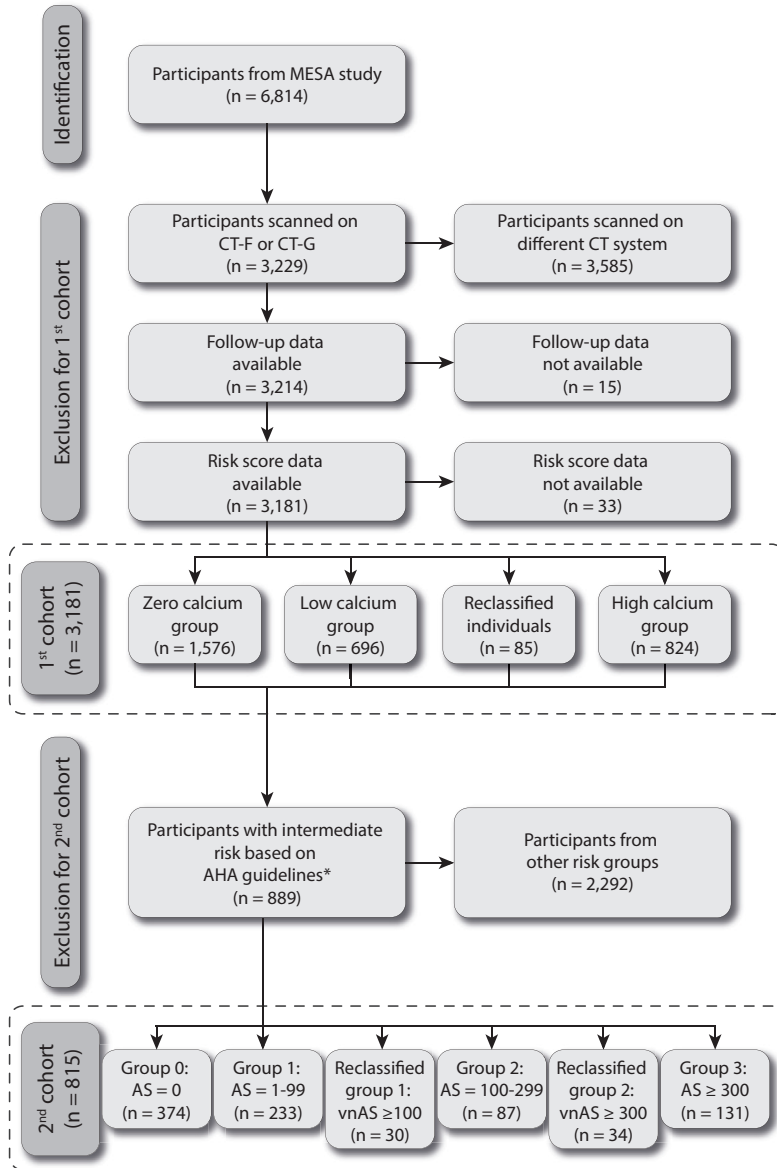
Figure 2 Flowchart of included participants from the MESA study. Both individuals included in the first and second cohort are indicated between the dashed lines. For the second cohort, all used groups are indicated: group 1 includes individuals with an Agatston score (AS) of 0; group 2 includes participants with an AS of 1 to 99; group 3 includes participants with an AS < 100 but a vendor-neutral AS (vnAS) of ≥100; group 4 includes participants with an AS of 100 to 299; group 5 includes participants with an AS < 300 and a vnAS of ≥300; finally, group 6 includes participants with AS ≥ 300.

CHD and ASCVD event prediction: cohort 1

Based on baseline MDCT-AS, participants from cohort 1 (n = 3,181) were assigned to zero, low, or high coronary calcium groups, with AS = 0, AS 1 – 99, or AS ≥ 100, respectively (*Table 2*). Participants with AS 1 – 99 and vnAS ≥ 100 were considered reclassified individuals (*Table 2*). For all four groups (zero, low, high, reclassified) CHD and ASCVD event rates were compared.

Table 2 Cohort 1 group definition for CHD and ASCVD event prediction based on Agatston score (AS) and vendor-neutral AS (vnAS)

Group definition	AS	vnAS	Number of MESA participants [#]
Zero coronary calcium	0	0	1576
Low coronary calcium	< 100	< 100	696
Reclassified	< 100	≥ 100	85
High coronary calcium	≥ 100	≥ 100	824

Statin therapy assessment: cohort 2

In addition to studying event rates, the potential benefit from statin therapy was assessed for MESA participants at intermediate cardiovascular risk (cohort 2, n = 889). The impact of using the vnAS was quantitatively expressed as change in the number needed to treat (NNT).[22] The NNT was calculated as the reciprocal of the absolute risk reduction (ARR) for an ASCVD event between individuals with and without statin therapy. The ARR is the absolute value of the difference between the ASCVD events in the non-statin group and statin group.[23] For the NNT calculation, ASCVD events during the entire 18 years of follow-up were used.

To calculate the NNT, cohort 2 was divided in four groups based on three thresholds: AS = 0, AS ≥ 100 for statin therapy initiation, or AS ≥ 300 for intensive statin therapy initiation (*Table 3*).[5,11] The resulting groups were AS = 0, AS 1-99, AS ≥ 100, and AS ≥ 300 which were designated group 0 to 3, respectively (*Figure 2* and *Table 3*). Based on vnAS, two additional groups were formed: reclassified group 1 with AS 1-99 and vnAS ≥ 100, and reclassified group 2 with AS < 300 and vnAS ≥ 300 (*Figure 2* and *Table 3*).

15

Table 3 Cohort 2 group definition for statin therapy assessment based on Agatston score (AS) and vendor-neutral AS (vnAS). For each group, the percentage of atherosclerotic cardiovascular (ASCVD) events is indicated

Group definition	AS	vnAS	Number of participants [#]	ASCVD events [%]
Group 0	0	0	374	10.7
Group 1	< 100	< 100	233	17.6
Reclassified group 1	< 100	≥ 100	30	23.3
Group 2	100-299	100-299	87	27.6
Reclassified group 2	100-299	≥ 300	34	38.2
Group 3	≥ 300	≥ 300	131	32.8

Statistical analysis

Study population characteristics were stratified according to calcium score groups (*Table 4*) and statin therapy group (*Table 6*) for cohorts 1 and 2, respectively. For continuous variables either means (with standard deviation) or medians (with interquartile range (IQR) were calculated. Normality was visually assessed based on histograms and Q-Q plots. To compare variables between groups, Chi-square test, Kruskal-Wallis test, and one-way analysis of variance were used, as appropriate.

Differences in both CHD and ASCVD events between the four coronary calcium groups was assessed with Kaplan-Meier curves and log-rank tests. The association of age, race, gender, systolic blood pressure, anti-hypertensive medication, high-density lipoprotein, total cholesterol, lipid-lowering medication, cigarette smoking status, diabetes, and coronary calcium group (zero, low, reclassified individuals, high) with first-ever CHD or ASCVD event was assessed using univariable Cox proportional hazard regressions models. Only the first CHD or ASCVD event was included, within 18 years of follow-up data. Next, all parameters were simultaneously used for a multivariable Cox proportional hazard regression model.

All statistical analyses were conducted using SPSS 27 (IBM, Armonk NY) or MedCalc 15.8 (MedCalc Software). Statistical significance was defined as $p < 0.05$.

15

Table 4 Characteristics of 3,181 individuals (cohort 1) of the MESA study with CT scans on CT-F or CT-G, stratified by calcium group

	Zero calcium	Low calcium	Reclassified individuals	High calcium	P value[1]
Total number	1,576	696	85	824	
Age (year)	58 ± 9	63 ± 9	67 ± 8	69 ± 9	< 0.001
Race/Ethnicity					< 0.001
White	735 (46.6)	368 (52.9)	40 (47.1)	549 (66.6)	
Black	572 (36.3)	232 (33.3)	37 (43.5)	190 (23.1)	
Hispanic	269 (17.1)	96 (13.8)	8 (9.4	85 (10.3)	
Gender					< 0.001
Female	998 (63.3)	336 (48.3)	34 (40.0)	303 (36.8)	
Male	578 (36.7)	360 (51.7)	51 (60.0)	521 (63.2)	
Diabetes mellitus	142 (9.0)	86 (12.4)	20 (23.5)	135 (16.4)	< 0.001
Systolic blood pressure	124 ± 21	126 ± 21	134 ± 22	134 ± 22	< 0.001
Hypertension medication	500 (31.7)	283 (40.7)	45 (52.9)	421 (51.1)	< 0.001
HDL	52 ± 15	49 ± 14	48 ± 13	49 ± 14	< 0.001
LDL	116.4 ± 32	119 ± 32	113 ± 31	117 ± 30	0.175
Total Cholesterol	194 ± 36	194 ± 37	185 ± 33	192 ± 33	0.086
Triglycerides	108 (76.0, 19.0)	113.0 (80.0, 162.0)	105.0 (78.0, 160.0)	134.2 (77.2, 166.0)	0.078
Any lipid-lowering therapy	200 (12.7)	137 (19.7)	28 (32.9)	193 (23.4)	< 0.001
Cigarettes smoking	244 (15.5)	696 (15.4)	7 (8.2)	116 (14.1)	0.266
Family history of heart attack	614 (39.0)	334 (48.0)	36 (42.4)	439 (53.3)	< 0.001
Coronary calcium score					
vnAS	0.0	30.5 (14.1, 56.7)	117.1 (105.3, 126.1)	801.6 (243.0, 910.1)	
AS	0.0	22.4 (10.2, 42,4)	87.2 (78.1, 94.2)	619.2 (186.2, 736.9)	
Risk calculators					
ASCVD risk	5.5 (2.4, 11.6)	13.9 (5.7, 19.6)	20.4 (11.3, 25.7)	22.1 (11.3, 29.3)	< 0.001
Events					
CHD	54 (3.4)	49 (7.0)	13 (15.3)	152 (18.4)	< 0.001
ASCVD	138 (8.8)	127 (18.2)	20 (23.5)	290 (35.2)	< 0.001

[1] Chi-square test

AS – Agatston score; ASCVD – atherosclerotic cardiovascular disease; CHD – coronary heart disease, HDL – high density lipoprotein, LDL- low density lipoprotein, Reclassified individuals – individuals with AS < 100 and vnAS ≥ 100; vnAS – vendor-neutral Agatston score,

15

15

Table 5 Characteristics of 889 individuals (cohort 2) of the MESA study with CT scans on CT-F or CT-G and at intermediate cardiovascular risk, stratified by statin therapy group

	Group 0 (AS = 0)	Group 1 (AS = 1 - 99)	Reclassified group 1 (vnAS ≥ 100)	Group 2 (AS = 100 - 299)	Reclassified group 2 (vnAS ≥ 300)	Group 3 (AS ≥ 300)
Total number	374	233	30	87	34	131
Age (year)	64 ± 7	65 ± 7	67 ± 6	67 ± 7	67 ± 7	64 ± 8
Race/Ethnicity						
White	133 (35.6)	118 (50.6)	14 (46.7)	60 (69.0)	20 (58.8)	94 (71.8)
Black	186 (49.7)	87 (37.3)	14 (46.7)	22 (25.3)	14 (41.2)	24 (18.3)
Hispanic	55 (14.7)	28 (12.0)	2 (6.7)	5 (5.7)	0 (0.0)	13 (9.9)
Gender						
Female	197 (52.7)	101 (43.3)	14 (46.7)	51 (58.6)	12 (35.3)	42 (32.1)
Male	177 (47.3)	132 (56.7)	16 (53.3)	51 (58.6)	22 (64.7)	89 (67.9)
Diabetes mellitus						
Systolic blood pressure	132.6 ± 18.4	129.3 ± 17.4	132.7 ± 19.8	126.2 ± 17.9	134.5 ± 23.2	127.6 ± 16.6
Hypertension medication	151 (40.4)	88 (37.8)	17 (56.7)	42 (48.3)	16 (47.1)	46 (35.1)
HDL	51.3 ± 15.5	49.1 ± 14.8	48.7 ± 12.1	50.3 ± 15.5	44.3 ± 11.2	49.5 ± 13.4
LDL	121 ± 26	121 ± 27	110.3 ± 23.8	120.5 ± 27.1	120.1 ± 23.1	120.9 ± 26.5
Total Cholesterol	198.5 ± 30.1	195.8 ± 31.8	179.3 ± 27.3	194.1 ± 27.8	194.5 ± 27.4	196.7 ± 28.3
Triglycerides	112.0 (83.0, 158.2)	112.0 (79.0, 161.0)	91.0 (67.5, 114.7)	102.0 (68.0, 143.0)	122.5 (77.7, 201.2)	116.0 (84.0, 162.0)
Any lipid-lowering therapy	65 (17.4)	37 (15.9)	9 (30.0)	22 (25.3)	7 (20.6)	35 (26.7)
Cigarettes smoking	74 (19.8)	41 (17.6)	2 (6.7)	7 (8.0)	9 (26.5)	18 (13.7)
Family history of heart attack	145 (38.8)	112 (48.1)	17 (56.7)	48 (55.2)	15 (44.1)	75 (57.3)
Coronary calcium score						
vnAS	0.0	39.5 (14.7, 64.7)	117.1 (110.2, 127.1)	198.6 (160.4, 239.4)	338.9 (322.6, 368.5)	825.1 (533.3, 1586.8)
AS	0.0	30.2 (10.3, 49.3)	87.2 (81.8, 94.9)	154.2 (128.9, 181.5)	253.5 (241.0, 274.1)	674.1 (413.9, 1263.3)
Risk calculators						
ASCVD risk	11.1 ± 3.3	12.5 ± 3.5	14.1 ± 3.8	12.7 ± 3.7	13.8 ± 3.3	14.1 ± 3.3
Events						
CHD	13 (3.5)	17 (7.3)	3 (10.0)	15 (17.2)	8 (23.5)	20 (15.3)
ASCVD	40 (10.7)	41 (17.6)	7 (23.3)	24 (27.6)	13 (38.2)	43 (32.8)
	374	233	30	87	34	131

Table 6 Example vendor-neutral Agatston scores for a CT Agatston score of 100 for both small (body mass index (BMI) < 25 kg/m²) and large (BMI ≥ 25 kg/m²) patients

Patient size	CT-A	CT-B	CT-C	CT-D	CT-E	CT-F	CT-G
Small	105	107	94	112	93	120	102
Large	129	122	121	148	88	139	123

RESULTS

vnAS calculation: phantom study

Irrespective of MDCT system, a high degree of correlation with EBT derived AS was found ($R^2 \geq 0.932$) in *Figure 3* and *Figure 4* for the small and large phantom, respectively. For all MDCT systems, the linear regression models, which predicted EBT AS, were statistically significantly (all $P < 0.001$). These prediction models, or vnAS, convert MDCT AS to AS acquired on EBT. For the small phantom size, vnAS were lower than non-corrected AS for CT-C and CT-E, while for the large phantom size, only CT-E showed lower vnAS than AS.

For example, for a patient with a BMI > 25 kg/m² and an MDCT-derived AS of 100, the vnAS varied between 88 - 148 depending on the specific CT system that was used (*Table 6*). For patients with an AS of 0, vnAS was always 0.

Validation study: MESA participants

CHD and ASCVD event prediction: cohort 1

The mean age of cohort 1 participants (n = 3,181) was 62 ± 10 years, 52.5% were women, and 46.8% of participants were of non-Caucasian ethnicity (*Table 4*). For both CT-F and CT-G, vnAS was higher than AS (up to 39%). This resulted in 85 (85 / (696 + 85) = 11%) reclassified individuals, who were reclassified from the low to the high coronary calcium group (*Table 2*).

Of the reclassified individuals, 13 (15.3%) and 20 (23.5%) experienced a CHD or ASCVD event during 16.7 (IQR 4.8) years of follow-up (*Figure 5*). In comparison with the low coronary calcium group, CHD and ASCVD event rates of the reclassified individuals were 8.3% (P = 0.008) and 5.3% (P = 0.24) higher, respectively. This was also reflected by the Kaplan-Meier curves, which confirmed substantial differences

15

15

$$vnAS = \begin{cases} 0 & \text{if CT-X} = 0 \\ 1.351 + 1.041 \cdot \text{CT-A} \\ 1.449 + 1.059 \cdot \text{CT-B} \\ 1.184 + 0.924 \cdot \text{CT-C} \\ 0.320 + 1.112 \cdot \text{CT-D} \\ 0.062 + 0.929 \cdot \text{CT-E} \\ 0.049 + 1.197 \cdot \text{CT-F} \\ 1.195 + 1.070 \cdot \text{CT-G} \end{cases}$$

Figure 3 Small phantom CT specific Agatston scores (AS) versus electron beam tomography (EBT) AS for both the CCI (light-grey) and D100 (dark grey) insert. For each combination of CT system and EBT the coefficient of determination (R^2) and linear regression model parameters are indicated. CT-X refers to the AS of any CT, CT-A is the AS of CT-A, etc. The dashed line shows the linear regression model which is used to calculate the vendor-neutral AS (vnAS)
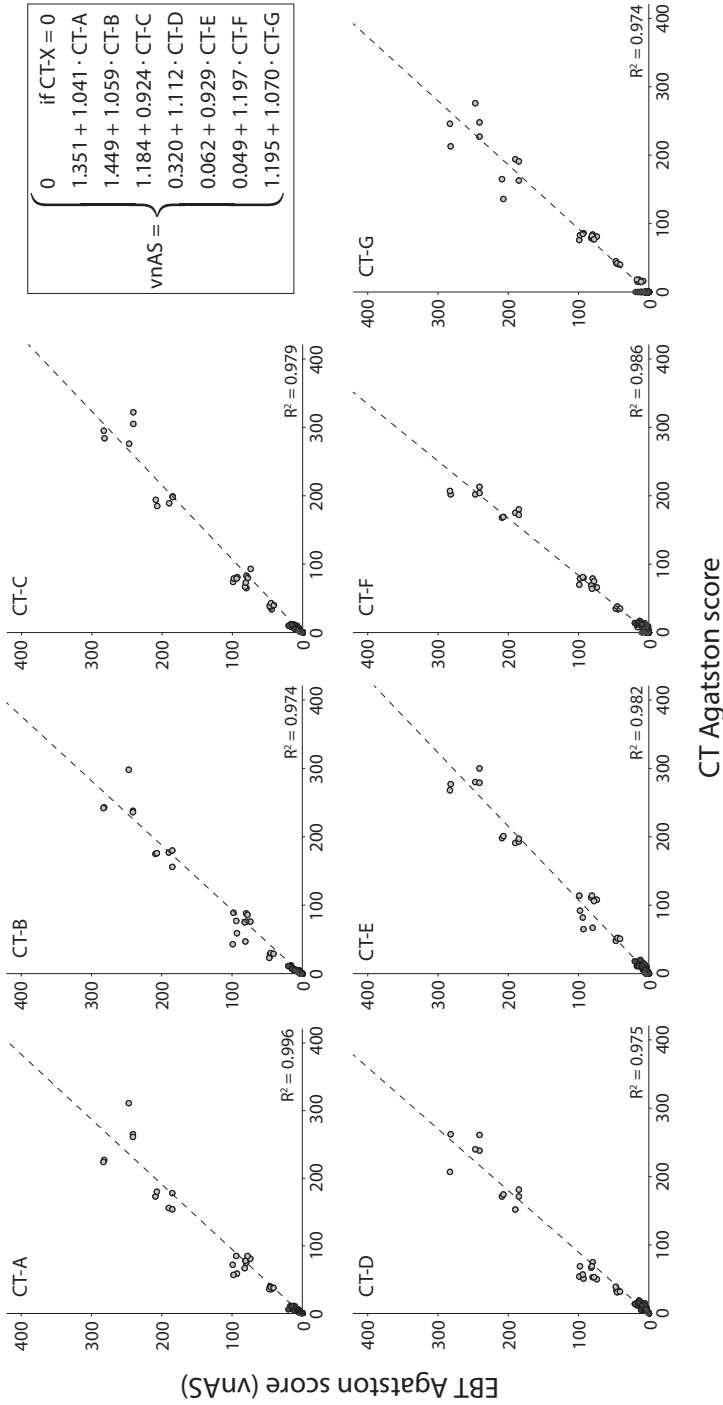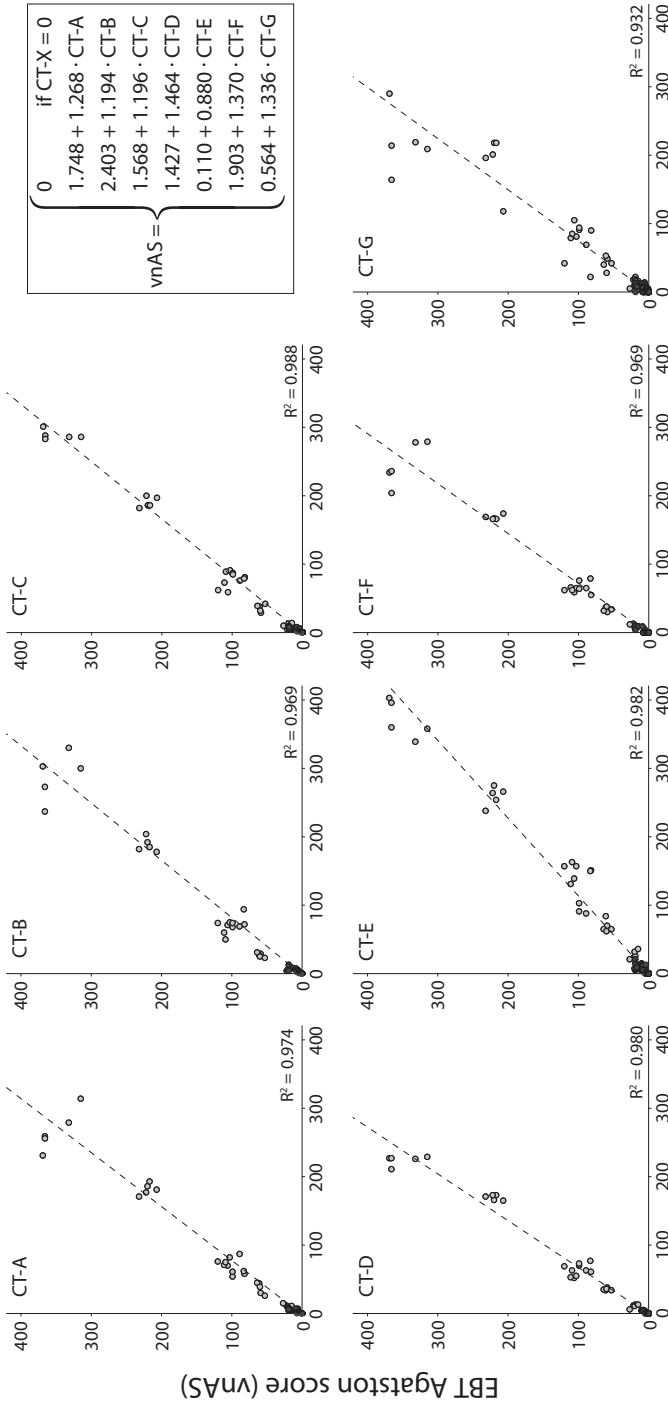
Figure 4 Large phantom CT specific Agatston scores (AS) versus electron beam tomography (EBT) AS for both the CCI (light-grey) and D100 (dark grey) insert. For each combination of CT system and EBT the coefficient of determination (R²) and linear regression model parameters are indicated. CT-X refers to the AS of any CT, CT-A is the AS of CT-A, etc. The dashed line shows the linear regression model which is used to calculate the vendor-neutral AS (vnAS)

$$vnAS = \begin{cases} 0 & if\ CT\text{-}X = 0 \\ 1.748 + 1.268 \cdot CT\text{-}A \\ 2.403 + 1.194 \cdot CT\text{-}B \\ 1.568 + 1.196 \cdot CT\text{-}C \\ 1.427 + 1.464 \cdot CT\text{-}D \\ 0.110 + 0.880 \cdot CT\text{-}E \\ 1.903 + 1.370 \cdot CT\text{-}F \\ 0.564 + 1.336 \cdot CT\text{-}G \end{cases}$$
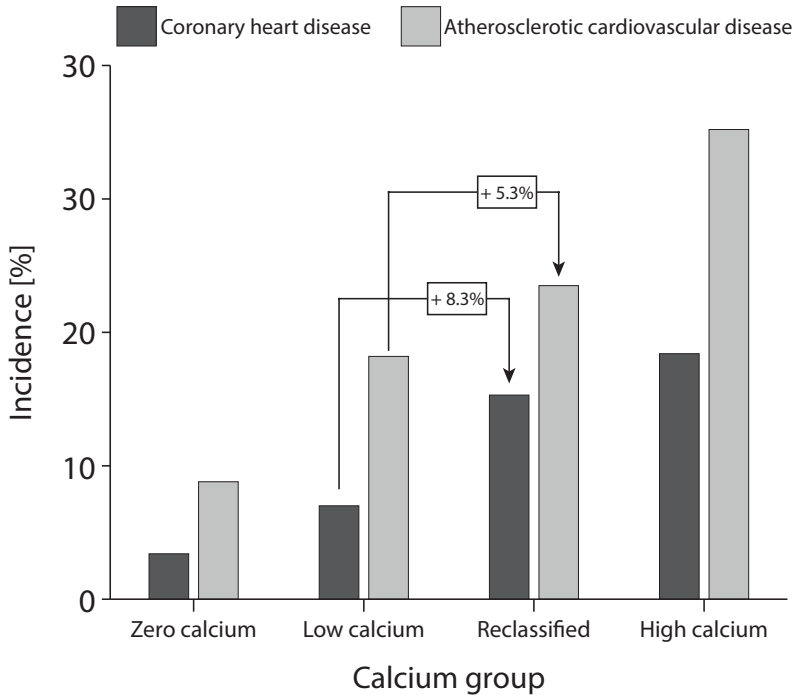
15

Figure 5 Coronary heart disease (CHD) and atherosclerotic cardiovascular disease (ASCVD) event rates during 18 years of follow up as per recommended by ESC and ACC for all calcium groups: zero calcium (AS = 0), low calcium (AS = 1 – 99, vnAS = 1 - 99), reclassified individuals (AS = 1 – 99; vnAS ≥ 100), high calcium group (AS ≥ 300). Reclassified individuals experienced increased CHD (+ 8.3%, P = 0.008) and ASCVD (+ 5.3%, P = 0.24) event rates in comparison with the baseline low calcium group

in CHD event rates between the reclassified and low calcium groups (p = 0.004), while there was no substantial difference in event rate of the reclassified group with the high coronary calcium group (P = 0.319) (*Figure 6*). For ASCVD, the opposite was true, with non-significant differences between reclassified individuals and the low coronary calcium group (P = 0.116), but significant differences of the reclassified group with the high coronary calcium group (P = 0.025).

Based on multivariable Cox regression, the hazard ratio of CHD for the reclassified individuals was 3.39 (95% CI 1.82 – 6.35, p = 0.001) (*Figure 7* and *Table 7*). The hazard ratio of CVD for reclassified individuals was 1.97 (95% CI 1.22, 3.18) (*Supplementary figure 1* and *Supplementary table 1*). The univariable Cox regression is depicted in *Supplementary table 2*.

Calcium	Coronary heart disease				Cardiovascular disease			
group	0	5	10	15	0	5	10	15
Zero	1576	1479	1352	1164	1576	1464	1328	1127
Low	696	632	552	460	696	612	525	429
Reclassified	85	75	61	44	85	71	54	38
High	824	691	539	348	824	639	471	406
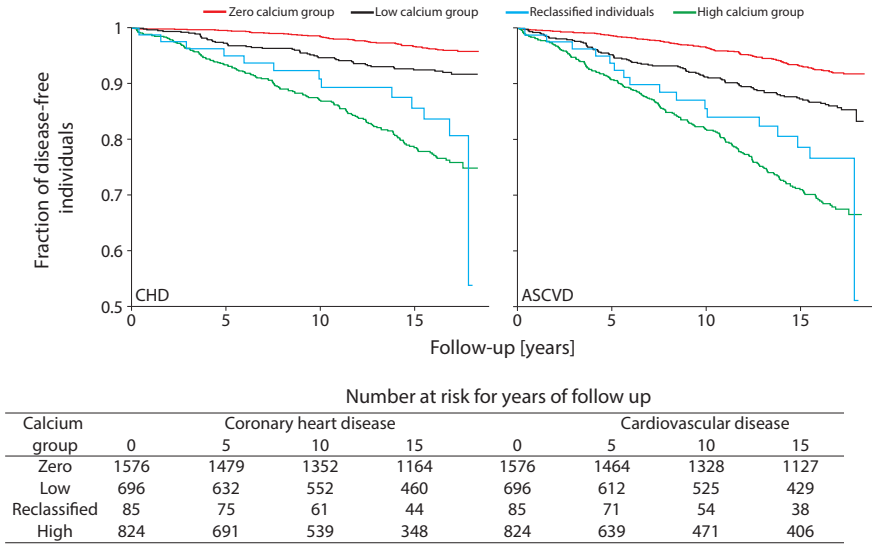
Figure 6 Kaplan-Meier survival estimates for all coronary calcium groups, for both coronary heart disease (CHD) (upper left) and atherosclerotic cardiovascular disease (ASCVD) (upper right). Total number of participants at risk are indicated in the table (below) for each coronary calcium group for 0, 5, 10, and 15 years of follow-up

Figure 7 Summary of the Cox regression model for the prediction of coronary heart disease events. The red dots depict a significant association between a variable and an event.

15

Table 7 Hazard ratio (95% confidence interval (CI)) examining the likelihood of coronary heart disease. The proportional multi-variable Cox regression was adjusted for all variables

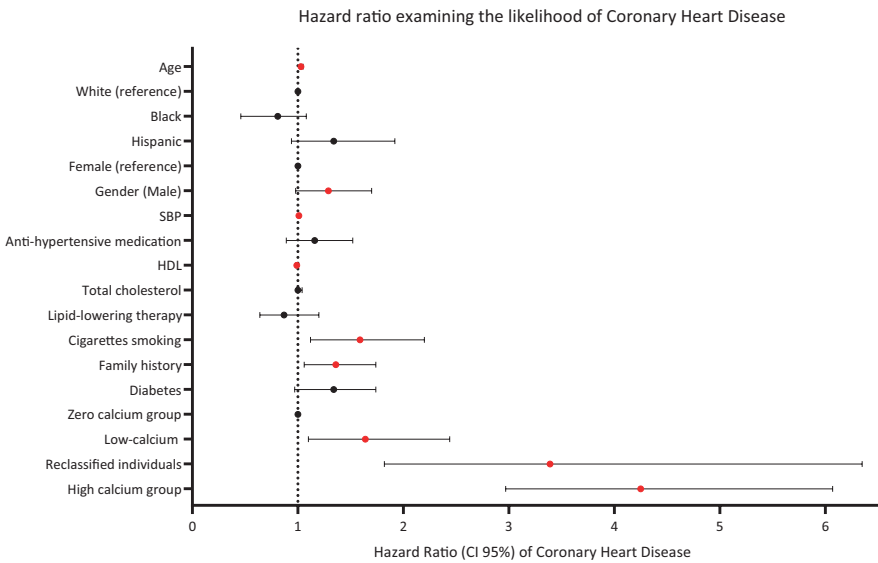|	Hazard ratio (95% CI)	P-value
Race		
White	Reference	
Black	1.08 (0.81, 1.46)	0.586
Hispanic	1.34 (0.94, 1.92)	0.110
Sex		
Female	Reference	
Male	1.29 (0.98, 1.70)	0.077
Age	1.03 (1.02., 1.05)	< 0.001
Anti-hypertensive medication		
Yes	1.16 (0.89, 1.52)	0.249
Systolic blood pressure	1.01 (1.01, 1.02)	<0.001
HDL	0.99 (0.98, 0.99)	0.014
Total cholesterol	1.00 (0.99, 1.04)	0.954
Lipid-lowering therapy		
Yes	0.87 (0.64, 1.20)	0.385
Cigarette smoking		
Yes	1.59 (1.12, 2.20)	0.008
Family history		
Yes	1.36 (1.06, 1.74)	0.014
Diabetes		
Yes	1.34 (0.97, 1.84)	0.074
Coronary calcium group		
Zero calcium	Reference	
Low calcium	1.64 (1.10, 2.44)	0.014
Reclassified individuals	3.39 (1.82, 6.35)	< 0.001
High calcium	4.25 (2.97, 6.07)	< 0.001

HDL = high-density lipoprotein

Statin therapy assessment: cohort 2

The mean age of cohort 2 participants (n = 889) was 64 ± 7 years, 47% were women, and 51% of participants were of non-white ethnicity (*Table 5*). Our vnAS calculation increased AS for the used CT systems, which resulted in 30 (11%) and 34 (28%) reclassified individuals from the baseline group 1 and 2 into reclassified group 1 and 2, respectively (*Table 3*). For both groups of reclassified individuals, ASCVD event rates were higher (*Table 3*).

In comparison with their baseline group, the statin therapy NNT reduced by n = 5 and n = 13 for reclassified groups 1 and 2, respectively (*Figure 8*).
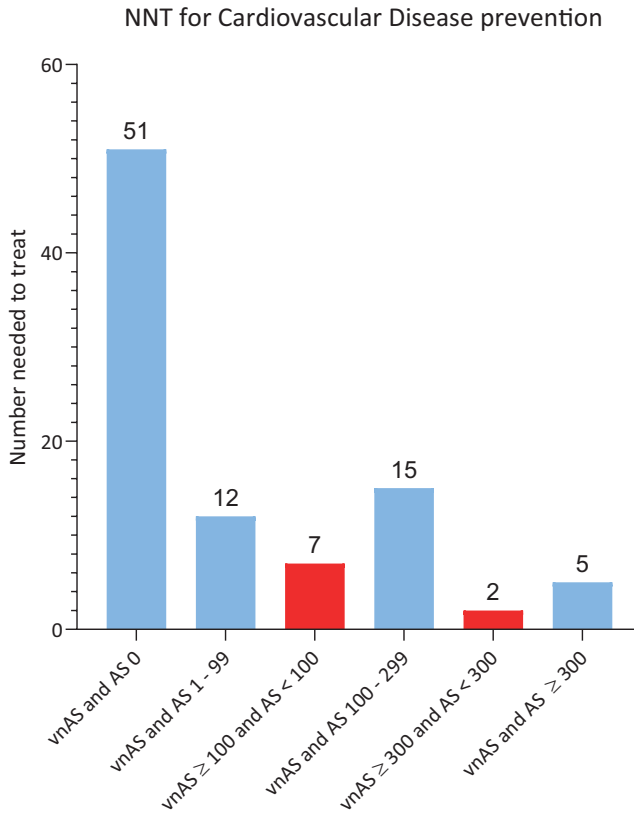
15

NNT for Cardiovascular Disease prevention

Figure 8 The estimated benefit of statins in primary prevention by AS and vnAS in intermediate risk group individuals from Multi-Ethnic study on Atherosclerosis. The ASCVD risk was calculated based on 10-year risk calculator. The 16-year number needed to treat (NNT) show that in reclassified groups the NNT was lower as compared to baseline groups

DISCUSSION

We developed a calibration tool which allows for vendor-neutral Agatston score (vnAS) calculation. Based on vnAS, MESA participants who were reclassified from a low to a high calcium group (11%) experienced more CHD and CVD events as compared to participants from the baseline group. This indicated improved risk categorization with vnAS. Moreover, vnAS also improved statin therapy classification for patients at intermediate ASCVD risk, with reduced NNT. Therefore, our vendor-neutral CAC assessment may improve patient risk classification and subsequent patient management.

15

This study showed large differences in AS between different CT manufacturers, but also between different CT systems from the same manufacturer. To overcome these discrepancies, we created a vendor-neutral calibration tool, which converts the vendor-specific AS to the gold-standard EBT AS. This calibration tool was validated in a large, long-term follow-up cohort. Since the advent of the AS in 1990, for which EBT images were used, CT technology has dramatically changed. Despite the present-day use of advanced MDCT systems in clinical practice, current clinical guidelines and risk calculators are based on studies, which almost exclusively used EBT systems to study the relation between the presence and amount of coronary calcium and subsequent cardiovascular events. This approach was correct from a methodological point of view, with EBT serving as the gold standard. However, it does thus not reflect current clinical practice.

Other studies also assessed the vendor dependency of AS, with contradictory findings. Mao et al. showed that the difference between AS on EBT and MDCT was about 8.3% for 102 patients, and was considered clinically insignificant.[24] In addition, Ghadri et al. showed that there were no substantial differences in CAC scores acquired on single source and dual source CT scanners.[25] Importantly, neither of these studies included information about patient's follow-up. On the other hand, in a study performed by Willemink and colleagues it was shown that differences in AS acquired with different vendors resulted in risk category reclassifications in up to 6.5% of individuals.[17] This is in agreement with our phantom analysis, in which we compared EBT to both older and state-of-the-art MDCT systems, and showed underestimations in AS of up to 48% on CT with respect to EBT. Importantly, in our study we were able to use the exact same phantom setup for both EBT and MDCT, which allowed for a direct AS comparison.

Nowadays, both American and European guidelines apply an AS threshold of 100 for consideration of pharmacological treatment. Therefore, to assess whether the vnAS outperforms AS for risk classifications, we also applied this commonly used AS threshold of 100. With this threshold, 85 (11%) MESA participants were reclassified from the low to the high calcium group. Out of these reclassified individuals, 15.3% and 23.5% experienced a CHD or ASCVD event, respectively. This was equal to an increase of 8.3% and 5.3% compared to the low calcium group, respectively, to which they were originally assigned based on the vendor specific AS. As further depicted in our Kaplan-Meier curves, the difference was only significant for CHD, but not for ASCVD events. The hazard ratio of CHD increased for reclassified individuals in comparison with individuals from the low calcium

15

group. For ASCVD events, this was less pronounced. The stronger association of vnAS with CHD in comparison with ASCVD might be explained by the fact that coronary calcium directly reflects atherosclerotic processes within coronary arteries. Therefore, the association of vnAS with CHD in a relatively small sample, will be more pronounced. This stronger association of CAC with CHD than with ASCVD was previously shown by Folsom et al.[26] Importantly, as the mean age of reclassified individuals was 67, this difference may have long-term consequences.

As previously indicated, independent studies showed that CAC quantification improved risk classification of asymptomatic individuals at risk by 14-30%.[16] Therefore, the American Heart Association recommends CAC measurement in asymptomatic individuals at intermediate ASCVD risk to plan lipid-lowering therapy.[13] To investigate the influence of vnAS on statin therapy, we used a sub-cohort of MESA participants which met the criteria of intermediate ASCVD risk group patients, designated as cohort 2.[11,13] As indicated by current guidelines, initiation of statin therapy is recommended in individuals with AS > 100, and for those with AS > 300 an intensive statin therapy is recommended.[11–13] In our sample, 30 (11%) participants were reclassified from the non-statin therapy group into the statin therapy group, and 34 (28%) participants were reclassified from the statin therapy group to the intensive statin therapy group. The NNT for both reclassified groups dropped as compared to the groups to which they were classified based on AS. This further indicates that for reclassified individuals, (intensive) statin therapy would be more efficient. Therefore, vnAS might be a valuable tool allowing for more appropriate treatment decision making for patients at intermediate ASCVD risk.

CAC evaluation is not only important for initial risk stratification, but also in follow-up analysis as indicated by Lehmann and colleagues, who showed that CAC progression was more pronounced in patients with CHD events than in those without events.[14] Patients with baseline and follow-up AS below 400 had 9.1% risk of ASCVD events within 10 years, while for those with baseline AS below 400 but follow-up AS exceeding 400, the ASCVD risk was 19.1%. These results served the National Lipid Association to recommend repeated CAC measurements depending on ASCVD risk.[11] However, in the abovementioned study by Lehmann et al., both scans were performed on the same EBT system.[14] In clinical practice, it is almost impossible to scan a patient every time on the same CT scanner. Knowing that the AS might differ up to 48% between different vendors, the importance of vnAS is self-evident.

15

We would like to underscore that our goal was not to develop a new calcium scoring method, but to provide a method which enables clinicians to calculate a vendor-neutral AS that closely reflects the EBT calcium score, based on which guidelines and risk calculators have been created. This approach can be implemented easily in existing workflows and leverages the strengths of the well-validated AS, while simultaneously helping to further improve patients risk classification.

This study has limitations that merit consideration. First, our vnAS is based on static anthropomorphic phantom scans only. Despite the fact that the linear attenuation coefficients of the phantoms were in line with human materials, a phantom does not completely simulate an actual human, with all internal organs, patient-specific variations, or coronary artery motion. Further, large-scale, in-vivo validation of the vnAS is therefore recommended in advance of wide-spread clinical acceptance. Second, for the vnAS validation we utilized older MDCT systems, as these were used to scan participants of the MESA. Other cohort studies which both employed state-of-the-art MDCT and had long-term follow-up, are not available yet. However, the MESA cohort used in our study did employ two MDCT systems from different manufacturers. Third, as a consequence of 50% of participants exhibiting a zero calcium score, the absolute number of reclassified MESA participants was relatively small at 2.7% for the first cohort. On the other hand, these patients constituted 11% of the low risk individuals group. In addition to the large and rising number of CT CAC assessments, this reclassification percentage is clinically relevant.

In conclusion, we developed a calibration tool which enables to calculate a vendor-neutral Agatston score which was validated in the MESA cohort. Based on the vnAS, reclassified individuals experienced increased CHD event rates in comparison with the baseline group. Moreover, the number needed to treat for statin therapy was reduced by using vnAS for MESA participants at intermediate cardiovascular risk. Therefore, our calibration tool for modern CT systems, if applied in daily clinical practice, may improve patient management and outcome.

Acknowledgements

REFERENCES

1. Oei HHS, Vliegenthart R, Hak AE, et al. The association between coronary calcification assessed by electron beam computed tomography and measures of extracoronary atherosclerosis: The rotterdam coronary calcification study. J Am Coll Cardiol. 2002;39(11):1745-1751. doi:10.1016/S0735-1097(02)01853-3

2. Bild DE, Bluemke DA, Burke GL, et al. Multi-Ethnic Study of Atherosclerosis: Objectives and Design. American Journal of Epidemiology. 2002;156(9):871-881. doi:10.1093/aje/kwf113

3. Schmermund A, Möhlenkamp S, Stang A, et al. Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged subjects: Rationale and design of the Heinz Nixdorf RECALL Study. American Heart Journal. 2002;144(2):212-218. doi:10.1067/MHJ.2002.123579

4. Demer LL, Tintut Y. Vascular Calcification Pathobiology of a Multifaceted Disease. Published online 2008. doi:10.1161/CIRCULATIONAHA.107.743161

5. Budoff MJ, Young R, Burke G, et al. Ten-year association of coronary artery calcium with atherosclerotic cardiovascular disease (ASCVD) events: the multi-ethnic study of atherosclerosis (MESA). doi:10.1093/eurheartj/ehy343

6. McClelland RL, Jorgensen NW, Budoff M, et al. 10-Year Coronary Heart Disease Risk Prediction Using Coronary Artery Calcium and Traditional Risk Factors: Derivation in the MESA (Multi-Ethnic Study of Atherosclerosis) With Validation in the HNR (Heinz Nixdorf Recall) Study and the DHS (Dallas Heart Study). J Am Coll Cardiol. 2015;66(15):1643-1653. doi:10.1016/J.JACC.2015.08.035

7. Yeboah J, McClelland RL, Polonsky TS, et al. Comparison of novel risk markers for improvement in cardiovascular risk assessment in intermediate-risk individuals. JAMA. 2012;308(8):788-795. doi:10.1001/jama.2012.9624

8. Dzaye O, Razavi AC, Dardari ZA, et al. Mean Versus Peak Coronary Calcium Density on Non-Contrast CT: Calcium Scoring and ASCVD Risk Prediction. JACC Cardiovasc Imaging. 2022;15(3):489-500. doi:10.1016/J.JCMG.2021.09.018

9. Greenland P, Blaha MJ, Budoff MJ, Erbel R, Watson KE. Coronary Calcium Score and Cardiovascular Risk. J Am Coll Cardiol. 2018;72(4):434-447. doi:10.1016/j.jacc.2018.05.027

10. Mitchell JD, Fergestrom N, Gage BF, et al. Impact of Statins on Cardiovascular Outcomes Following Coronary Artery Calcium Scoring. J Am Coll Cardiol. 2018;72(25):3233-3242. doi:10.1016/j.jacc.2018.09.051

11. Orringer CE, Blaha MJ, Blankstein R, et al. The National Lipid Association scientific statement on coronary artery calcium scoring to guide preventive strategies for ASCVD risk reduction. Journal of Clinical Lipidology. 2021;15(1):33-60. doi:10.1016/j.jacl.2020.12.005

12. Mach F, Baigent C, Catapano AL, et al. 2019 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce cardiovascular risk: The Task Force for the management of dyslipidaemias of the European Society of Cardiology (ESC) and European Atherosclerosis Society (EAS). European Heart Journal. 2020;41(1):111-188. doi:10.1093/eurheartj/ehz455

15

13. Grundy SM, Stone NJ, Bailey AL, et al. 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA Guideline on the Management of Blood Cholesterol: Executive Summary: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. J Am Coll Cardiol. 2019;73(24):3168-3209. doi:10.1016/J.JACC.2018.11.002

14. Lehmann N, Erbel R, Mahabadi AA, et al. Value of Progression of Coronary Artery Calcification for Risk Prediction of Coronary and Cardiovascular Events. Circulation. 2018;137(7):665-679. doi:10.1161/CIRCULATIONAHA.116.027034

15. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. J Am Coll Cardiol. 1990;15(4):827-832. doi:10.1016/0735-1097(90)90282-T

16. Erbel R, Möhlenkamp S, Moebus S, et al. Coronary Risk Stratification, Discrimination, and Reclassification Improvement Based on Quantification of Subclinical Coronary Atherosclerosis. J Am Coll Cardiol. 2010;56(17):1397-1406. doi:10.1016/j.jacc.2010.06.030

17. Willemink MJ, Vliegenthart R, Takx RAP, et al. Coronary Artery Calcification Scoring with State-of-the-Art CT Scanners from Different Vendors Has Substantial Effect on Risk Classification. Radiology. 2014;273(3):695-702. doi:10.1148/radiol.14140066

18. McCollough CH, Ulzheimer S, Halliburton SS, Shanneik K, White RD, Kalender WA. Coronary Artery Calcium: A Multi-institutional, Multimanufacturer International Standard for Quantification at Cardiac CT. Radiology. 2007;243(2):527-538. doi:10.1148/radiol.2432050808

19. Carr JJ, Nelson JC, Wong ND, et al. Calcified coronary artery plaque measurement with cardiac CT in population-based studies: standardized protocol of Multi-Ethnic Study of Atherosclerosis (MESA) and Coronary Artery Risk Development in Young Adults (CARDIA) study. Radiology. 2005;234(1):35-43. doi:10.1148/RADIOL.2341040439

20. Praagh GD, Werf NR, Wang J, et al. Fully automated quantification method (FQM) of coronary calcium in an anthropomorphic phantom. Medical Physics. 2021;48(7):3730-3740. doi:10.1002/mp.14912

21. Muntner P, Colantonio LD, Cushman M, et al. Validation of the atherosclerotic cardiovascular disease Pooled Cohort risk equations. JAMA. 2014;311(14):1406-1415. doi:10.1001/JAMA.2014.2630

22. Adèr H, Mellenbergh G. Research Methodology in the Social, Behavioural and Life Sciences. SAGE Publications, Ltd; 1999. doi:10.4135/9780857029027

23. Altman DG, Andersen PK. Calculating the number needed to treat for trials where the outcome is time to an event. BMJ. 1999;319(7223):1492-1495. doi:10.1136/bmj.319.7223.1492

24. Mao SS, Pal RS, McKay CR, et al. Comparison of Coronary Artery Calcium Scores Between Electron Beam Computed Tomography and 64-Multidetector Computed Tomographic Scanner. Journal of Computer Assisted Tomography. 2009;33(2):175-178. doi:10.1097/RCT.0b013e31817579ee

25. Ghadri JR, Goetti R, Fiechter M, et al. Inter-scan variability of coronary artery calcium scoring assessed on 64-multidetector computed tomography vs. dual-source computed tomography: A head-to-head comparison. European Heart Journal. 2011;32(15):1865-1874. doi:10.1093/eurheartj/ehr157

15

26. Folsom AR, Kronmal RA, Detrano RC, et al. Coronary Artery Calcification Compared With Carotid Intima-Media Thickness in the Prediction of Cardiovascular Disease Incidence The Multi-Ethnic Study of Atherosclerosis (MESA). Accessed April 22, 2022. www.jamaarchivescme.com

15

SUPPLEMENTAL MATERIAL

Supplementary table 1 Hazard ratio examining the likelihood of atherosclerotic cardiovascular disease events. The proportional multi-variable Cox regression was adjusted for all variables

	Hazard ratio (95% CI)	P-value
Race		
White	Reference	
Black	1.07 (0.88, 1.31)	0.501
Hispanic	1.28 (0.99, 1.66)	0.051
Sex		
Female	Reference	
Male	1.32 (0.99, 1.54)	0.004
Age	1.04 (1.03, 1.05)	< 0.001
Anti-hypertensive medication		
Yes	1.24 (1.03, 1.48)	0.019
Systolic blood pressure	1.01 (1.01, 1.02)	< 0.001
HDL	0.98 (0.98, 0.99)	< 0.001
Total cholesterol	1.00 (0.99, 1.01)	0.115
Lipid-lowering therapy		
Yes	1.05 (0.85, 1.28)	0.658
Cigarette smoking		
Yes	1.55 (1.23, 1.95)	< 0.001
Family history		
Yes	1.24 (1.02, 1.51)	0.027
Diabetes		
Yes	1.44 (1.16, 1.79)	< 0.001
Coronary calcium group		
Zero calcium	Reference	
Low calcium	1.63 (1.27, 2.09)	< 0.001
Reclassified individuals	1.97 (1.22, 3.18)	0.006
High calcium	3.14 (2.48, 3.97)	< 0.001

15

Supplementary table 2 Hazard ratio examining the likelihood of coronary heart disease (CHD) and atherosclerotic cardiovascular disease (ASCVD) events. The proportional uni-variable Cox regression was adjusted for all variables

	CHD		ASCVD	
	Hazard ratio (95% CI)	**P-value**	**Hazard ratio (95% CI)**	**P-value**
Race				
White	Reference		Reference	
Black	0.97 (0.74, 1.28)	0.840	0.98 (0.81, 1.18)	0.857
Hispanic	1.06 (0.75, 1.51)	0.735	1.02 (0.80, 1.29)	0.882
Sex				
Female	Reference		Reference	
Male	1.86 (1.45, 2.38)	0.000	1.7 (1.44, 2.01)	< 0.001
Age	1.05 (1.04, 1.07)	0.000	1.06 (1.05, 1.07)	< 0.001
Anti-hypertensive medication				
Yes	1.77 (1.39, 2.25)	0.000	1.87 (1.58, 2.21)	< 0.001
Systolic blood pressure	1.02 (1.01, 1.02)	0.000	1.02 (1.01, 1.02)	< 0.001
HDL	0.98 (0.97, 0.99)	0.000	0.98 (0.97, 0.99)	< 0.001
Total cholesterol	0.99 (0.99, 1.01)	0.133	0.99 (0.99, 1.01)	0.296
Lipid-lowering therapy				
Yes	1.18 (0.88, 1.59)	0.272	1.39 (1.14, 1.69)	< 0.001
Cigarette smoking				
Yes	1.26 (0.92, 1.74)	0.148	1.18 (0.95, 1.48)	0.134
Family history				
Yes	1.57 (1.24, 2.00)	0.000	1.43 (1.22, 1.69)	< 0.001
Diabetes				
Yes	2.05 (1.51, 2.77)	0.000	2.1 (1.71, 2.59)	< 0.001
Coronary calcium group				
Zero calcium	Reference		Reference	
Low calcium	2.19 (1.49, 3.22)	0.000	2.27 (1.78, 2.89)	< 0.001
Reclassified individuals	5.32 (2.91, 9.75)	0.000	3.38 (2.11, 5.41)	< 0.001
High calcium	7.01 (5.13, 9.56)	0.000	5.61 (4.57, 6.87)	< 0.001

15
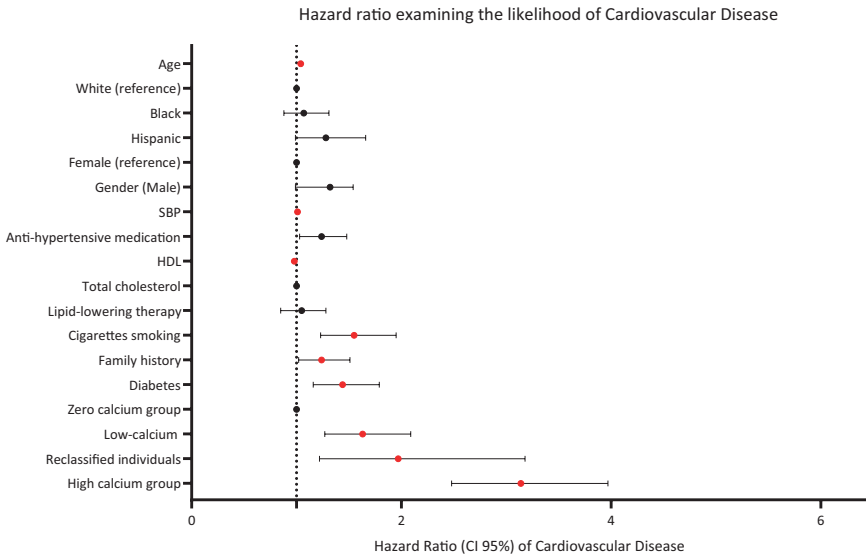
Supplementary figure 1 Summary of the Cox regression model for the prediction of atherosclerotic cardiovascular disease events. The red dots depict a significant association between a variable and an event.

CHAPTER 16

Fully automated Quantification Method (FQM) of coronary calcium in an anthropomorphic phantom

Gijs D. van Praagh*, MSc

Niels R. van der Werf*, MSc

Jia Wang, PhD

Fasco van Ommen, PhD

Keris Poelhekken, BSc

Riemer H.J.A. Slart, MD PhD

Dominik Fleischmann, MD

Marcel J.W. Greuter, PhD

Tim Leiner, PhD

Martin J. Willemink, MD PhD

* Shared first authors

ABSTRACT

Objective

Coronary artery calcium (CAC) score is a strong predictor for future adverse cardiovascular events. Anthropomorphic phantoms are often used for CAC studies on computed tomography (CT) to allow for evaluation or variation of scanning or reconstruction parameters within or across scanners against a reference standard. This often results in large number of datasets. Manual assessment of these large datasets is time consuming and cumbersome. Therefore, this study aimed to develop and validate a fully automated, open-source quantification method (FQM) for coronary calcium in a standardized phantom.

Materials and Methods

A standard, commercially available anthropomorphic thorax phantom was used with an insert containing nine calcifications with different sizes and densities. To simulate two different patient sizes, an extension ring was used. Image data was acquired with four state-of-the-art CT systems using routine CAC scoring acquisition protocols. For inter-scan variability, each acquisition was repeated five times with small translations and/or rotations. Vendor-specific CAC scores (Agatston, volume, and mass) were calculated as reference scores using vendor-specific software. Both the international standard CAC quantification methods as well as vendor-specific adjustments were implemented in FQM. Reference and FQM scores were compared using Bland-Altman analysis, intraclass correlation coefficients, risk reclassifications, and Cohen's kappa. Also, robustness of FQM was assessed using varied acquisitions and reconstruction settings and validation on a dynamic phantom. Further, image quality metrics were implemented: noise power spectrum, task transfer function, and contrast- and signal-to-noise ratio among others. Results were validated using imQuest software.

Results

Three parameters in CAC scoring methods varied among the different vendor-specific software packages: the Hounsfield unit (HU) threshold, the minimum area used to designate a group of voxels as calcium, and the usage of isotropic voxels for the volume score. The FQM was in high agreement with vendor-specific scores and ICC's (median [95% CI]) were excellent (1.000 [0.999-1.000] to 1.000

16

[1.000-1.000]). An excellent inter-platform reliability of κ = 0.969 and κ = 0.973 was found. TTF results gave a maximum deviation of 3.8% and NPS results were comparable to imQuest.

Conclusions

We developed a fully automated, open-source, robust method to quantify CAC on CT scans in a commercially available phantom. Also, the automated algorithm contains image quality assessment for fast comparison of differences in acquisition and reconstruction parameters.

16

INTRODUCTION

Coronary artery calcium (CAC) score is a strong predictive value for future adverse cardiovascular events, including myocardial infarction and sudden cardiac death, and a powerful tool in primary prevention.[1–3] In 1990, Agatston and colleagues developed a specific quantification method for CAC using electron beam tomography (EBT).[4] This so-called Agatston score – currently quantified using cardiac computed tomography (CT) – is clinically used for further risk classification of asymptomatic individuals at intermediate risk.[5–7] In addition to the Agatston score, two other metrics were introduced to quantify CAC, namely the volume and mass score.[8,9]

It is well known that CAC scores vary between different CT scanners. Not only do CAC scores differ between scanners of different vendors, but also between different systems from the same vendor, and between the same systems from the same vendor if a slightly different starting position is applied.[10–12] Moreover, CAC scores can vary greatly due to motion of the coronary arteries during the scan phase of a CAC scoring CT acquisition.[13] In order to study these differences, their possible impact on clinical outcome, and to optimize acquisition protocols, dedicated coronary calcium phantoms are frequently used. In the well-established international standard developed for CAC quantification by McCollough and colleagues, a commonly evaluated commercially available anthropomorphic phantom was used (thorax and CCI phantom, QRM, Möhrendorf, Germany).[8] With this phantom not only the Agatston score, but also the volume and mass score of the calcifications in the phantom can be studied among different scanners and different vendors for influences of acquisition and reconstruction parameters. This phantom also contains calibration rods, allowing for adequate mass score assessment.

However, manual assessment of the CAC scores is time consuming and cumbersome, especially when several scan-and/or reconstruction parameters have been systematically varied resulting in a large number of scans. Therefore, the aim of our study was to develop and validate a fully automated quantification method (FQM) for coronary calcium in a standardized phantom. In order to be useful for a variety of CT scanners of different vendors, we sought to develop an automated scoring method that replicates CAC scores.

16

METHODS

Phantom

We used a standard, commercially available anthropomorphic thorax phantom (QRM-Thorax, QRM, Möhrendorf, Germany) (*Figure 1a*). The static phantom comprises artificial lungs, a spine, and a shell of soft tissue equivalent material. X-ray attenuation of the phantom's materials is similar to human tissues when data are acquired at a peak tube potential of 120 kVp. To simulate two different patient sizes, an extension ring (QRM-Extension ring, QRM, Möhrendorf, Germany) of fat equivalent material (-100 Hounsfield Units (HU)) was used. With this extension ring, outer dimensions of the phantom increased from 300 x 200 mm to 400 x 300 mm, similar to a small and large patient size, respectively.[14] Within the thorax, a commercially available calcium containing insert (Cardiac Calcification Insert (CCI), QRM, Möhrendorf, Germany) was placed, which is commonly used in coronary calcium studies (*Figure 1b*).[8,15–20] The insert consisted of nine hydroxyapatite (HA) containing calcifications and two large calibration rods. These calibration rods consisted of water-equivalent material and 200 mgHAcm^{-3}. The calcifications had diameters and lengths of 1.0, 3.0, and 5.0 mm, defined as small, medium, and large, respectively. For each calcification size, three densities were present in the phantom: 200, 400, and 800 mgHAcm^{-3}, defined as low, medium, and high density, respectively.

To assess the performance of our automatic scoring method on dynamic data, a robotic arm (QRM Sim2D, QRM, Möhrendorf, Germany) moved an artificial coronary artery in a water-filled compartment, which was positioned in the center of the anthropomorphic thorax phantom. Two artificial arteries were used, where each artery consisted of two calcifications. These calcifications were equal in dimensions (5.0±0.1 mm in diameter, with a length of 10.0±0.1 mm), but different in density. Densities were 196±3, 380±2, 408±2, and 800±2 mgHAcm^{-3}. The arteries were moved at four constant velocities (0 – 30 mm/s, increment of 10 mm/s) along the x-axis, comparable to heart rates of 0, <60, 60-75, and >75 bpm.[21] Electrocardiography trigger output was used to ensure that acquisition was done during linear motion of the calcifications.[13]

Acquisition and Reconstruction

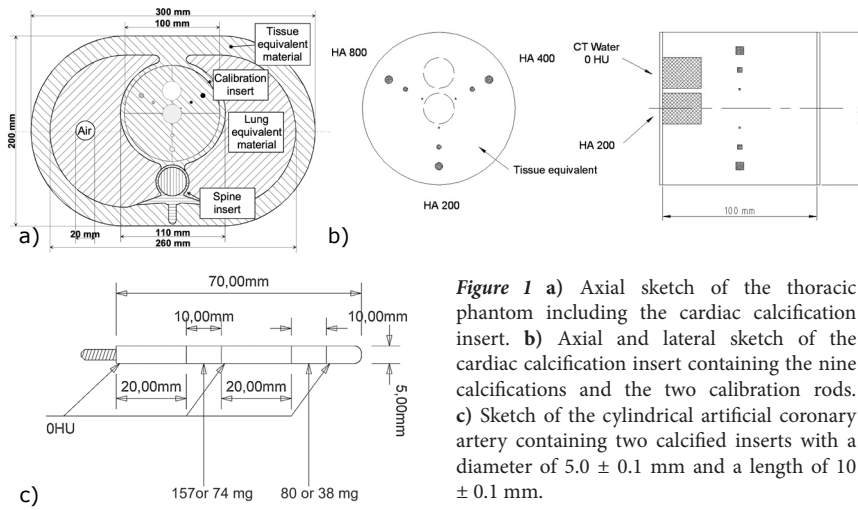Static phantom image data was acquired with four state-of-the-art CT systems,

16

Figure 1 **a)** Axial sketch of the thoracic phantom including the cardiac calcification insert. **b)** Axial and lateral sketch of the cardiac calcification insert containing the nine calcifications and the two calibration rods. **c)** Sketch of the cylindrical artificial coronary artery containing two calcified inserts with a diameter of 5.0 ± 0.1 mm and a length of 10 ± 0.1 mm.

Table 1 Acquisition and reconstruction parameters for all CT systems used in this study.

Parameter	CT1	CT2	CT3	CT4	Dynamic
Manufacturer	Canon	Philips	GE	Siemens	Siemens
CT system	Aquilion One Vision	Brilliance iCT	Revolution	SOMATOM Force	SOMATOM Flash
Acquisition mode	Axial	Axial	Axial	Axial	Axial
Tube voltage [kVp]	120	120	120	120	120
Tube current time product [mAs]	Small: 15 Large: 84	Small: 50 Large: 50	Small: 30 Large: 161	Small: 44 Large: 194	80
Automatic exposure correction	SD=55	Off	Off	Off	Off
CTDI$_{vol}$ [mGy]	Small: 2.3 Large: 12.8	Small: 4.7 Large: 4.4	Small: 1.49 Large: 7.2	Small: 1.5 Large: 6.7	Large: 2.8
Collimation [mm]	280x0.5	128x0.625	224x0.625	160x0.6	128x0.6
Field of View [mm]	250	250	250	250	250
Rotation time [s]	0.35	0.27	0.28	0.25	0.28
Slice thickness [mm]	3.0	3.0	2.5	3.0	3.0
Increment [mm]	3.0	3.0	2.5	3.0	3.0
Reconstruction kernel	FC12	XCA	Standard	Qr36d	B35f*
Matrix size [pixels]	512x512	512x512	512x512	512x512	512x512
Reconstruction	FBP	FBP	FBP	FBP	FBP
Calcium scoring software	Vitrea FX 6.5.0 (**S1**)	Heartbeat-CS (**S2**)	SmartScore 4.0 (**S3**)	Syngo Calcium Scoring (**S4**)	Syngo Calcium Scoring (**S4**)

*Based on vendor recommended protocol of earlier software version than used for the static phantom

one from each of the main CT manufacturers: CT-1: Aquilion One Vision (Canon Medical Systems, Otawara, Japan); CT-2: Brilliance iCT (Philips Healthcare, Best, The Netherlands); CT-3: Revolution CT (GE Healthcare, Waukesha, Wisconsin, USA); and CT-4: SOMATOM Force (Siemens Healthineers, Erlangen, Germany), respectively. Routine CAC scoring acquisition protocols for small and large patients were used (*Table 1*). To simulate inter-scan variability each acquisition of the thorax phantom with and without extension ring was done five times with small translations and/or rotations of approximately 2 mm and 2 degrees, respectively. Raw data were reconstructed with filtered back projection (FBP) (*Table 1*).

Vendor-specific CAC scores

For all acquisitions, vendor-specific CAC scores were derived using each vendor's commercial software implementation (*Table 1*). These CAC scores included Agatston, volume, and mass scores. For each vendor, CAC scores derived with their respective software were used as reference CAC scores for the analysis. The CT specific mass calibration factor was determined for each CT system according to standard methodology.[8]

CAC score standard: automated algorithms

The international standard for quantification of CAC scores was implemented in a fully automated algorithm (FQM) for CAC scoring of the CCI phantom. This was done in two popular programming languages to allow for wide usage: MATLAB® R2020a (Mathworks, Natick, Massachusetts, USA) and Python (Python 3.7.3). Both algorithms were made publicly available via Github (https://github.com/nwerf/FQM_Analysis) to assist in any research where the CCI insert is used.

After importing a DICOM series into FQM (module 1), the center of the insert (module 2) and two main locations in the CCI were found: the largest calcifications (module 3) and the 200 mg HA calibration rod (module 4; *Figure 2*). These calcified areas were found using a connected component analysis (4-connected) with the standard CAC scoring threshold of 130 HU.[4] Next, a mask based on the locations of the nine calcifications was determined. First, the largest calcifications were defined based on the area of the connected components. For each density, the locations of the other calcifications were determined using the known distances between the calcifications of different sizes, on the connecting lines between the center of the insert and the center of the large calcification. The mean HU value of each of the large calcifications was used to determine the density of the calcifications, with
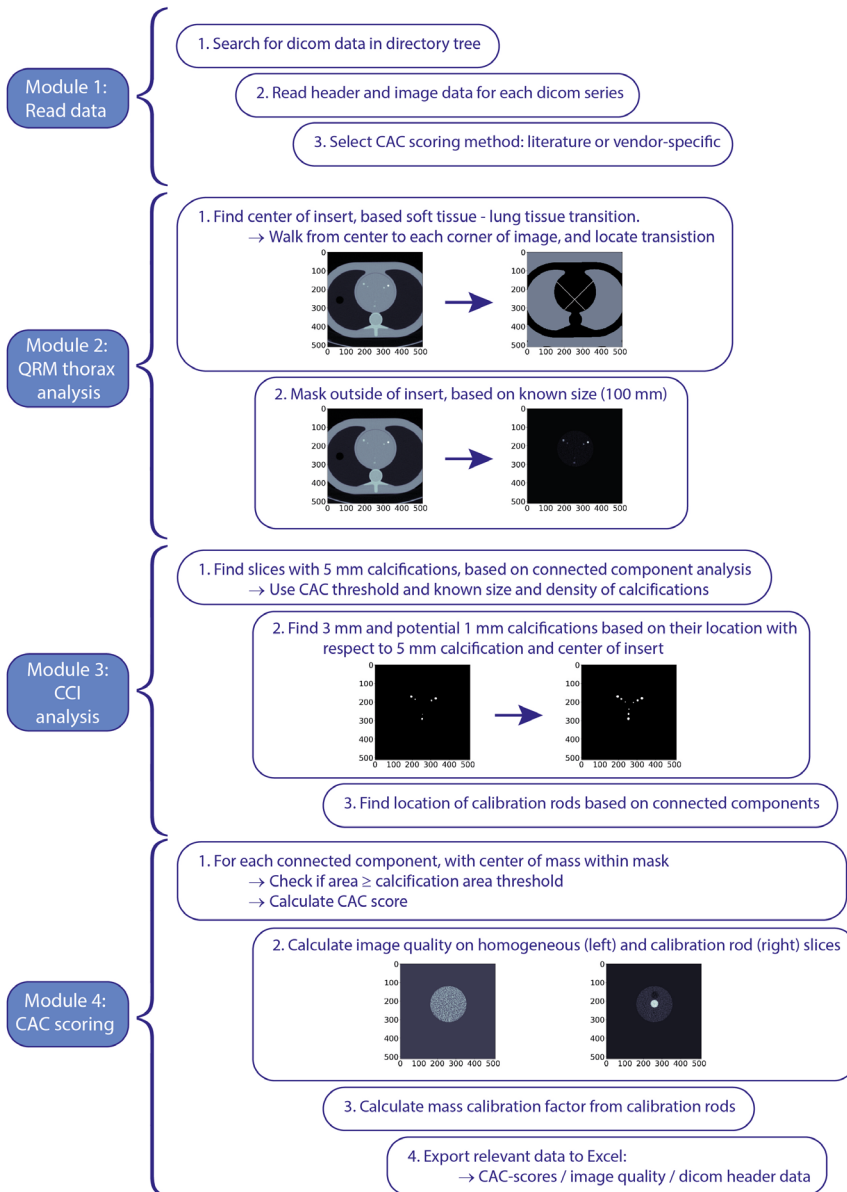
16

Figure 2 Flowchart of FQM.

the highest mean HU value corresponding to the highest density etc. By using this methodology, the exact position of the phantom within the CT system, and any rotation of the CCI insert within the thorax phantom, was made irrelevant, consequently adding to the robustness of FQM.

The international standard implementation for all three CAC scoring methods (Agatston, volume, and mass scores) were in accordance with their respective definitions from literature.[4,8,9] All methods used a minimum in-plane area of 1 mm^2 for pixels > 130 HU to identify calcium-containing lesions. The Agatston scores were derived for each calcified area per slice from a multiplication of that area with an associated weighting factor depending on the maximum HU within the area: 130 to 200 HU = 1; 200 to 300 HU = 2; 300 to 400 HU = 3; and ^{3} 400 HU = 4. The Agatston score per calcification was defined as the summation of all Agatston scores per slice.

The volume score was determined according to Callister et al., based on a linear interpolation to create isotropic voxels.[9] To achieve this, the slice thickness was decreased to match in-plane pixel spacing by means of a linear grid interpolation in 3D. To limit computation time, this was only performed for the slices containing the calcifications. For each slice, the volume score was calculated by multiplication of the number of voxels per lesion with the interpolated voxel volume.

Lastly, mass scores were determined according to McCollough et al., using scan specific mass calibration factors.[8] Mean CT numbers (HU) for the calibration factor calculation were measured in the center slice of the large cylinder-shaped calibration rods with a region of interest of 1.5 cm^2. The calibration rods were automatically located, based on the known specifications of the phantom. Then, mean CT numbers (HU) for both calibration rods were used to calculate the scan-specific calibration factor. Finally, mass scores of the calcifications were calculated by multiplication of the calibration factor with the calcified volume (without interpolation) and the mean CT number of the lesion.

To assess robustness of FQM, additional acquisitions with varying acquisition settings were made on CT-4 and scored with its vendor-specific software. In these acquisitions, parameters that have a well-known influence on CAC scores were changed: tube potential was changed from 120 to 100 and 80 kVp, tube current time product was changed from 44 to 34 and 22 mAs, convolution kernel was changed from Qr36 to Qr32 and Qr44, iterative reconstruction was applied at levels 2 and 4, and lastly, field-of-view was changed from 250 to 200 and 320 mm. Finally, robustness was assessed for a dynamic phantom on another CT system: SOMATOM Definition Flash (Siemens Healthineers, Erlangen, Germany). A routinely used clinical CT CAC protocol was used for acquisition and reconstruction (*Table 1*; *Figure 3*).
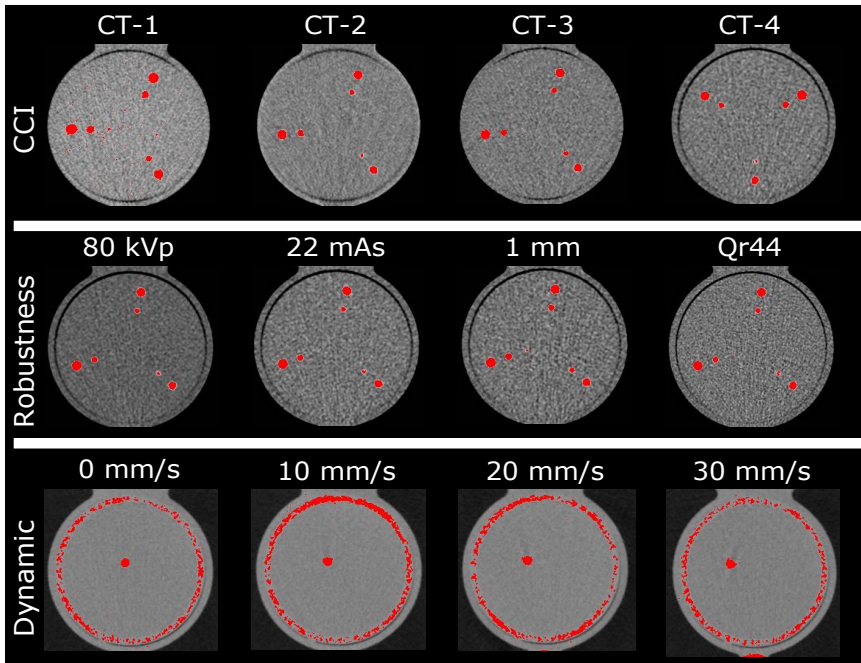
16

Figure 3 Axial views of the cardiac calcification insert from the four CT systems used in static experiments (top row), a few examples of the robustness scans where acquisition or reconstruction settings were changed (middle row; from left to right: tube voltage, tube current, slice thickness, and kernel), and the dynamic phantom with four different speed settings (bottom row). Red overlay is used to highlight the pixels above the 130 HU threshold. Screenshots are made with ImageJ (U.S. National Institutes of Health, Bethesda, Maryland, USA).

Vendor-specific CAC scores: automated algorithms

In addition, FQM was adapted in such a way that the calculation of the calcium scores matched the methodology used in the vendor-specific software packages. These adjustments were based on scoring mechanism descriptions in manuals, and information provided by the vendors. The following parameters were adapted: HU threshold used to designate a pixel as calcium, the threshold used to indicate the minimum area necessary for calcium scoring, and the use of interpolation for specific CAC scores (*Table 2*). Vendor-specific parameters were automatically extracted by FQM from the DICOM header information, which also identified the vendor-specific CT system that was used to acquire the data. In addition, the algorithms allowed for manual selection of vendor-specific scoring parameters. With this, images from any of the four vendors can be evaluated with scoring parameters from any other vendor.

16

Table 2 International standard and vendor-specific CAC scoring parameters for all vendors and commercial vendor-neutral software. Light-grey entries indicate equal parameter values with respect to literature. Darker-grey entries are vendor-specific parameters.

Parameters	CAC score	International standard	S1	S2	S3	S4
Connectivity	All	4	4	4	4	4
HU threshold	Agatston	130	130	130	130	130
	Volume	130	130	130 or 100/c[a]	Patented	130
	Mass	130	130	100/c	Patented	130
Calcification area threshold	All	1 mm^2	3 pixels	0.5 mm^2	1 mm^2	0
Interpolation[b]	Volume	Yes	No	No	Patented	Yes

c = calibration factor
[a] Depending on availability of CT system specific calibration factor within the scoring software
[b] Linear interpolation algorithm used to calculate isotropic voxels

Image quality assessment

For the automated analysis of the CCI phantom, several image quality metrics were included to assess image quality differences for changing acquisition or reconstruction parameters. These image quality metrics both concerned image noise and contrast measurement. For the image noise, first the standard deviation (SD) of the mean CT-value (HU) of a square region-of-interest (ROI) of 55 x 55 mm in a non-calcium containing slice of the CCI insert was calculated. Second, image noise was characterized with a noise power spectrum (NPS) analysis. This analysis was implemented according to the methodology of the International Commission on Radiation Units and Measurements (ICRU), as previously implemented by Van Ommen et al.[22,23] For this, 18 radially dispersed ROIs of 15 x 15 mm were used. Both 2D and 1D NPS results were extracted.

For the contrast-related image quality metrics, first the mean HU and SD of the three large calcifications and two calibration rods were calculated. For each calcification, the mean HU was calculated over the entire volume of each calcification. The mean HU and SD of a circular ROI of 1.5 cm^2 in the calibration rod were calculated within the center slice of these rods. Second, the signal-to-noise ratio (SNR) and contrast-to-noise ratio (CNR) were calculated for the calcifications.

Lastly, the task-transfer-function (TTF) was computed. The TTF is a type of modulation-transfer-function, which is also valid for non-linear systems and incorporates contrast and noise.[24] For this, the ICRU implementation for

16

modulation transfer function calculation was used.[22] For robustness, the TTF was calculated by radially averaging the edge-spread-function (ESF) of the calibration rod, as described previously by Van Ommen et al.[23] Due to the proximity of the water-equivalent calibration rod, the ESF in the direction of this rod were excluded from the analysis. In addition, image data were linearly interpolated by a factor four, to reduce pixel size effects. For quick evaluation purposes, 50% and 10% TTF were also calculated.

NPS and TTF results were validated by comparison with the CT image analysis tool (imQuest (Duke University, Durham, 2018)) described in Task Group 233 of the American Association of Physicists in Medicine (AAPM) for two datasets, reconstructed with different reconstruction kernels (Qr44, Qr32). For the NPS calculation, only one ROI was placed at the center of the insert for both tools for the current comparison, due to potential measurement errors resulting from manual placement of 18 ROIs for imQuest.

Statistical analysis

To assess the accuracy of our FQM, automatically quantified CAC scores were compared with reference scores obtained with vendor-specific software. Agreement between FQM and reference CAC scores was assessed using Bland-Altman analyses. Reliability between the methods was determined by calculating intraclass correlation coefficients (ICCs) and root mean square error (RMSE). Reference and FQM scores were classified per calcification according to the Agatston risk stratification: 0 – absent; > 0 and < 10 – minimal; ³ 10 and < 100 – mild; ³ 100 and < 400 – moderate; ³ 400 – severe. Calcifications classified differently by FQM from the reference classifications were defined as reclassifications. Subsequently, reliability of reclassification between FQM and reference scores was determined by calculating Cohen's kappa (κ). All statistical analyses were performed with SPSS for Windows, version 26.0. A p-value <0.05 was used to determine significant differences.

16

RESULTS

Vendor-specific CAC scores: automated algorithms

Vendor-specific adjustments to our generic CAC scoring methods were necessary to match vendor-specific scores. An overview of all parameters, including vendor-specific parameters, is shown in *Table 2*. Three parameters varied among the

different vendor-specific software packages. First, the HU threshold, used to indicate whether a pixel contains CAC, varied. In general, a threshold of 130 HU was used for all vendors, for all CAC scores. However, for one vendor the threshold was 100 mg HA, when a CT system specific calibration factor was available in the software. When this calibration factor was not available, the normal threshold of 130 HU was used. Second, the minimum area used to designate a group of pixels as calcium varied. For a group of pixels with HU above the CAC scoring threshold, this minimum area varied between >0 pixels and 1 mm^2. Last, some vendors used an interpolation algorithm to create isotropic voxels for the volume score, and some vendors did not. Parameters for the volume and mass score of CT-3 were kept confidential by the vendor and could therefore not be determined.

With these vendor-specific CAC scoring parameters implemented, FQM scores were in high agreement with the vendor-specific software scores for all CAC scoring methods (*Figure 4*). Smallest confidence interval (CI) (95%) range of absolute differences between the FQM and vendor-specific scores was 0.000 to 0.000 mg for the mass score when FQM was compared to S4. Largest CI range was -2.480 to 1.827 mm^3 for the volume score when FQM was compared to S4. ICCs were excellent for all comparisons between FQM and the vendor-specific software. The ICC of the volume score of S4 and FQM was 1.000 (0.999-1.000); all other comparisons gave an ICC of 1.000 (1.000-1.000). RMSE for Agatston, volume, and mass score ranged between 0.02 – 1.01, 0.80 – 1.64 mm^3, and 0.00 – 0.22 mg, respectively. Reclassification of the calcifications occurred seven times out of ninety calcifications (7.8%) at CT-1 and three times out of ninety calcifications (3.3%) at CT-3. All reclassifications were from zero to minimal or vice versa. No reclassifications occurred with CT-2 and CT-4. This gave an inter-platform reliability of κ = 0.969 (p<0.0001), 95% CI [0.947, 0.991] between FQM$_{MATLAB}$ and the vendor-specific software and κ = 0.973 (p<0.0001), 95% CI [0.953, 0.993] between FQM$_{Python}$ and the vendor-specific software.

Algorithm robustness

FQM scores were also in high agreement with the vendor-specific software packages after varying the acquisition settings for all CAC scores. When FQM scores were compared with the vendor-specific scores, mean (95% CI) differences for Agatston, volume, and mass scores were -0.001 (-0.033 to 0.031), -0.2 (-0.365 to -0.035) mm^3, and -0.071 (-1.086 to 0.944) mg HA, respectively (*Figure 5*). ICC's (mean [95% CI]) were excellent (1.000 [1.000-1.000] for all CAC scores). No reclassifications occurred.
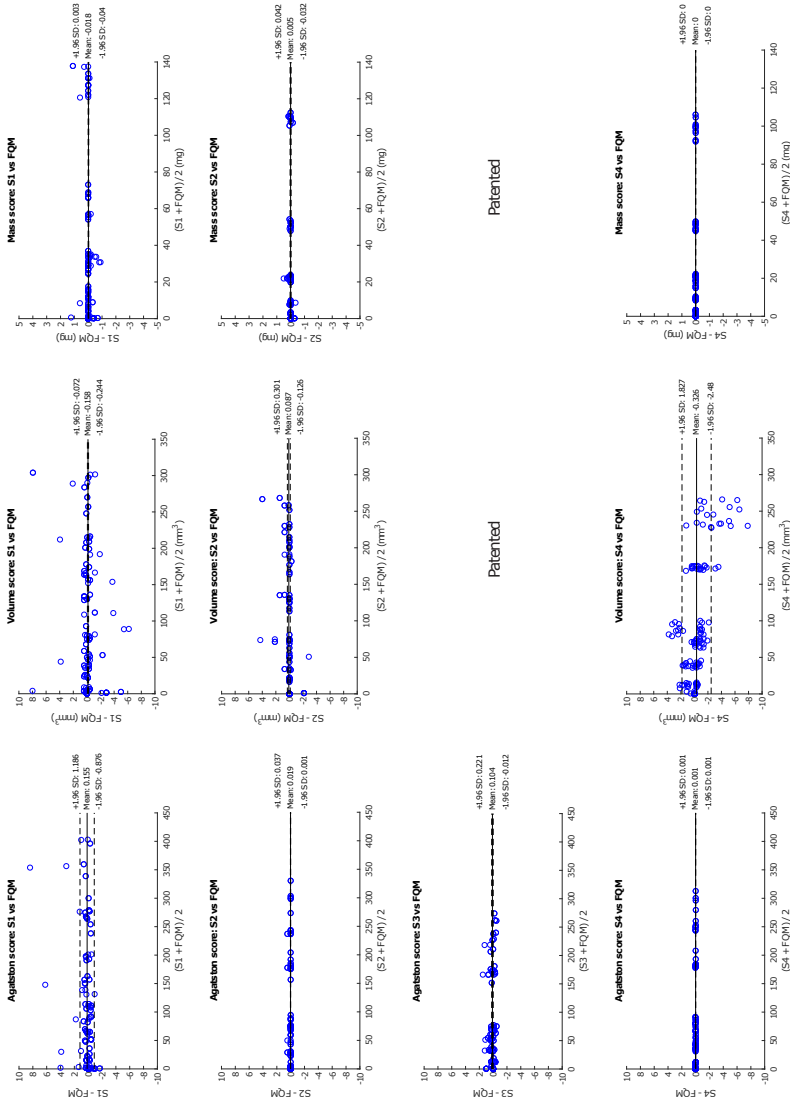
16

Figure 4 Bland-Altman plots of all CAC scoring software compared to the FQM. From left to right the Agatston, volume, and mass scores are shown, respectively. From top to bottom S1 to S4 are shown. Volume and mass scoring method of S3 were patented (the manufacturer was not able to provide any information) and could therefore not be implemented into the FQM.
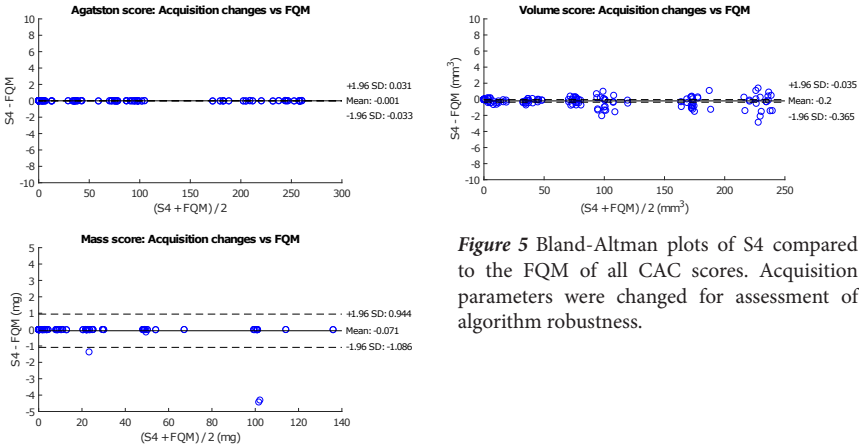
Figure 5 Bland-Altman plots of S4 compared to the FQM of all CAC scores. Acquisition parameters were changed for assessment of algorithm robustness.

RMSE were between 0.012 and 0.020 for Agatston scores, 0.220 and 0.835 mm^3 for volume scores, and 0.003 and 1.063 mg for mass scores. Remarkably, all RMSEs of mass scores were below 0.034 mg except for field of view changes, where RMSE scores were 1.042 and 1.063 mg for FOV 320 and 200, respectively.

For the dynamic phantom, FQM scores were in high agreement with vendor-specific software too. When FQM scores were compared with the vendor-specific scores, mean (95% CI) differences for Agatston, volume, and mass scores were -0.393 (-2.502 to 1.716), -0.514 (-10.177 to 9.15) mm^3, and -0.283 (-0.651 to 1.181) mg HA, respectively (*Figure 6*). ICCs (mean [95% CI]) were excellent (1.000 [1.000-1.000] for Agatston and mass scores and 0.999 [0.999-1.000] for volume scores). RMSE was 1.139, 4.926 mm^3, and 0.536 mg for Agatston, volume, and mass scores, respectively.

On a regular desktop computer (Windows 7, i5-6500 CPU 3.2 GHz, 8 GB RAM), evaluating a single scan with FQM took on average 3 or 6 seconds without or with interpolation for the volume score, respectively. In contrast, manual analysis of the phantom (without advanced image quality assessment) is in the order of minutes.

Image quality

For two datasets which were reconstructed with different reconstruction kernels (Qr44, Qr32), 10% and 50% TTF results were calculated (*Figure 7*). For the NPS analysis, images, ROI placement, and resulting 1D NPS curve results are shown in *Figure 8*. For both reconstruction kernels, NPS results were comparable between FQM and imQuest.
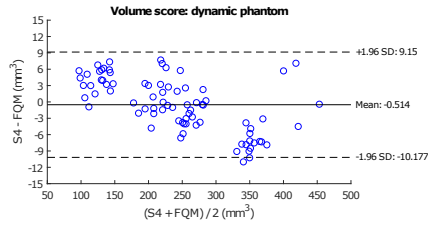
16
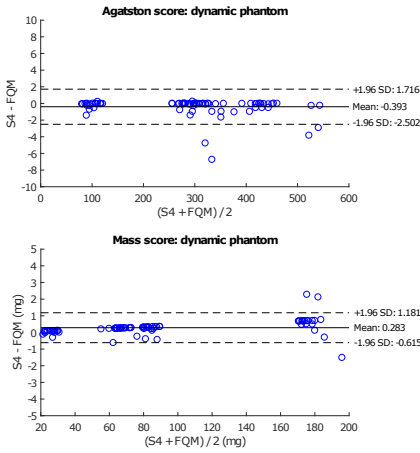
Figure 6 Bland-Altman plots of S4 compared to the FQM of all CAC scores. Scans were acquired with a dynamic phantom and scored with FQM for assessment of algorithm robustness.
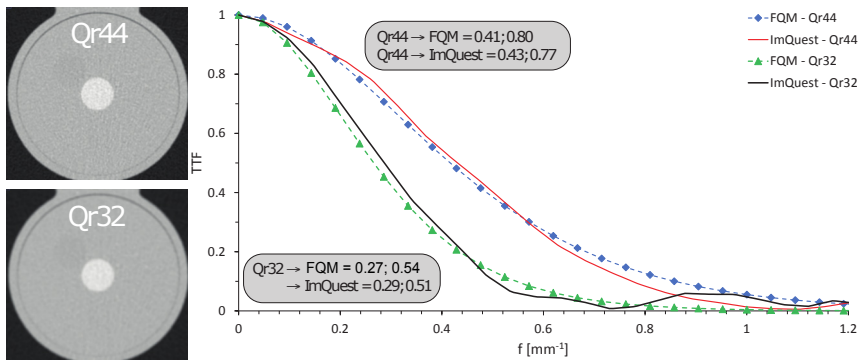
Figure 7 TTF results for both FQM and imQuest for two datasets, reconstructed with different reconstruction kernels. In addition, deviations at 50% and 10% TTF between results from both analyses are shown.

16
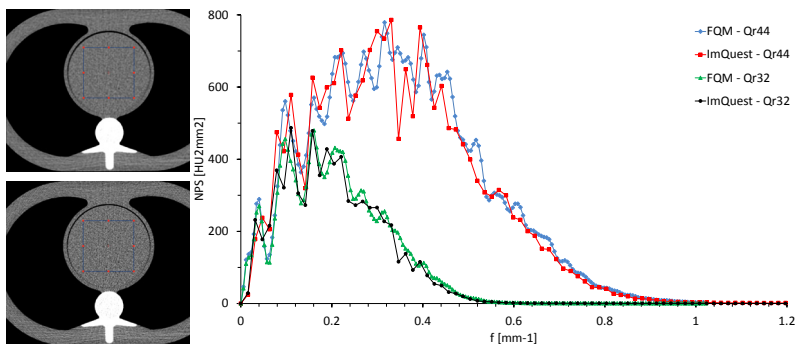
Figure 8 NPS results for both FQM and imQuest. Left, images for the Qr32 (upper) and Qr44 (lower) reconstruction kernel are shown, together with the placed ROI. Right, resulting 1D NPS results are shown. Small differences between both results are expected to be due to small differences in ROI placement.

DISCUSSION

In this study, we successfully developed an open-source, fully automated, vendor-independent, robust method to quantify CAC in two commonly used commercially available phantoms. In addition, we implemented vendor-specific scoring methods from four major calcium scoring software vendors with excellent agreement. Two scoring methods could not be implemented in our method due to non-disclosures. Also, image quality metrics, useful for comparison of CT scans with varying imaging parameters, were automatically extracted from the image data. These advanced image quality metrics can aid in assessing the influence of non-linear (post)processing steps on CAC scores.

Our algorithm is focused on a fully automated analysis of a standard anthropomorphic cardiac phantom. The main reasons for this focus, are the substantial reduction of evaluation time and the lack of inter- and intra-observer variability, manual notation errors, and software problems because of acquisition settings. An example of the latter is that some software programs are not able to process calcium scoring scans with a slice thickness different from the usual 3 mm, which can be rather inconvenient for research purposes. This in contrast to FQM, which is written in both MATLAB and Python, making it widely usable, depending on programming-language preference. This phantom is often used for careful evaluation of novel technical advances in CT, e.g., acquisition techniques, such as novel photon-counting detector elements, or reconstruction techniques, such as kernels which allow for tube voltage independent CAC acquisitions, before clinical usage.[16,25] FQM can aid in these experiments, as larger number of scans can easily be analyzed in a fully automatic manner.

Although the predictive role in risk stratification of low nonzero calcium scores caused by microcalcifications is still unknown, zero CAC scores are proven to be a strong negative predictor of CAD.[26–28] Also, Criqui and colleagues found an inversely proportional association of density on future cardiovascular events.[29] Therefore, the detection of small and low-density calcifications is of utmost importance. In our study, we found three main software parameters, which influence CAC detection and, therefore, quantification. First, the threshold used to discriminate calcium-containing voxels from non-calcium containing voxels. Second, the minimum calcification area threshold used to discriminate between noise and calcium containing voxels. And third, the use of isotropic interpolation for volume scores. All factors have an important impact on the detection of microcalcifications, especially for high noise acquisitions. For these acquisitions,

16

lower thresholds and use of interpolation will increase CAC area, and smaller minimum calcification areas will increase the number of false positives due to noise effects. It is thus important to investigate the exact influence of these parameters on CAC scores and the impact of scoring method standardization on differences in CAC scores between scanners. Besides that, the need of an improved CAC scoring method is high.[28,30,31] Both Agatston and volume scores show high variability in scores within and between CT systems.[11,12] The mass score is a more reliable score in terms of variability although small differences still exist.[32] FQM is thus a helpful tool in the development of new CT acquisition/reconstruction protocols and new scoring methods.

A few studies developed an automatic CAC scoring algorithm for patient CT angiography scans.[33–36] However, to the best of our knowledge, this is the first study that developed a fully automated, vendor-neutral method for quantification of CAC scores in a phantom. Also, no other study examined and reproduced the exact scoring methods of the four major calcium scoring software vendors. Only a few studies compared software platforms in CAC scores. However, these were either with platforms that are nowadays no longer widely used, or they compared scores, but did not go into detail about the parameters.[37–39] Weininger and colleagues used three different workstations, Syngo Calcium Scoring (Siemens), Aquarius (TeraRecon), and Vitrea (Vital Images), to acquire CAC scores of 59 patients.[39] Total Agatston and volume scores were compared between these systems. Although all results were numerically different, they found excellent correlations between the three workstations for both scoring methods.[39]

Our study has limitations. First, we were not able to implement the volume and mass quantification method of GE Healthcare. The vendor explained that they make use of a patented algorithm which adapts the threshold to help correct for beam hardening and overestimation. This adaptive threshold is used for both volume and mass scores. Another limitation of this study is that, currently, FQM can only be used in the described phantoms and not in patients or other phantoms as it makes use of the physical properties of these phantoms. However, these are commonly used phantoms for coronary calcium studies and FQM provides simple and fast analyses. Also, the main body of FQM can be rewritten to include other phantoms as we have shown in our flowchart and by validating both a static and a dynamic phantom. This increases the usability of FQM. Finally, only in-plane resolution measurements were added to the current version of FQM. Longitudinal measurements, based on the edge of the calibration rod, could be added in a future release.

CONCLUSIONS

In conclusion, we developed a fully automated, open-source, robust method in MATLAB and Python to quantify CAC in a commercially available and widely used phantom. The algorithm contains the international standard quantification methods described in literature, as well as almost all scoring methods of four major calcium scoring software vendors with an excellent agreement. The need for manual calcium scoring was completely eliminated with our fully automated method. Also, the automated algorithm contains image quality assessment for fast comparison of differences in acquisition and reconstruction parameters.

16

REFERENCES

1. Malguria N, Zimmerman S, Fishman EK. Coronary Artery Calcium Scoring: Current Status and Review of Literature. J Comput Assist Tomogr. 2018;42(6):887-897. doi:10.1097/RCT.0000000000000825

2. Van Der Bijl N, De Bruin PW, Geleijns J, et al. Assessment of coronary artery calcium by using volumetric 320-row multi-detector computed tomography: Comparison of 0.5 mm with 3.0 mm slice reconstructions. Int J Cardiovasc Imaging. 2010;26(4):473-482. doi:10.1007/s10554-010-9581-8

3. Keelan PC, Bielak LF, Ashai K, et al. Long-term prognostic value of coronary calcification detected by electron-beam computed tomography in patients undergoing coronary angiography. Circulation. 2001;104(4):412-417. doi:10.1161/hc2901.093112

4. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. J Am Coll Cardiol. 1990;15(4):827-832. doi:10.1016/0735-1097(90)90282-T

5. Divakaran S, Cheezum MK, Hulten EA, et al. Use of cardiac CT and calcium scoring for detecting coronary plaque: Implications on prognosis and patient management. Br J Radiol. 2015;88(1046). doi:10.1259/bjr.20140594

6. Hecht H, Blaha MJ, Berman DS, et al. Clinical indications for coronary artery calcium scoring in asymptomatic patients: Expert consensus statement from the Society of Cardiovascular Computed Tomography. J Cardiovasc Comput Tomogr. 2017;11(2):157-168. doi:10.1016/j.jcct.2017.02.010

7. Greenland P, Blaha MJ, Budoff MJ, Erbel R, Watson KE. Coronary Calcium Score and Cardiovascular Risk. J Am Coll Cardiol. 2018;72(4):434-447. doi:10.1016/j.jacc.2018.05.027

8. McCollough CH, Ulzheimer S, Halliburton SS, Shanneik K, White RD, Kalender WA. Coronary Artery Calcium: A Multi-institutional, Multimanufacturer International Standard for Quantification at Cardiac CT. Radiology. 2007;243(2):527-538. doi:10.1148/radiol.2432050808

9. Callister TQ, Cooil B, Raya SP, Lippolis NJ, Russo DJ, Raggi P. Coronary artery disease: improved reproducibility of calcium scoring with an electron-beam CT volumetric method. Radiology. 1998;208(3):807-814. doi:10.1148/radiology.208.3.9722864

10. van der Werf NR, Willemink MJ, Willems TP, Greuter MJW, Leiner T. Influence of dose reduction and iterative reconstruction on CT calcium scores: a multi-manufacturer dynamic phantom study. Int J Cardiovasc Imaging. 2017;33(6):899-914. doi:10.1007/s10554-017-1061-y

11. Willemink MJ, Vliegenthart R, Takx RAP, et al. Coronary Artery Calcification Scoring with State-of-the-Art CT Scanners from Different Vendors Has Substantial Effect on Risk Classification. Radiology. 2014;273(3):695-702. doi:10.1148/radiol.14140066

12. Rutten A, Isgum I, Prokop M. Coronary calcification: effect of small variation of scan starting position on Agatston, volume, and mass scores. Radiology. 2008;246(1):90-98. doi:10.1148/radiol.2461070006

13. van der Werf NR, Willemink MJ, Willems TP, Vliegenthart R, Greuter MJW, Leiner T. Influence of heart rate on coronary calcium scores: a multi-manufacturer phantom study. Int J Cardiovasc Imaging. 2017;34(6):959-966. doi:10.1007/s10554-017-1293-x

16

14. Willemink MJ, Abramiuc B, den Harder AM, et al. Coronary calcium scores are systematically underestimated at a large chest size: A multivendor phantom study. J Cardiovasc Comput Tomogr. 2015;9(5):415-421. doi:10.1016/j.jcct.2015.03.010

15. McCollough CH, Primak AN, Saba O, et al. Dose performance of a 64-channel dual-source CT scanner. Radiology. 2007;243(3):775-784. doi:10.1148/radiol.2433061165

16. Booij R, van der Werf NR, Budde RPJ, Bos D, van Straten M. Dose reduction for CT coronary calcium scoring with a calcium-aware image reconstruction technique: a phantom study. Eur Radiol. 2020;30(6):3346-3355. doi:10.1007/s00330-020-06709-9

17. Tang Y-C, Liu Y-C, Hsu M-Y, Tsai H-Y, Chen C-M. Adaptive Iterative Dose Reduction 3D Integrated with Automatic Tube Current Modulation for CT Coronary Artery Calcium Quantification: Comparison to Traditional Filtered Back Projection in an Anthropomorphic Phantom and Patients. Acad Radiol. Published online January 2018. doi:10.1016/j.acra.2017.12.018

18. Vonder M, Pelgrim GJ, Huijsse SEM, et al. Coronary artery calcium quantification on first, second and third generation dual source CT: A comparison study. J Cardiovasc Comput Tomogr. 2017;11(6):444-448. doi:10.1016/j.jcct.2017.09.002

19. Blobel J, Mews J, Goatman KA, Schuijf JD, Overlaet W. Calibration of coronary calcium scores determined using iterative image reconstruction (AIDR 3D) at 120, 100, and 80 kVp. Med Phys. 2016;43(4):1921-1932. doi:10.1118/1.4942484

20. Schindler A, Vliegenthart R, Schoepf UJ, et al. Iterative Image Reconstruction Techniques for CT Coronary Artery Calcium Quantification: Comparison with Traditional Filtered Back Projection in Vitro and in Vivo. Radiology. 2014;270(2):387-393. doi:10.1148/radiol.13130233

21. Husmann L, Leschka S, Desbiolles L, et al. Coronary artery motion and cardiac phases: dependency on heart rate -- implications for CT image reconstruction. Radiology. 2007;245(2):567-576. doi:10.1148/radiol.2451061791

22. The International Commision on Radiation Units and Measurements. ICRU Report no.87 - Radiation Dose and Image-Quality Assessment in Computed Tomography. J ICRU. 2012;12(1):1-149. doi:10.1093/jicru/ndsxxx

23. van Ommen F, Bennink E, Vlassenbroek A, et al. Image quality of conventional images of dual-layer SPECTRAL CT: A phantom study. Med Phys. 2018;45(7):3031-3042. doi:10.1002/mp.12959

24. Robins M, Solomon J, Richards T, Samei E. 3D task-transfer function representation of the signal transfer properties of low-contrast lesions in FBP- and iterative-reconstructed CT. Med Phys. 2018;45(11):4977-4985. doi:10.1002/mp.13205

25. Sandfort V, Persson M, Pourmorteza A, Noël PB, Fleischmann D, Willemink MJ. Spectral photon-counting CT in cardiovascular imaging. J Cardiovasc Comput Tomogr. 2020;(September). doi:10.1016/j.jcct.2020.12.005

26. Blaha M, Budoff MJ, Shaw LJ, et al. Absence of Coronary Artery Calcification and All-Cause Mortality. JACC Cardiovasc Imaging. 2009;2(6):692-700. doi:10.1016/j.jcmg.2009.03.009

27. Sarwar A, Shaw LJ, Shapiro MD, et al. Diagnostic and Prognostic Value of Absence of Coronary Artery Calcification. JACC Cardiovasc Imaging. 2009;2(6):675-688. doi:10.1016/j.jcmg.2008.12.031

16

28. Blaha MJ, Mortensen MB, Kianoush S, Tota-Maharaj R, Cainzos-Achirica M. Coronary Artery Calcium Scoring: Is It Time for a Change in Methodology? JACC Cardiovasc Imaging. Published online 2017. doi:10.1016/j.jcmg.2017.05.007

29. Criqui MH, Denenberg JO, Ix JH, et al. Calcium Density of Coronary Artery Plaque and Risk of Incident Cardiovascular Events. Jama. 2014;311(3):271. doi:10.1001/jama.2013.282535

30. Willemink MJ, van der Werf NR, Nieman K, Greuter MJW, Koweek LM, Fleischmann D. Coronary artery calcium: A technical argument for a new scoring method. J Cardiovasc Comput Tomogr. 2019;13(6):347-352. doi:10.1016/j.jcct.2018.10.014

31. Arnold BA, Budoff MJ, Child J, Xiang P, Mao SS. Coronary calcium test phantom containing true CaHA microspheres for evaluation of advanced CT calcium scoring methods. doi:10.1016/j.jcct.2010.08.004

32. Dijkstra H, Greuter MJW, Groen JM, et al. Coronary calcium mass scores measured by identical 64-slice MDCT scanners are comparable: A cardiac phantom study. Int J Cardiovasc Imaging. 2010;26(1):89-98. doi:10.1007/s10554-009-9503-9

33. Yang G, Chen Y, Ning X, Sun Q, Shu H, Coatrieux JL. Automatic coronary calcium scoring using noncontrast and contrast CT images. Med Phys. 2016;43(5):2174-2186. doi:10.1118/1.4945045

34. Lessmann N, van Ginneken B, Zreik M, et al. Automatic Calcium Scoring in Low-Dose Chest CT Using Deep Neural Networks With Dilated Convolutions. IEEE Trans Med Imaging. 2018;37(2):615—625. doi:10.1109/tmi.2017.2769839

35. de Vos BD, Wolterink JM, Leiner T, de Jong PA, Lessmann N, Isgum I. Direct Automatic Coronary Calcium Scoring in Cardiac and Chest CT. IEEE Trans Med Imaging. 2019;38(9):2127-2138. doi:10.1109/TMI.2019.2899534

36. Wolterink JM, Leiner T, de Vos BD, van Hamersvelt RW, Viergever MA, Išgum I. Automatic coronary artery calcium scoring in cardiac CT angiography using paired convolutional neural networks. Med Image Anal. 2016;34:123-136. doi:10.1016/j.media.2016.04.004

37. Yamamoto H, Budoff MJ, Lu B, Takasu J, Oudiz RJ, Mao S. Reproducibility of three different scoring systems for measurement of coronary calcium. Int J Cardiovasc Imaging. 2002;18(5):391-397. doi:10.1023/a:1016051606758

38. Adamzik M, Schmermund A, Reed JE, Adamzik S, Behrenbeck T, Sheedy PF. Comparison of two different software systems for electron-beam CT- derived quantification of coronary calcification. Invest Radiol. 1999;34(12):767-773. doi:10.1097/00004424-199912000-00006

39. Weininger M, Ritz KS, Schoepf UJ, et al. Interplatform Reproducibility of CT Coronary Calcium Scoring Software. Radiology. 2012;265(1):70-77. doi:10.1148/radiol.12112532

16

Online material is available via:
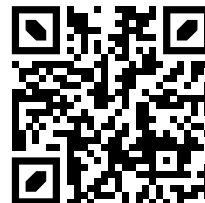
CHAPTER 17

Summary and General Discussion

SUMMARY

In this thesis the diagnostic performance of coronary artery calcium (CAC) detection and quantification with computed tomography (CT) was studied. First, spurious variations in CAC detection and quantification were identified and studied for three categories: patient specific factors, acquisition and reconstruction parameters, and post-processing of reconstructed images. Next, novel technical advances in CT were used to reduce CAC scoring variability, improve the prognostic value of CT examinations and/or reduce radiation dose exposure for patients.

PART 1 PATIENT SPECIFIC FACTORS

In the first part of this thesis, CAC detection and quantification as a function of patient specific parameters were assessed. One important parameter is the heart rate of a patient during the scan phase of a CAC CT assessment. Depending on heart rate and anatomical location, coronary arteries move at velocities of 10 to 30 mm/s during the CT acquisition, whereas coronary arteries are erroneously assumed to be stationary during CT acquisition in many studies. This motion of CAC leads to motion artefacts. For high density CAC, these motion artefacts increase the apparent size of a calcification. This results in an increased number of voxels exceeding the conventional CAC scoring threshold of 130 Hounsfield Units (HU), which subsequently increases CAC scores. For low density CAC the apparent size of a calcification is also increased. However, for these CAC, the HU only marginally exceed the 130 HU threshold when no motion is present. Motion artefacts result in increased partial volume effects, thereby decreasing the number of voxels which exceed the CAC scoring threshold and a potential reduction in the resulting CAC score for low density CAC.

In **chapter 2**, the influence of heart rate on CAC scores for high-end CT systems from four major manufacturers was assessed with a dynamic phantom. An artificial coronary artery, containing CAC of different densities, was translated in an anthropomorphic chest phantom at linear velocities (10 – 30 mm/s). These velocities corresponded to heart rates of <60, 60-75, >75 beats per minute (bpm). Raw data was acquired with CT vendor specific routinely used clinical CAC CT protocols. CAC scores, quantified as Agatston and mass scores, were compared to the reference scores at <60 bpm. While low density CAC Agatston scores were comparable to the reference, medium and high-density CAC Agatston scores increased significantly (up to 50%) with increasing heart rate. Susceptibility of CAC scores to motion was shown to differ significantly between different CT

systems. Follow-up CAC CT scans should, therefore, be acquired on the same CT system, with similar protocols and comparable heart rates.

In **chapters 3 & 4**, a deep learning network was used to identify CAC motion artefacts on CT images. These identified motion artefacts could be looked up in a library of motion artefacts for different CT systems and CAC densities. Next, the impact of the motion artefact on the CAC score could be nullified by taking the CAC score without motion from the library. To be able to identify the motion artefact, first, a deep learning network (an inception v3 convolutional neural network (CNN)) was trained to identify motion blurring of nine CAC categories of combinations of different densities (high, medium, and low) and different sizes (small, medium, and large), acquired on two CT systems from two manufacturers with non-triggered chest protocols. For each of the nine combinations of CAC density and size, motion artefacts from acquisitions with CAC movement at 0 – 90 mm/s were used. After classification, CAC scores (Agatston, volume, and mass scores) were corrected for motion by assigning the motion-free CAC score to the CAC of that specific category. A calcification by CNN correction was considered if the uncorrected score was zero, but the corrected score was non-zero. The overall accuracy of the CNN classification was 79 ± 6%. Compared to resting CAC scores, the overall Agatston score variation was 38% for CAC scores obtained in motion. This variation was decreased to 4% by CNN correction. Sensitivity, defined as the percentage of non-zero Agatston scores among all scores, increased from 65% to 85% with CNN correction. Second, three deep learning networks were trained on CAC CT data from four major CT manufacturers to again identify CAC motion artefacts. For these networks, raw data was acquired at reference clinical radiation dose, and 40% and 80% reduced radiation dose levels for CAC (low, medium, and high density) which were translated in an anthropomorphic phantom at linear velocities of 0 – 60 mm/s. Subsequently raw data was reconstructed with filtered back projection (FBP) and at least three levels of iterative reconstruction (IR). Our CNNs categorized motion artefacts from the reconstructed images in categories based on combinations of CT system and CAC density. Overall motion artefact classification of the CNN was 90% to 94%, with the highest accuracy associated with the DenseNet architecture. All deep CNN architectures were capable of classifying motion-blurred images to the correct CT – CAC density category with a high accuracy, regardless of CT vendor, velocity, radiation dose, and reconstruction algorithm. Clinical application of these results is hampered by the absence of a library with in-vivo motion artefacts for CAC for different heart rates. If present, it would allow for CAC scoring acquisitions with CT independent of heart rate, i.e.

17

as used in a screening setting or with attenuation correction scans at the Nuclear Medicine department.

PART 2 ACQUISITION AND RECONSTRUCTION PARAMETERS

In part 2, the influence of acquisition and reconstruction parameters on CAC detection and quantification was assessed.

In **chapter 5** a systematic review on the impact of dose reduction on CAC quantification and risk categorization was presented. Twenty-eight studies were included, among which 17 patient studies, 10 phantom studies and 1 both patient and phantom study. In these studies, tube voltage reduction, tube current reduction, IR, spectral shaping with tin filtration and combinations of these techniques were used to reduce patient radiation dose, without compromising CAC quantification. In 78% of the studies, radiation dose was reduced by ≥50% with Agatston score risk reclassification rates between 3% and 21%. Specific combinations of tube current reduction and IR, or spectral shaping with tin filtration, that showed low reclassification rates may potentially be used in CAC CT protocols or future population-based screening for cardiovascular risk stratification.

In **chapter 6**, a technical argument was presented for the optimization of CAC acquisition and reconstruction protocols for newer CT technology. Optimized CAC quantification should be acquired at a low tube voltage allowing for detection of small calcifications with lower densities and images that are reconstructed with IR and thin slices allowing for more precise 3-dimensional CAC assessment with improved reproducibility. A lower tube voltage also decreases radiation exposure. Optimized CAC quantification methods should be evaluated in large observational studies with long follow-up time. Alternatively, anthropomorphic phantom studies and clinical trials should indicate whether novel CAC quantification methods improve small and low-density CAC detection, while maintaining comparable CAC quantification results. We expect that optimized protocols will result in improved reproducibility and will allow for improved evaluation of CAC density, shape, and distribution within the coronary artery tree.

In **chapter 7** we evaluated the impact of IR, as an alternative to FBP, on CAC scores at different heart rates, using the same setup as in chapter 2. Here the artificial coronary artery was translated at 0, 10, 20, and 30 mm/s, corresponding to heart rates of 0, <60, 60-75 and >75 bpm, respectively. At each heart rate, raw data was reconstructed with FBP, and three increasing levels of IR. CAC was quantified

17

as Agatston and mass scores. Furthermore, an IR susceptibility index (IRS) was introduced, to assess the susceptibility of CAC scores to IR. CAC scores originating from FBP reconstructions were used as the reference, with which IR CAC scores were compared. In general, Agatston scores were more susceptible to increased strength of IR when compared to mass scores. The effect of IR on CAC scores was shown to not only depend on CT system and calcification density, but also on heart rate. Therefore, follow-up CAC assessments should be performed on the same CT system, with the same reconstruction type and, preferably, at similar heart rate.

Chapter 8 describes a study that assessed the influence of radiation dose reduction and IR on CAC scores for all major CT manufacturers. Again, a dynamic anthropomorphic phantom was used. Artificial coronary arteries were translated at 20 mm/s or 60-75 bpm. Raw data was acquired with CT specific routinely used CAC CT protocols on four high-end CT systems from four major CT manufacturers. Besides reference radiation dose, data was acquired at 40% and 80% reduced radiation dose. Decreased radiation dose resulted in increased CAC scores, while at increased IR levels, CAC scores decreased again. For all CT systems, similar CAC scores were found at 40% reduced radiation dose. For some CT systems, CAC scores were not affected at 80% radiation dose reduction in combination with IR. So, for all CT systems, a radiation dose reduction of at least 40% was feasible, without affecting CAC quantification.

Chapters 9 and 10 describe CAC detection and quantification for a CT system equipped with a novel photon-counting detector. First, the CAC scoring potential of this spectral photon-counting CT (SPCCT) was assessed, by comparison with a conventional CT at routine clinical protocol. Additionally, improved CAC detection and quantification at reduced slice thickness was assessed. For this, two CAC containing cylindrical inserts were positioned in an anthropomorphic chest phantom, which was scanned on both a conventional CT and a SPCCT system. To assess the influence of patient size, the chest phantom was scanned without and with extension ring, to simulate a small and large patient, respectively. CAC was quantified as Agatston and volume scores. Routine clinical protocol CAC scores were comparable between both CT systems. Small phantom CAC detection increased at reduced slice thickness by 142% and 169% for CT and SPCCT, respectively. For large phantom size, CAC detection increased by 31% and 110%. In comparison with CT, SPCCT physical volume approximation was superior. Overall, CAC assessment with SPCCT yields comparable quantification results as CT for routine clinical protocols. Furthermore, the increased spatial resolution of SPCCT allows

17

for increased detectability and more accurate CAC volume estimation. Second, improved CAC detection and quantification with SPCCT was assessed for reduced slice thickness and reduced radiation dose. The same phantom setup and clinical protocol was used as for the previous paper. Additionally, data was acquired at 50% radiation dose, by reducing tube current. CAC quantification was performed with volume and mass scores. At both reduced radiation dose and reduced slice thickness, small phantom CAC detection increased by 108% and 150% for CT and SPCCT, respectively. For the large phantom size, noise levels interfered with CAC detection. At routine clinical protocols, median volume score deviations from physical CAC volume were large (up to 134%). At reduced radiation dose and slice thickness, physical volume overestimations were decreased to 96% and 72% for CT and SPCCT, respectively. In comparison with volume scores, mass score deviation from physical quantities were smaller. In conclusion, CAC detection on SPCCT is superior to CT, and was even preserved at reduced radiation dose. Furthermore, SPCCT allows for improved physical volume approximation.

In the following two chapters, a novel calcium-aware (Ca-aware) reconstruction technique was assessed. CT numbers, and therefore CAC scores, depend on the used tube potential. However, acquiring data at a reduced tube potential can increase calcium contrast and reduce patient radiation dose. With the Ca-aware reconstruction kernel, CT numbers of calcium are scaled to match the CT numbers that would have been measured at reference tube potential (defined at 120 kVp), enabling the use of patient specific tube potentials.

First, in **chapter 11**, the dose reduction potential of this novel reconstruction technique was assessed using a static anthropomorphic phantom mimicking three patient sizes (small, medium, and large). Two CAC containing inserts were used, containing CAC varying in size and density. Tube potentials were varied from 70 – 150 kVp. In addition, a dedicated tin-filtered (100Sn kVp) protocol was used. For all acquisitions, tube current was automatically adapted to maintain reference image quality. Standard and Ca-aware reconstruction techniques were applied to reconstruct raw data. In comparison with the reference protocol, phantom size specific tube potential selection resulted in 22%, 15% and 12% radiation dose reduction for the small, medium, and large phantom, respectively. Tin-filtered protocols led to a radiation dose reduction of 55%, 55% and 60%. In comparison with the reference, CT numbers deviated up to 64% for the standard reconstruction technique for varying tube potentials. This deviation was reduced to 11% when the Ca-aware reconstruction technique was applied. Agatston scores deviated up to 40% for the standard reconstruction technique, which was reduced to 8% for

the Ca-aware reconstruction technique. Overall, CT numbers remained consistent with comparable calcium scores when the calcium-aware image reconstruction technique was applied with a varying tube voltage. Less consistency was observed in small calcifications with low density. Automatic reduction of tube voltage resulted in a dose reduction of up to 22%.

Second, in **chapter 12**, a dynamic phantom was used to assess the influence of the Ca-aware reconstruction technique on CAC scores of moving calcifications with different densities surrounded by a variety of patient equivalent tissues. Raw data was acquired with three protocols: a routine clinical CAC protocol at 120 kVp, a reduced tube voltage protocol at 100 kVp, and a tin-filtration protocol at 150Sn kVp. Again, raw data was reconstructed with the standard and a Ca-aware reconstruction technique. For each protocol, the center compartment of an anthropomorphic phantom was filled with mixtures, which resembled fat, water, and soft tissue CT numbers. Due to the tube potential dependency of HU, the center compartment was refilled for each resembling material for each tube potential. For all heart rates, Agatston scores of CAC submerged in fat were comparable (<10% deviation) to the reference (120 kVp + standard reconstruction) for the reduced-kVp acquisition with a Ca-aware reconstruction kernel. For water and soft tissue, medium density Agatston scores were again comparable to the reference for all heart rates, while low density Agatston scores showed relevant deviations (>10%). In conclusion, the Ca-aware reconstruction technique allows for the use of patient-specific tube voltages for CAC surrounded by fatty materials, as found in-vivo, regardless of CAC density.

Chapter 13 evaluated the influence of four generations of CT reconstruction algorithms from one vendor on coronary calcium detection and quantification. For this, a phantom and 50 patients were scanned with routine settings from a clinical CAC protocol. Data was reconstructed with FBP, hybrid IR (HIR), model-based IR (MBIR) and deep-learning reconstruction (DLR) algorithms. For the first part of this study, both a static and a dynamic anthropomorphic phantom were used. The static phantom, containing 100 small calcifications, was used to evaluate CAC detection with different reconstruction algorithms. The dynamic phantom, in which low, medium, and high-density CAC was translated at 60-75 bpm, was used to assess the influence of the four generations of reconstruction methods on Agatston scores. Besides standard radiation dose, tube current was reduced for the phantom scans to assess the influence of radiation dose reduction with IR or DLR. The second part of the study concerned fifty patient scans, for which data were also reconstructed with FBP, HIR, MBIR, and DLR. In comparison with FBP,

17

CAC detection reduced for all reconstruction algorithms, except for HIR. For the standard radiation dose, CAC detection decreased up to 22%. CAC detection further decreased with decreasing radiation dose. For the dynamic phantom, differences in Agatston score between HIR, MBIR, and DLR and the reference (FBP) were not clinically relevant. For the patient study, the cardiovascular risk classification resulted for all reconstruction techniques in excellent agreement with the reference (κ=0.97 [95% CI: 0.94-1.0]; κ=0.96 [95% CI: 0.92-1.0]; κ=0.97 [95% CI: 0.94-1.0] for HIR, MBIR, and DLR, respectively), although MBIR resulted in significantly higher Agatston scores (p<0.001), resulting in 3/50(6%) patients with risk reclassifications. Overall, Agatston score agreement between FBP, HIR, MBIR, and DLR was excellent, with low risk reclassification rates. However, large variation in CAC detection of small CAC was found for the four reconstruction generations.

In **chapter 14**, an update for current CAC scoring protocols is proposed, which is aimed at increased reproducibility and reduced radiation dose for CAC assessment with CT for the four main CT manufacturers. For this, an anthropomorphic thorax phantom with two extension rings, which contained an international standardized cardiac insert was scanned using routine clinical protocols on state-of-the-art CT systems from all four main CT manufacturers. Additionally, tube voltages (80-120 kVp), tube currents (100% to 25% dose levels), slice thicknesses (3/2.5 and 1/1.25-mm), and reconstruction techniques (FBP and IR) parameters were varied. CAC was quantified as Agatston scores. Inter- and intrascanner variability was assessed using interquartile ranges (IQR). In comparison with the reference routine clinical protocol, a protocol based on a tube voltage of 100 kVp, 75% radiation dose, slice thickness of 1 or 1.25-mm, and increased IR-levels resulted in an on average 36% lower intrascanner IQR. This protocol led to increased CAC detection, with 6.2±0.4 and 7.0±0.4 detected CAC for the reference and proposed protocol, respectively. Pairwise comparisons of Agatston scores between scanners within the same phantom size demonstrated three significantly different comparisons at the standard protocol (P<0.05), whereas no significantly different comparisons arose at the proposed protocol (P>0.05). Therefore, we proposed to use an updated CAC protocol with 100 kVp, 25% reduced dose, and thin slices of 1 mm. This led to improved intra- and inter-scanner reproducibility and increased detectability of small and low-density calcifications. Before clinical use, this protocol should be extensively validated in a real world setting. After this has been done, it could potentially improve clinical inter-scanner/interinstitutional reproducibility and enable more consistent risk assessment and treatment strategies.

17

PART 3 POST-PROCESSING PARAMETERS

Part 3 of this thesis was concerned with the influence of post-processing parameters on CAC detection and quantification.

In order to reduce the current substantial variation in CAC scores acquired with different computed tomography (CT) scanners, in **chapter 15** a vendor-neutral Agatston score (vnAS) is presented and validated. The vnAS calibration tool was derived from scans of two static anthropomorphic CAC containing phantoms on seven different CT systems and one EBT system. Regression analysis was used to convert the actual CT Agatston scores to virtual EBT Agatston scores, or vnAS. To assess the effect of vnAS on event prediction we included 3,181 participants from the Multi-Ethnic Study on Atherosclerosis (MESA). Chi-square was used to compare coronary heart disease (CHD) event rates between low calcium (vnAS < 100) and high calcium groups (vnAS ≥ 100). Multivariable Cox proportional hazard regression models were used to assess the added value of vnAS. For a sub-cohort of 890 participants at intermediate atherosclerotic cardiovascular disease risk (ASCVD risk between 7.5-19.9%, no diabetes, LDL-C 70-189 mg/dl), ASCVD event rate and statin therapy number needed to treat (NNT) were determined for patients with vnAS ≥ 100 (group 1) and vnAS ≥ 300 (group 2) with original Agatston scores below 100 and 300, respectively. For all CT systems, CT Agatston scores showed a high degree of correlation with EBT Agatston scores ($R^2 > 0.932$). For large patients, Agatston scores in the range of 88-148 obtained from state-of-the-art CT converted to a vnAS of 100. For MESA participants, vnAS increased Agatston scores with 85 participants (11%) reclassified from the low calcium group (vnAS < 100) to the high calcium group (vnAS ≥ 100). For these reclassified participants, CHD event rates increased significantly from 7% to 15% (p = 0.008) with a CHD hazard ratio of 3.39 (95%CI 1.82–6.35, p = 0.001). For participants at intermediate cardiovascular risk, ASCVD event rates increased from 17.6% to 23.3% and from 27.6% to 38.2% for group 1 and 2, respectively. This increase in ASVCD event rate was accompanied by a decrease in statin therapy NNT from 12 to 7 and 15 to 2, again for group 1 and 2, respectively. Overall, MESA participants who were reclassified to a higher calcium group based on vnAS, showed increased CHD event rates. Also, the potential benefit from statin therapy was increased for reclassified participants at intermediate cardiovascular risk. In conclusion, calculation of a vnAS improves risk prediction and thus may improve patient treatment and outcome.

17

Finally, in **chapter 16**, a fully automated, open-source quantification method (FQM) for coronary calcium in a standardized phantom was developed and validated. Both the international standard CAC quantification methods as well as vendor-specific adjustments were implemented in FQM. Further, image quality metrics were implemented: noise power spectrum, task transfer function, and contrast- and signal-to-noise ratio among others. Three parameters in CAC scoring methods varied among different vendor-specific software packages: the Hounsfield unit (HU) threshold, the minimum area used to designate a group of voxels as calcium, and the usage of isotropic voxels for the volume score. All these parameters were implemented in FQM. Overall, FQM was in excellent agreement with vendor specific CAC scores. The need for manual calcium scoring was therefore completely eliminated with our fully automated method. Also, the automated algorithm contains image quality assessment for fast comparison of the influence of differences in acquisition and reconstruction parameters on resulting CT image quality.

17

GENERAL DISCUSSION

The overall aims of the research described in this thesis are to better understand spurious variations in coronary artery calcium (CAC) scoring and to provide strategies to address this problem. To better understand sources of variation we investigated the extent to which differences in CAC scores are attributable to the type of CT system, patient size and heart rate, level of radiation dose reduction, iterative reconstruction (IR), deep-learning reconstruction (DLR), calcium-aware reconstruction kernel, slice thickness, CAC scoring algorithms, and CT detector technology, rather than true variations in the amount of CAC. Importantly, we demonstrate that technical advances in CT can improve clinical CAC scoring, by reducing CAC scoring variability and/or reducing patient radiation dose.

Challenges in CAC scoring

CAC detection and quantification with the Agatston methodology on CT is strongly related to future adverse cardiovascular events.[1,2] Agatston scores are used for patient risk stratification and potential subsequent pharmacological treatment. Moreover, progression in CAC burden over multiple subsequent CT scans is associated with increased risks for adverse events.[3] To be able to define CAC progression, reproducible CAC scores are essential. This is also true for CAC scores in general, as two patients with the same amount of coronary calcium undergoing a CT CAC exam should receive equal Agatston scores, independent of the specific CT system used over different vendors.

However, many studies, including those performed as part of the current thesis, found poor Agatston score reproducibility with differences in Agatston scores up to 50%[4,5] based on several parameters such as patient size[5], heart rate[4,6-9], coronary vessel displacement and vessel trajectory[6], CT system[4,10-15], CT detector technology[16,17], spectral beam shaping[18,19], scan starting position[20], cardiac phase[7], reconstruction technique[14,21-24], slice increments[25], slice thickness[9], CAC scoring software[26,27], and CAC quantification parameters[27-29]. The poor Agatston score reproducibility even occurs when using the vendor recommended CT CAC protocol. Likewise, several studies showed that a change in one of these parameters can result in risk reclassifications.[19,30] For clinical examinations, especially when ionizing radiation is used, robustness is key.

17

Enhancing Agatston score reproducibility

To enhance robustness, or improve CAC scoring reproducibility, technical advances in CT technology can be employed. Since the introduction of current clinical CAC protocol parameters by McCollough et al. in 2007, several significant advances in CT technology have been made. These include increased detector coverage, more efficient detector technology, increased spatial and temporal resolution, and improved reconstruction algorithms. An important argument for continued use of current protocols is the wide breadth of available clinical outcome data across age, gender, and race.[3,31,32] However, these technical developments in CT can be used to further improve CAC detection and reproducibility of CAC quantification by, for example, motion artefact reduction, reduced slice thickness reconstruction, or by applying other advanced reconstruction algorithms.[33]

Motion artefacts

In this thesis, we found that Agatston scores obtained with high-end CT systems of four major CT manufacturers were substantially, but not equally, influenced up to 50% by heart rate.[4] The magnitude and direction of this influence depended on CAC density. For low-density CAC, motion blurring resulted in less voxels exceeding the CAC scoring threshold of 130 Hounsfield Units (HU), thereby decreasing the Agatston score, while the opposite was true for high-density CAC. Furthermore, the Agatston score uses weighting factors based on the maximum CT number within each slice of a CAC. Motion artefacts can also decrease this maximum number, thereby decreasing the resulting Agatston score. If, for example, the maximum CT number was 201 HU without coronary artery motion and 198 HU with coronary artery motion, this would result in a reduction of the Agatston score by a factor 2.[2,3]

Our results are in line with previous studies, which found that Agatston scores are not only affected by heart rate, but also by the specific coronary artery trajectory and vessel displacement.[6,34] Several solutions have been presented to reduce motion artefacts resulting from coronary artery motion during the CT scan phase. One solution is to acquire data at specific cardiac phases as a function of heart rate: 70% or 40% of the RR interval for heart rates below and above 70 beats per minute (bpm), respectively.[9,35] Greuter et al. presented another solution to overcome the influence of motion on CAC scores, based on the linear relation between coronary artery velocity and Agatston scores.[36] With this method, however, it is not possible to correct for reduced maximum voxel values and its potential effect on the Agatston weighting factor.

We have presented another solution to overcome motion artefacts in this thesis, by using a convolution neural network (CNN) approach to nullify motion artefacts.[37,38] For this approach, a library of motion artefacts from a dynamic phantom with moving calcifications was necessary. However, a direct clinical application of this technique is hampered by the absence of such a library with patient data. If available, it would allow for heart rate independent CAC scoring acquisitions with CT, as used in screening settings or in attenuation correction scans used in positron emission tomography (PET) or single photon emission tomography (SPECT) scans.[39–42]

Slice thickness reduction

To further increase Agatston reproducibility, we propose a novel CAC protocol on state-of-the-art CT systems from four vendors, using thin-slice reconstructions. For this protocol, improved intra- and inter-scanner reproducibility and increased detectability of small and low-density calcifications was shown. This phantom study should be seen as a first step towards a new CAC scoring standard with CT, after extensive clinical validation. We found increased CAC detectability of small calcifications for reduced slice thickness reconstructions of 142% and 169% for conventional energy integrating detector (EID) CT and photon-counting detector (PCD) CT, respectively.[17] The proposed thin-slice reconstruction protocol is also in line with several other studies which show improved reproducibility by reducing partial volume effects due to reduced slice thickness reconstructions.[12,20,34] In addition to increased reproducibility, reduced slice thickness reconstructions also allow for improved confidence in a zero CAC score, with less false negative findings (5% of subjects), as previously shown by Groen et al.[43,44] Near zero CAC scores result from small and/or low-density CAC. These results are in agreement with a recent study by Sandstedt et al., which indicated more accurate CAC volume approximation with PCD CT.[45] Another method to decrease false negative findings was used by Schlosser et al., which showed that the use of overlapping slices may help to detect small CAC.[25] Conventional Agatston scores for reduced slice thickness reconstructions can still be calculated, by multiplying the scores with the reduced slice thickness divided by the conventional 3 mm slice thickness. This approach is used clinically by two CT manufacturers, where Siemens recommends CAC protocols with overlapping slices (3mm with 1.5 mm increment), and where GE uses 2.5 mm slice thickness reconstructions.

17

Reconstruction algorithm

Conventionally, filtered back projection (FBP) is used to reconstruct raw CT data from CAC examinations. In this thesis, we showed that the impact of the IR algorithm depends on the CAC density.[14] For low-density CAC, especially at an increased heart rate >75 bpm, the use of IR resulted in reduced Agatston scores up to 21% for all major CT manufacturers. For medium and high-density CAC, a general decline in Agatston scores was shown for increased IR levels. For these densities, the differences with FBP Agatston scores were only minor. This finding is in line with a study by Takahashi et al., who also found decreased Agatston scores when IR was applied.[46] Despite changes in Agatston scores, we also showed in this thesis that the impact of four generations of reconstruction algorithms (including hybrid IR (HIR), model-based IR (MBIR) and DLR) from the same vendor did not result in clinically relevant differences in Agatston score for heart rates of 60-75 bpm. However, in comparison with FBP, changes in reconstruction algorithm did result in decreased CAC detection. These results are in line with findings by Schindler et al., who also found that the application of IR did not have a profound effect on reproducibility of Agatston scores in comparison with FBP for a patient study.[21] In addition, their in-vitro assessment did not show significant differences when all CAC densities of their phantom setup were summed. Only for high-density CAC significant differences between FBP and IR were shown. Oda et al. showed a decrease in Agatston variability when MBIR was applied, yielding more stable and reproducible measurements.[47]

Post-processing

We have shown in this thesis that CAC scoring reproducibility can also be enhanced by recalculating vendor and CT specific Agatston scores to a vendor-neutral electron beam tomography (EBT) Agatston score (or vnAS). Willemink et al. previously showed that CAC detection and quantification differed between modern CT scanners currently used in every day clinical practice.[12] This difference was associated with a potential of risk misclassification and improper treatment of up to 6.5% of asymptomatic individuals. Validation of our vnAS for participants of the Multi-Ethnic Study of Atherosclerosis showed significantly increased hazard ratios for coronary heart disease events for reclassified individuals. Furthermore, for individuals at intermediate atherosclerotic cardiovascular disease, the number needed to treat was smaller for reclassified individuals based on vnAS.

17

Conclusion

Overall, the simplest method to improve Agatston score reproducibility in individual patients is by using the same CT system and similar acquisition and reconstruction protocols as well as low heart rates for subsequent CAC examinations. Furthermore, reduced slice thickness reconstructions should be used to reduce variability and reduce false negative findings. Finally, a vendor and CT neutral score should be used, given the absence of large cohort studies with well-documented outcomes using CAC scoring derived from different modern CT scanners.

Reduced radiation dose CAC scoring

Another important aspect regarding CAC detection and quantification with CT is related to the use of ionizing radiation. In general, the utilization of CT in diagnostic radiology has increased dramatically, with CT examinations in the United States of America (USA) rising from 20 million in 1995 to almost 79 million in 2015, which amounts to a growth rate of more than 10% per year.[48,49] This increased usage of CT is associated with an increased exposure to ionizing radiation, where 75.4% of radiation dose burden in the United States of America (USA) originated from CT examinations in 2009.[50] This CT radiation dose was estimated to be responsible for 1.5 to 2.0% of all cancers in the USA.[51]

For CAC examinations specifically, measurements are recommended by several guidelines to improve clinical risk prediction in appropriately selected asymptomatic individuals, which therefore results in a high number of examinations.[32,52,53] For a screening program with an interval of five years, Kim et al. estimated the excess lifetime cancer risk for a typical CT CAC examination with an effective dose of 2.3 mSv to be 42 cases per 100 000 men, or 62 cases per 100 000 women.[54] According to the as-low-as-reasonably-achievable (ALARA) principle, any potential reduction of these number of cases by reducing the radiation dose should be considered.

To be able to reduce radiation dose in CT CAC exams, the associated influence on increased image noise should be assessed. Agatston et al. introduced two parameters to discriminate between noise and actual CAC containing voxels: a minimum CAC size of 1 mm^2 and a CAC scoring threshold of 130 HU.[2]

For the first CAC scoring parameter, vendor-specific implementations were shown to range from no threshold to 1 mm^2 in this thesis.[27] For one software platform,

17

a minimum of 3 connected voxels was used. This results in a patient specific minimum area, as a function of the used field-of-view (FOV) and matrix size (i.e. voxel dimensions = FOV / matrix size). Two studies confirm our results, as they also show numerical deviations in Agatston scores for the same images, when analyzed with different software platforms.[26,55] These studies even showed that the influence of CAC scoring software in CAC volume calculation was larger than intersystem variability for CT CAC scoring. Szilveszter et al. showed that model-based IR (MBIR) techniques allowed for more robust identification of CAC in comparison with FBP.[22] This is because of the reduced noise levels for MBIR, which thereby decreases the possibility of erroneous identification of noise as CAC.

The second CAC scoring parameter, the threshold of 130 HU, is used by the CAC scoring software of all main CT manufacturers.[27] However, this parameter was originally developed for CAC scans from EBT acquisitions, which were known for their higher image noise levels.[36] According to the ALARA principle, McCollough et al. proposed a minimum noise level in water to keep radiation dose levels at an acceptable level.[56] These levels were 20 and 23 HU for small and large patients, respectively. An upper limit for noise levels in water, which would indicate a minimum radiation dose level needed for diagnostic image quality, is lacking.

The current thesis proposes such an upper noise limit, above which the noise in a combination of acquisition and reconstruction settings would result in false positive CAC detection. We have termed this measure the Background Agatston Score (BAS), for which the Agatston score was calculated in a large region-of-interest in a homogeneous, non-CAC containing region.[27,57] Reconstructions with BAS > 0 can potentially lead to spurious identification of noise as CAC and should therefore not be used for clinical CAC detection and quantification.

Due to vendor-specific implementations of CAC scoring parameters, the optimal combination of acquisition and reconstruction parameters vary by vendor, as illustrated in *Figure 1*. Here, acquisitions with full and 50% of the routinely used radiation dose on one CT system were reconstructed with the same parameters. At full dose, all but one manufacturer showed no false positives. The influence of image noise on CAC assessment is smaller when the parameters from literature or GE are applied. For these scoring methods, the BAS is equal to zero at reduced dose. Therefore, with these scoring methods, radiation dose reduction is possible, without changing the CAC outcome. For all other manufacturers, lowering radiation dose led to false positive identification of CAC, with a maximum Agatston score of 40 based on only noise.

17
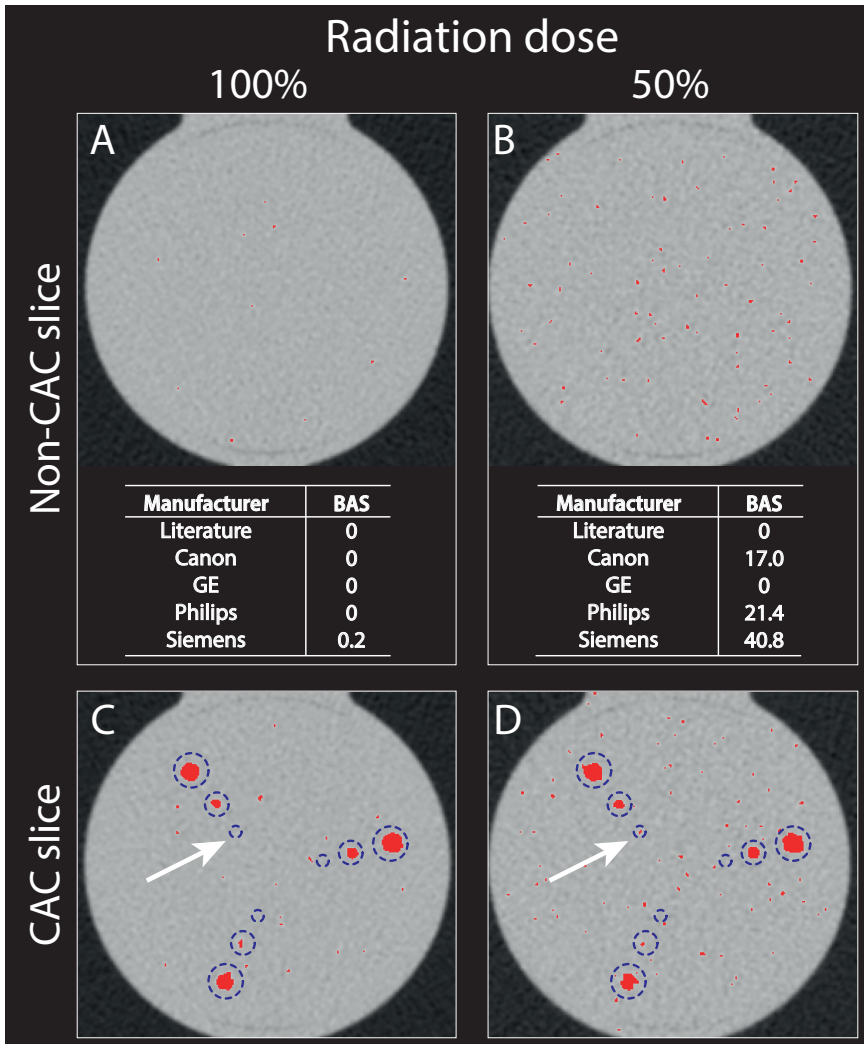
Figure 1 Influence of dose reduction on false positive CAC scores as indicated by the Background Agatston score (BAS) when calculated for literature and manufacturer-specific CAC scoring parameters for slices without CAC (subfigures **A** and **B**). In subfigures **C** and **D**, CAC containing slices at the corresponding dose level are shown, with the known CAC location of the cylindrical calcium insert (CCI, QRM) as indicated by the blue dashed circles. The white arrow indicates a CAC location which cannot be detected with current acquisition and reconstruction settings for 100% dose. At 50% radiation dose, however, increased noise levels result in a false positive CAC assessment at the location of the white arrow

17

Another technical advancement in CT which could allow for reduced radiation dose acquisitions is related to advanced image reconstruction such as IR and deep-learning reconstruction (DLR) techniques. In comparison with conventional FBP, both IR and DLR allow for reduced image noise.[23,58–61] In the current thesis, IR algorithms from all main CT manufacturers were used to reduce image noise for reduced radiation dose CT CAC exams by reducing the tube current in a dynamic phantom study.[15] For all CT systems, a radiation dose reduction of 40% in combination with specific IR settings did not result in significantly different Agatston scores. For one CT vendor, Agatston scores with IR were not affected even at 80% dose reduction. We also showed with a systematic review that this method of dose reduction (tube current reduction with increased IR levels) resulted in low risk reclassification rates.[19]

PCD CT, as currently developed by all major CT manufacturers, is a major technical advancement in CT in reducing image noise.[62–65] This noise reduction characteristic originates from individual photons being counted within predefined energy specific bins. The threshold of the lowest energy bin threshold is placed above the electronic noise. With this, the effect of electronic noise is minimized. In this thesis, it was shown on a small phantom that for 50% reduced tube current acquisitions CAC detection for small CAC improves at a reduced slice thickness of 1 mm by up to 108% and 150% for conventional CT and PCD CT, respectively.[66]

Besides tube current reduction, other dose reduction methods were also assessed. In this thesis, it was shown that a calcium-aware (Ca-aware) reconstruction technique yielded only small differences in Agatston score, when data was acquired at patient-size-specific tube potentials.[57,67] These results were in-line with two studies by Vingiani et al., which also showed a high concordance between Agatston scores with standard tube voltage acquisitions at 120 kVp and patient-size-specific tube voltages reconstructed with the Ca-aware reconstruction technique. Furthermore, Jubran et al. showed excellent correlation (r = 0.99) between Agatston scores of patients scanned with the standard protocol (120 kVp and standard kernel) and a protocol with a patient-specific tube voltage and a Ca-aware reconstruction technique.[68] The patient-specific protocol was associated with a 21% decrease in median volumetric CT dose index.

One important limitation of the work in this thesis is related to the extensive use of current anthropomorphic phantoms for the majority of radiation dose reduction technique studies, as these allow for repeated and systematic assessments. With the advent of reconstruction techniques such as the Ca-aware technique, or for

17

novel PCD CT systems, their use may be limited. The reason for this is that the x-ray attenuation of the used materials in the phantoms should simulate in-vivo tissues for specific tube potentials. For the Ca-aware reconstruction technique, a specific correction factor for the phantom base material was needed to use the international standard cardiac calcification phantom (CCI, QRM).[57,69] In our dynamic phantom setup, which uses a fillable water compartment, we were able to show that the material surrounding the calcium influences the resulting Agatston scores.[67] Ideally, anthropomorphic phantoms should mimic human tissue for all photon energies in the spectrum of the x-ray tube. For conventional CT, human tissues should be mimicked for the applied tube potential and resulting x-ray tube output. For PCD CT this also has to be true for the different energy bins. Currently, the latter is not true for most phantoms in general, and specifically for the before-mentioned CCI phantom, which hampers direct conversion from phantom studies on PCD CT to patient protocols.

Generally, several radiation dose reduction strategies for CAC CT examinations are available. These strategies not only depend on technical advances in CT (i.e. novel reconstruction kernels, detector designs, advanced reconstruction algorithms, etc.), but also on vendor-specific implementations. As parameter settings in CAC software affect the outcome, these settings for calcium scoring should be standardized in order to yield reproducible calcium scores.[70]

Future of CAC scoring

For further optimization of patient risk stratification, the major influencing clinical risk factors for CAC have to be determined first. With the Agatston methodology, a CT number based factor is used to increase scores for high-density structures. Criqui et al. found that high-density CAC appeared to be more stable, and should therefore be associated with a reduced cardiovascular risk.[71] Contrary, Peng et al. showed that individuals with very high Agatston scores (>1000) are at a substantially higher risk for adverse cardiovascular events.[72,73] In addition, spatial distribution of CAC is not assessed with the Agatston methodology. However, diffuse multi-vessel atherosclerosis is indicated to be associated with increased risk for cardiovascular events.[32,74,75]

As stated, technical advances in CT allow for improved CAC detection for both conventional and PCD CT. Potentially, new scoring methods could be derived which benefit from these technical advances, and which are based on the major influencing clinical risk factors for CAC such as spatial distribution of CAC or

17

CAC density. These scoring methods could build on the already existing volume and mass score, which are both related to actual physical quantities and have much better reproducibility.[4,15,29,32,33,56,76]

These new scoring methods could potentially also be assessed from other regularly acquired CT acquisitions, namely coronary CT angiography (CCTA). For CCTA, an iodine contrast medium is administered to the patient to visualize the coronary arteries. For conventional CT, attenuation of iodine and CAC can be comparable depending on the densities of both, which hampers CAC delineation. Moreover, CAC blooming hinders stenosis grade evaluation. For dual layer CT and PCD CT, however, spectral information is always available without the cost of reduced temporal and/or spatial resolution, which allows for material decomposition.[77] With this, CAC-containing voxels could be identified, allowing for so-called calcium-only maps based on physical quantities such as mass or volume. These maps can subsequently be used to quantify CAC, which would provide the prognostic information of CAC assessment, without the need of performing an additional (non-contrast) acquisition. In addition, virtual-non-contrast maps can be generated, which allow for conventional CAC assessment as shown by Nadjiri et al.[78]

Irrespective of the used scoring method, the method itself should be robust. For example, a zero Agatston score has a very high negative predictive value for adverse cardiovascular events, with a decrease of risk for cardiovascular disease from 15% to <7.5% for a zero Agatston score result.[43] For patients with a small but high density CAC, or large but low density CAC, partial volume effects might result in a zero Agatston score, while coronary artery disease may be present.[79] For PCD CT, however, spatial resolution is inherently improved in comparison with conventional CT. This may result in a non-zero Agatston score, as the same CAC might be resolved at increased spatial resolution. The same holds for improvements in other parameters, such as higher temporal resolution or patient-specific CAC scoring parameters. For the latter, the BAS could potentially be used to reduce the conventional CAC scoring threshold of 130HU, to fully exploit the sensitivity of CAC detection of a given CT system. In *Figure 2*, for example, the CAC scoring threshold was reduced to 35HU, which is just above the level which would result in false positives or BAS > 0. With this, the available data in the image is fully exploited, with increased CAC detectability as a result.
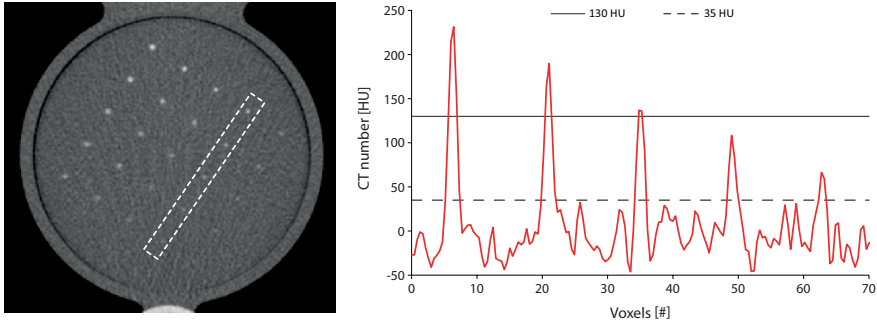
Figure 2 Example image of the D100 (QRM D100, QRM) insert with 25 small calcifications, placed at the center of an anthropomorphic thorax phantom (left), with window level at 90 and window width at 750. CAC detectability increases from 3 to 5 detected CAC by reducing the conventional CAC scoring threshold of 130HU to 35HU as illustrated with a profile plot (right) through five cylindrical calcifications of 1.4 mm in diameter and length, with densities from left to right of 540, 450, 370, 300, and 240 mg hydroxyapatite cm^{-3}

To improve patient risk stratification for adverse cardiovascular events, clinical studies are needed. These studies should be performed on state-of-the-art CT and PCCT systems, with acquisition and reconstruction protocols leading to the highest possible image quality at conventional radiation dose levels. This means that the smallest available detector elements should be used, for acquisitions at both 120 kVp and patient size specific reduced tube potentials. For the included patients, conventional Agatston scores should be calculated on standard CAC scoring reconstructions (120 kVp, 3 mm slice thickness). Next, per calcification Agatston, volume, mass and density (mass / volume) scores should be calculated on reconstructions with sharp reconstruction kernels, DL reconstruction algorithm, and high reconstruction algorithm levels. For dual layer CT and PCCT, CAC density should be calculated with the use of the available spectral data which can result in for example calcium-only maps. Finally, clinical data and follow-up scans should be used to optimize current risk stratification models. For this, coronary heart disease events, atherosclerotic cardiovascular disease events, and mortality should be monitored during a long-term follow-up time of 5 – 10 years.

17

Final conclusion

With recent substantial technical advances in CT, CAC detection can and must be improved, CAC quantification must be made more reproducible, and radiation dose should be reduced. These technical advances include increased spatial resolution for PCD and novel image reconstruction techniques such as DLR. However, current CAC quantification methods do not readily allow for implementation of these technical novelties in clinical protocols. Given the evidence presented in this thesis, I therefore call upon the cardiac CT field to update the current CAC quantification methods, to further improve risk stratification. For this, sensitivity and specificity of current and next-generation CT techniques for CAC detection should be evaluated. Next, major outcome studies are needed, to verify risk stratification improvements based on these novel CAC assessment protocols.

17

REFERENCES

1. Budoff MJ, Achenbach S, Blumenthal RS, et al. Assessment of coronary artery disease by cardiac computed tomography: A scientific statement from the American Heart Association Committee on Cardiovascular Imaging and Intervention, Council on Cardiovascular Radiology and Intervention, and Committee on C. Circulation. 2006;114(16):1761-1791. doi:10.1161/CIRCULATIONAHA.106.178458

2. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. J Am Coll Cardiol. 1990;15(4):827-832. doi:10.1016/0735-1097(90)90282-T

3. Alluri K, Joshi PH, Henry TS, Blumenthal RS, Nasir K, Blaha MJ. Scoring of coronary artery calcium scans: History, assumptions, current limitations, and future directions. Atherosclerosis. 2015;239(1):109-117. doi:10.1016/j.atherosclerosis.2014.12.040

4. van der Werf NR, Willemink MJ, Willems TP, Vliegenthart R, Greuter MJW, Leiner T. Influence of heart rate on coronary calcium scores: a multi-manufacturer phantom study. International Journal of Cardiovascular Imaging. 2017;34(6):959-966. doi:10.1007/s10554-017-1293-x

5. Willemink MJ, Abramiuc B, den Harder AM, et al. Coronary calcium scores are systematically underestimated at a large chest size: A multivendor phantom study. Journal of Cardiovascular Computed Tomography. 2015;9(5):415-421. doi:10.1016/j.jcct.2015.03.010

6. Tigges S, Arepalli CD, Tridandapani S, et al. A phantom study of the effect of heart rate, coronary artery displacement and vessel trajectory on coronary artery calcium score: Potential for risk misclassification. Journal of Cardiovascular Computed Tomography. 2012;6(4):260-267. doi:10.1016/j.jcct.2012.01.005

7. Rutten A, Krul SPJ, Meijs MFL, de Vos AM, Cramer MJM, Prokop M. Variability of coronary calcium scores throughout the cardiac cycle: Implications for the appropriate use of electrocardiogram-dose modulation with retrospectively gated computed tomography. Investigative Radiology. 2008;43(3):187-194. doi:10.1097/RLI.0b013e31815cdd56

8. Funabashi N, Koide K, Mizuno N, et al. Influence of heart rate on the detectability and reproducibility of multislice computed tomography for measuring coronary calcium score using a pulsating calcified mock-vessel in comparison with electron beam tomography. Int J Cardiol. 2006;113(1):113-117. doi:10.1016/j.ijcard.2005.08.053

9. Groen JM, Greuter MJ, Schmidt B, Suess C, Vliegenthart R, Oudkerk M. The influence of heart rate, slice thickness, and calcification density on calcium scores using 64-slice multidetector computed tomography: A systematic phantom study. Investigative Radiology. 2007;42(12):848-855. doi:10.1097/RLI.0b013e318154c549

10. Greuter MJW, Dijkstra H, Groen JM, et al. 64 Slice MDCT generally underestimates coronary calcium scores as compared to EBT: A phantom study. Medical Physics. 2007;34(9):3510-3519. doi:10.1118/1.2750733

11. Kopp AF, Ohnesorge B, Becker C, et al. Reproducibility and accuracy of coronary calcium measurements with multi-detector row versus electron-beam CT. Radiology. 2002;225(1):113-119. doi:10.1148/radiol.2251010173

17

12. Willemink MJ, Vliegenthart R, Takx RAP, et al. Coronary Artery Calcification Scoring with State-of-the-Art CT Scanners from Different Vendors Has Substantial Effect on Risk Classification. Radiology. 2014;273(3):695-702. doi:10.1148/radiol.14140066

13. Detrano RC, Anderson M, Nelson J, et al. Coronary calcium measurements: effect of CT scanner type and calcium measure on rescan reproducibility--MESA study. Radiology. 2005;236(2):477-484. doi:10.1148/radiol.2362040513

14. van der Werf NR, Willemink MJ, Willems TP, Greuter MJW, Leiner T. Influence of iterative reconstruction on coronary calcium scores at multiple heart rates: a multivendor phantom study on state-of-the-art CT systems. International Journal of Cardiovascular Imaging. 2017;0(0):1-11. doi:10.1007/s10554-017-1292-y

15. van der Werf NR, Willemink MJ, Willems TP, Greuter MJW, Leiner T. Influence of dose reduction and iterative reconstruction on CT calcium scores: a multi-manufacturer dynamic phantom study. International Journal of Cardiovascular Imaging. 2017;33(6):899-914. doi:10.1007/s10554-017-1061-y

16. Symons R, Sandfort V, Mallek M, Ulzheimer S, Pourmorteza A. Coronary artery calcium scoring with photon-counting CT: first in vivo human experience. Int J Cardiovasc Imaging. 2019;35(4):733-739. doi:10.1007/s10554-018-1499-6

17. van der Werf NR, Si-Mohamed S, Rodesch PA, et al. Coronary calcium scoring potential of large field-of-view spectral photon-counting CT: a phantom study. European Radiology. 2021;Accepted f. doi:10.1007/s00330-021-08152-w

18. Tesche C, de Cecco CN, Schoepf UJ, et al. CT coronary calcium scoring with tin filtration using iterative beam-hardening calcium correction reconstruction. European Journal of Radiology. 2017;91(December 2016):29-34. doi:10.1016/j.ejrad.2017.03.011

19. Vonder M, van der Werf NR, Leiner T, et al. The impact of dose reduction on the quantification of coronary artery calcifications and risk categorization: A systematic review. Journal of Cardiovascular Computed Tomography. 2018;12(5):352-363. doi:10.1016/j.jcct.2018.06.001

20. Rutten A, Isgum I, Prokop M. Coronary calcification: effect of small variation of scan starting position on Agatston, volume, and mass scores. Radiology. 2008;246(1):90-98. doi:10.1148/radiol.2461070006

21. Schindler A, Vliegenthart R, Schoepf UJ, et al. Iterative Image Reconstruction Techniques for CT Coronary Artery Calcium Quantification: Comparison with Traditional Filtered Back Projection in Vitro and in Vivo. Radiology. 2014;270(2):387-393. doi:10.1148/radiol.13130233

22. Szilveszter B, Elzomor H, Károlyi M, et al. The effect of iterative model reconstruction on coronary artery calcium quantification. International Journal of Cardiovascular Imaging. 2016;32(1):153-160. doi:10.1007/s10554-015-0740-9

23. Takx RAP, Willemink MJ, Nathoe HM, et al. The effect of iterative reconstruction on quantitative computed tomography assessment of coronary plaque composition. Int J Cardiovasc Imaging. 2014;30(1):155-163. doi:10.1007/s10554-013-0293-8

24. Oda S, Utsunomiya D, Nakaura T, et al. The Influence of Iterative Reconstruction on Coronary Artery Calcium Scoring - Phantom and Clinical Studies. Academic Radiology. 2017;24(3):295-301. doi:10.1016/j.acra.2016.11.003

17

25. Schlosser T, Hunold P, Voigtländer T, Schmermund A, Barkhausen J. Coronary artery calcium scoring: Influence of reconstruction interval and reconstruction increment using 64-MDCT. American Journal of Roentgenology. 2007;188(4):1063-1068. doi:10.2214/AJR.05.1369

26. Weininger M, Ritz KS, Schoepf UJ, et al. Interplatform Reproducibility of CT Coronary Calcium Scoring Software. Radiology. 2012;265(1):70-77. doi:10.1148/radiol.12112532

27. Praagh GD, Werf NR, Wang J, et al. Fully automated quantification method (FQM) of coronary calcium in an anthropomorphic phantom. Medical Physics. 2021;48(7):3730-3740. doi:10.1002/mp.14912

28. Hong C, Bae KT, Pilgram TK, Zhu F. Coronary artery calcium quantification at multi-detector row CT: influence of heart rate and measurement methods on interacquisition variability initial experience. Radiology. 2003;228(1):95-100. doi:10.1148/radiol.2281020685

29. Hong C, Bae KT, Pilgram TK. Coronary Artery Calcium: Accuracy and Reproducibility of Measurements with Multi–Detector Row CT—Assessment of Effects of Different Thresholds and Quantification Methods. Radiology. 2003;227(3):795-801. doi:10.1148/radiol.2273020369

30. van Osch JAC, Mouden M, van Dalen JA, et al. Influence of iterative image reconstruction on CT-based calcium score measurements. International Journal of Cardiovascular Imaging. 2014;30(5):961-967. doi:10.1007/s10554-014-0409-9

31. Budoff MJ, Shaw LJ, Liu ST, et al. Long-term prognosis associated with coronary calcification: observations from a registry of 25,253 patients. J Am Coll Cardiol. 2007;49(18):1860-1870. doi:10.1016/j.jacc.2006.10.079

32. Blaha MJ, Mortensen MB, Kianoush S, Tota-Maharaj R, Cainzos-Achirica M. Coronary Artery Calcium Scoring: Is It Time for a Change in Methodology? JACC: Cardiovascular Imaging. 2017;10(8):923-937. doi:10.1016/j.jcmg.2017.05.007

33. Willemink MJ, van der Werf NR, Nieman K, Greuter MJW, Koweek LM, Fleischmann D. Coronary artery calcium: A technical argument for a new scoring method. Journal of Cardiovascular Computed Tomography. 2019;13(6):347-352. doi:10.1016/j.jcct.2018.10.014

34. Groen JM, Greuter MJW, Vliegenthart R, et al. Calcium scoring using 64-slice MDCT, dual source CT and EBT: A comparative phantom study. International Journal of Cardiovascular Imaging. 2008;24(5):547-556. doi:10.1007/s10554-007-9282-0

35. Rutten A, Krul SPJ, Meijs MFL, de Vos AM, Cramer MJM, Prokop M. Variability of Coronary Calcium Scores Throughout the Cardiac Cycle. Investigative Radiology. 2008;43(3):187-194. doi:10.1097/RLI.0b013e31815cdd56

36. Greuter MJW, Groen JM, Nicolai LJ, Dijkstra H, Oudkerk M. A model for quantitative correction of coronary calcium scores on multidetector, dual source, and electron beam computed tomography for influences of linear motion, calcification density, and temporal resolution: A cardiac phantom study. Medical Physics. 2009;36(11):5079-5088. doi:10.1118/1.3213536

37. Zhang Y, van der Werf NR, Jiang B, van Hamersvelt R, Greuter MJW, Xie X. Motion-corrected coronary calcium scores by a convolutional neural network: a robotic simulating study. European Radiology. 2020;30(2):1285-1294. doi:10.1007/s00330-019-06447-7

17

38. Dobrolińska M, van der Werf N, Greuter M, Jiang B, Slart R, Xie X. Classification of moving coronary calcified plaques based on motion artifacts using convolutional neural networks: a robotic simulating study on influential factors. BMC Medical Imaging. 2021;21(1):151. doi:10.1186/s12880-021-00680-7

39. Park J bin, Jeong YJ, Lee G, Lee NK, Kim JY, Lee JW. Influence of Heart Rate and Innovative Motion-Correction Algorithm on Coronary Artery Image Quality and Measurement Accuracy Using 256-Detector Row Computed Tomography Scanner: Phantom Study. Korean J Radiol. 2019;20(1):94-101. doi:10.3348/kjr.2018.0251

40. Balaney B, Vembar M, Grass M, et al. Improved visualization of the coronary arteries using motion correction during vasodilator stress CT myocardial perfusion imaging. Eur J Radiol. 2019;114:1-5. doi:10.1016/j.ejrad.2019.02.010

41. Lee D, Choi J, Kim H, Cho M, Lee KY. Validation of a novel cardiac motion correction algorithm for x-ray computed tomography: From phantom experiments to initial clinical experience. PLoS ONE. 2020;15(9 September):1-18. doi:10.1371/journal.pone.0239511

42. Waltz J, Kocher M, Kahn J, Dirr M, Burt JR. The Future of Concurrent Automated Coronary Artery Calcium Scoring on Screening Low-Dose Computed Tomography. Cureus. 2020;12(6):e8574. doi:10.7759/cureus.8574

43. Blaha MJ, Cainzos-Achirica M, Greenland P, et al. Role of Coronary Artery Calcium Score of Zero and Other Negative Risk Markers for Cardiovascular Disease: The Multi-Ethnic Study of Atherosclerosis (MESA). Circulation. 2016;133(9):849-858. doi:10.1161/CIRCULATIONAHA.115.018524

44. Groen JM, Kofoed KF, Zacho M, Vliegenthart R, Willems TP, Greuter MJW. Calcium score of small coronary calcifications on multidetector computed tomography: Results from a static phantom study. European Journal of Radiology. 2013;82(2):e58-e63. doi:10.1016/j.ejrad.2012.09.018

45. Sandstedt M, Marsh J, Rajendran K, et al. Improved coronary calcification quantification using photon-counting-detector CT: an ex vivo study in cadaveric specimens. European Radiology. 2021;31(9):6621-6630. doi:10.1007/s00330-021-07780-6

46. Takahashi M, Kimura F, Umezawa T, Watanabe Y, Ogawa H. Comparison of adaptive statistical iterative and filtered back projection reconstruction techniques in quantifying coronary calcium. Journal of Cardiovascular Computed Tomography. 2016;10(1):61-68. doi:10.1016/j.jcct.2015.07.012

47. Oda S, Utsunomiya D, Nakaura T, et al. The Influence of Iterative Reconstruction on Coronary Artery Calcium Scoring—Phantom and Clinical Studies. Academic Radiology. 2017;24(3):295-301. doi:10.1016/j.acra.2016.11.003

48. Brenner DJ, Hall EJ. Computed Tomography — An Increasing Source of Radiation Exposure. New England Journal of Medicine. 2007;357(22):2277-2284. doi:10.1056/nejmra072149

49. Ferrero A, Takahashi N, Vrtiska TJ, Krambeck AE, Lieske JC, McCollough CH. Understanding, justifying, and optimizing radiation exposure for CT imaging in nephrourology. Nature Reviews Urology. 2019;16(4):231-244. doi:10.1038/s41585-019-0148-8

50. Fazel R, Krumholz HM, Wang Y, et al. Exposure to low-dose ionizing radiation from medical imaging procedures. N Engl J Med. 2009;361(9):849-857. doi:10.1056/NEJMoa0901249

17

51. Brenner DJ, Hall EJ. Computed Tomography — An Increasing Source of Radiation Exposure. New England Journal of Medicine. 2007;357(22):2277-2284. doi:10.1056/nejmra072149

52. Greenland P, Alpert JS, Beller GA, et al. 2010 ACCF/AHA guideline for assessment of cardiovascular risk in asymptomatic adults: Executive summary: A report of the American College of cardiology foundation/American Heart association task force on practice guidelines. Circulation. 2010;122(25):2748-2764. doi:10.1161/CIR.0b013e3182051bab

53. Hecht H, Blaha MJ, Berman DS, et al. Clinical indications for coronary artery calcium scoring in asymptomatic patients: Expert consensus statement from the Society of Cardiovascular Computed Tomography. Journal of Cardiovascular Computed Tomography. 2017;11(2):157-168. doi:10.1016/j.jcct.2017.02.010

54. Kim KP, Einstein AJ, Berrington de González A. Coronary Artery Calcification Screening: Estimated Radiation Dose and Cancer Risk. Archives of Internal Medicine. 2009;169(13):1188-1194. doi:10.1001/archinternmed.2009.162

55. Ghadri JR, Goetti R, Fiechter M, et al. Inter-scan variability of coronary artery calcium scoring assessed on 64-multidetector computed tomography vs. dual-source computed tomography: A head-to-head comparison. European Heart Journal. 2011;32(15):1865-1874. doi:10.1093/eurheartj/ehr157

56. McCollough CH, Ulzheimer S, Halliburton SS, Shanneik K, White RD, Kalender WA. Coronary Artery Calcium: A Multi-institutional, Multimanufacturer International Standard for Quantification at Cardiac CT. Radiology. 2007;243(2):527-538. doi:10.1148/radiol.2432050808

57. Booij R, van der Werf NR, Budde RPJ, Bos D, van Straten M. Dose reduction for CT coronary calcium scoring with a calcium-aware image reconstruction technique: a phantom study. European Radiology. 2020;30(6):3346-3355. doi:10.1007/s00330-020-06709-9

58. Willemink MJ, de Jong P a, Leiner T, et al. Iterative reconstruction techniques for computed tomography Part 1: Technical principles. Eur Radiol. Published online January 2013. doi:10.1007/s00330-012-2765-y

59. Willemink MJ, Takx RAP, de Jong PA, et al. The impact of CT radiation dose reduction and iterative reconstruction algorithms from four different vendors on coronary calcium scoring. European Radiology. 2014;24(9):2201-2212. doi:10.1007/s00330-014-3217-7

60. den Harder AM, Wolterink JM, Willemink MJ, et al. Submillisievert coronary calcium quantification using model-based iterative reconstruction: A within-patient analysis. European Journal of Radiology. 2016;85(11):2152-2159. doi:10.1016/j.ejrad.2016.09.028

61. Higaki T, Nakamura Y, Zhou J, et al. Deep Learning Reconstruction at CT: Phantom Study of the Image Characteristics. Academic Radiology. 2020;27(1):82-87. doi:10.1016/j.acra.2019.09.008

62. Sandfort V, Persson M, Pourmorteza A, Noël PB, Fleischmann D, Willemink MJ. Spectral photon-counting CT in cardiovascular imaging. Journal of Cardiovascular Computed Tomography. 2020;(September). doi:10.1016/j.jcct.2020.12.005

63. Willemink MJ, Persson M, Pourmorteza A, Pelc NJ, Fleischmann D. Photon-counting CT: Technical principles and clinical prospects. Radiology. 2018;289(2):293-312. doi:10.1148/radiol.2018172656

17

64. Leng S, Yu Z, Halaweish A, et al. Dose-efficient ultrahigh-resolution scan mode using a photon counting detector computed tomography system. Journal of Medical Imaging. 2016;3(4):043504. doi:10.1117/1.jmi.3.4.043504

65. Pourmorteza A, Symons R, Henning A, Ulzheimer S, Bluemke DA. Dose Efficiency of Quarter-Millimeter Photon-Counting Computed Tomography: First-in-Human Results. Invest Radiol. 2018;53(6):365-372. doi:10.1097/RLI.0000000000000463

66. van der Werf N, Rodesch P, Si-Mohamed S, et al. Improved coronary calcium detection and quantification with low-dose full field-of-view photon-counting CT: a phantom study. European Radiology. Published online 2022. doi:10.1007/s00330-021-08421-8

67. Werf NR van der, Booij R, Schmidt B, et al. Evaluating a calcium-aware kernel for CT CAC scoring with varying surrounding materials and heart rates : a dynamic phantom study. European Radiology. 2021;31(12):9211-9220. doi:10.1007/s00330-021-08076-5

68. Jubran A, Mastrodicasa D, van Praagh GD, et al. Low-dose coronary calcium scoring CT using a dedicated reconstruction filter for kV-independent calcium measurements. European Radiology. Published online 2022. doi:10.1007/s00330-021-08451-2

69. Allmendinger T, Hamann A. Agatston score calcium quantification with arbitrary tube voltage [white paper]. Siemens Healthineers. Published 2019. https://www.siemens-healthineers.com/computed-tomography/clinical-imaging-solutions/cardiovascular-imaging#CLINICAL_USE

70. van Ooijen PMA, Vliegenthart R, Witteman JC, Oudkerk M. Influence of scoring parameter settings on Agatston and volume scores for coronary calcification. European Radiology. 2005;15(1):102-110. doi:10.1007/s00330-004-2479-x

71. Criqui MH, Denenberg JO, Ix JH, et al. Calcium Density of Coronary Artery Plaque and Risk of Incident Cardiovascular Events. Jama. 2014;311(3):271. doi:10.1001/jama.2013.282535

72. Peng AW, Dardari ZA, Blumenthal RS, et al. Very High Coronary Artery Calcium (≥1000) and Association With Cardiovascular Disease Events, Non-Cardiovascular Disease Outcomes, and Mortality: Results From MESA. Circulation. 2021;143(16):1571-1583. doi:10.1161/CIRCULATIONAHA.120.050545

73. Adelhoefer S, Uddin SMI, Osei AD, Obisesan OH, Blaha MJ, Dzaye O. Coronary Artery Calcium Scoring: New Insights into Clinical Interpretation-Lessons from the CAC Consortium. Radiol Cardiothorac Imaging. 2020;2(6):e200281-e200281. doi:10.1148/ryct.2020200281

74. Blaha MJ, Budoff MJ, Tota-Maharaj R, et al. Improving the CAC Score by Addition of Regional Measures of Calcium Distribution: Multi-Ethnic Study of Atherosclerosis. JACC Cardiovasc Imaging. 2016;9(12):1407-1416. doi:10.1016/j.jcmg.2016.03.001

75. Tota-Maharaj R, Al-Mallah MH, Nasir K, Qureshi WT, Blumenthal RS, Blaha MJ. Improving the relationship between coronary artery calcium score and coronary plaque burden: addition of regional measures of coronary artery calcium distribution. Atherosclerosis. 2015;238(1):126-131. doi:10.1016/j.atherosclerosis.2014.11.008

76. Callister TQ, Cooil B, Raya SP, Lippolis NJ, Russo DJ, Raggi P. Coronary artery disease: improved reproducibility of calcium scoring with an electron-beam CT volumetric method. Radiology. 1998;208(3):807-814. doi:10.1148/radiology.208.3.9722864

17

77. Alvarez RE, Macovski A. Engergy-selective reconstructions in X-ray computerized tomography. Phys Med Biol. 1976;21(5):733-744.

78. Nadjiri J, Kaissis G, Meurer F, et al. Accuracy of Calcium Scoring calculated from contrast-enhanced Coronary Computed Tomography Angiography using a dual-layer spectral CT: A comparison of Calcium Scoring from real and virtual non-contrast data. PLoS ONE. 2018;13(12). doi:10.1371/JOURNAL.PONE.0208588

79. Osborne-Grinter M, Kwiecinski J, Doris M, et al. Association of coronary artery calcium score with qualitatively and quantitatively assessed adverse plaque on coronary CT angiography in the SCOT-HEART trial. European Heart Journal - Cardiovascular Imaging. Published online September 16, 2021. doi:10.1093/EHJCI/JEAB135

17

CHAPTER 18

Dutch Summary
(Nederlandse samenvatting)

Hart- en vaatziekten zijn wereldwijd de voornaamste doodsoorzaak. De belangrijkste aanleiding van deze ziekten is slagaderverkalking. Dit wordt veroorzaakt door chronische ontsteking met vetafzettingen in de slagaderwand, waarin zich in een later stadium kalk in de vorm van calcium kan ophopen. Met behulp van radiologische beeldvorming is het mogelijk om slagaderverkalking waar te nemen, en de hoeveelheid calcium te kwantificeren. De belangrijkste beeldvormende techniek voor het detecteren en kwantificeren van calcium is computertomografie of CT. Met CT wordt met behulp van röntgenstraling een 3D beeld van het lichaam gemaakt. Hiervoor wordt gebruik gemaakt van de materiaal specifieke verstrooiing en absorptie (of attenuatie) van röntgenstraling. Dat wil zeggen dat de hoeveelheid röntgenstraling die 10 cm weefsel, of 10 cm calcium verlaat verschillend is. De hoeveelheid attenuatie wordt weergegeven met verschillende grijswaarden in een CT-beeld. Als voorbeeld zou 10 cm lucht, zacht weefsel of calcium als respectievelijk zwart, grijs en wit worden weergegeven.

Op CT-beelden van het hart is hierdoor een duidelijk contrast tussen calcium en omliggend weefsels zichtbaar, zoals bijvoorbeeld de hartspier, bloedvaten en bloed. Door dit duidelijke contrast is het mogelijk om calcium te detecteren en kwantificeren. Dit is belangrijk, aangezien de aanwezigheid en hoeveelheid van coronair calcium een goede voorspeller is van toekomstige cardiovasculaire complicaties, zoals bijvoorbeeld een hartinfarct. Voor opeenvolgende CT-scans geeft een oplopende hoeveelheid coronair kalk een verder verhoogde kans op toekomstige cardiovasculaire events.

Voor het bepalen van de aanwezige hoeveelheid calcium in de hartvaten wordt gebruik gemaakt van zogenaamde coronaire kalkscores. In de dagelijkse klinische praktijk wordt hiervoor de Agatston score gebruikt. Deze methode telt alle voxels (volume pixels) boven een bepaalde grijswaarde op, waarna het totaal wordt vermenigvuldigd met een factor die afhangt van de maximale voxel-waarde. Naast de Agatston score bestaan er nog andere methodes voor het kwantificeren van coronair kalk, zoals de volume en massa score. De Agatston score is voor het eerst beschreven in een wetenschappelijk artikel uit 1990 voor beelddata van de elektronenbundel tomograaf (EBT), een voorloper van de CT. In navolging van de Agatston score is in 2007 een internationale standaard voor calcium kwantificatie met CT gepubliceerd. Ondanks de grote vooruitgang in CT-technieken sinds 2007, wordt deze standaard nog altijd in de klinische praktijk gebruikt. De belangrijkste reden hiervoor is de goede voorspellende waarde voor toekomstige cardiovasculaire gebeurtenissen met behulp van de Agatston score. Hiervoor

18

wordt gebruik gemaakt van verschillende factoren die van invloed zijn op de kans van het optreden van cardiovasculaire gebeurtenissen, zoals bijvoorbeeld familiegeschiedenis, rookgedrag en gewicht. De Agatston score wordt gebruikt voor aanpassen van de risicoschatting. Deze voorspellende waarde zou potentieel nog verder kunnen worden verbeterd door gebruik te maken van de nieuwste technologische ontwikkelingen van CT. Een belangrijk argument hiervoor is dat bij gebruik van het huidige protocol de Agatston score een grote spreiding laat zien: twee direct na elkaar uitgevoerde CT-scans kunnen een verschil van wel 7.4% in score geven.

PROEFSCHRIFT

Dit proefschrift had als hoofddoel het evalueren van de diagnostische prestaties van CT voor de detectie en kwantificatie van coronair calcium. Als eerste stap werden pseudovariaties in de detectie en kwantificatie van coronair calcium geïdentificeerd en bestudeerd voor drie categorieën van variabelen: 1) patiënt-specifieke factoren, 2) CT-acquisitie en reconstructie parameters, en 3) nabewerking van CT-beelden. Vervolgens werden diverse recente technische ontwikkelingen op het gebied van CT-technologie onderzocht op de mogelijkheid om de gevonden variabiliteit in coronaire kalkscores te reduceren, en om de voorspellende waarde van de CT-kalkscore te vergroten en tegelijkertijd de stralingsdosis voor patiënten ten gevolge van de gebruikte röntgenstraling te verlagen.

Deel 1 – Patiënt specifieke invloed op kalkdetectie en kwantificatie

Een belangrijke variabele voor de detectie en kwantificatie van coronair calcium is de hartfrequentie tijdens een CT-scan. De snelheid van de kransslagaders varieert tussen de 10 en 30 mm/s afhankelijk van de positie van de ader rondom het hart en de hartfrequentie. Deze beweging van de kransslagaders wordt bij de interpretatie van de CT-beelden vaak ten onrechte verwaarloosd. Doordat er wel degelijk beweging van de vaten is tijdens een CT-scan, levert dit altijd enige bewegingsonscherpte in de CT-beelden op, en daarmee ook in het eventueel aanwezige coronaire calcium. Dit zorgt ervoor dat deze calcificaties groter lijken, wat resulteert in een hogere Agatston score. Voor calcificaties met een hoge dichtheid is dit effect nog groter. Voor een calcificatie met een lage densiteit, daarentegen, kan beweging zorgen voor een verlaging van de kalkscore. Dit komt doordat er een drempel voor grijswaarden van 130 Hounsfield Units (HU) wordt gebruikt voor de detectie van coronair calcium. Door bewegingsonscherpte bij lage dichtheid calcificaties worden de

18

grijswaarden verlaagd, waardoor potentieel minder voxels de drempelwaarde overschrijden, waardoor er een onterechte verlaging van de kalkscore optreedt.

In **hoofdstuk 2** is de invloed van de hartfrequentie op de coronaire kalkscore voor CT-systemen van de vier grote CT-fabrikanten onderzocht met behulp van een dynamisch antropomorf fantoom. Dit fantoom bevatte drie calcificaties van gelijke grootte, maar verschillende dichtheid: laag, gemiddeld en hoog. Het fantoom werd bewogen met snelheden die vergelijkbaar waren met hartfrequenties van <60, 60-75 en >75 slagen per seconde (bpm). Voor elk CT-systeem werd het door de fabrikant aanbevolen kalkscore protocol gebruikt. In vergelijking met de kalkscore bij <60 bpm werden er geen significante verschillen in Agatston score aangetoond voor de calcificaties met lage dichtheid. Voor de calcificaties met gemiddelde en hoge dichtheid werden significante verschillen tot wel 50% voor de verhoogde hartslag aangetoond. Omdat de gevoeligheid van verschillende CT-systemen voor beweging van coronair calcium duidelijk verschillend was concluderen we dat herhaalde CT-scans voor de detectie van coronair calcium daarom bij voorkeur uitgevoerd moeten worden op hetzelfde type CT-systeem, met hetzelfde protocol en vergelijkbare hartfrequentie.

In **hoofdstuk 3 en 4** werd de bewegingsonscherpte geanalyseerd met behulp van kunstmatige intelligentie in de vorm van een deep-learning (DL) netwerk. De bewegingsonscherpte van calcificaties met verschillende groottes en dichtheden werden door het DL-netwerk gecategoriseerd, waarna ze werden vergeleken met een bibliotheek van bekende bewegingsonscherpte. Deze bibliotheek bevatte ook de kalkscore van de specifieke calcificatie zonder beweging. Door het gebruik van deze kalkscore werd het effect van bewegingsonscherpte op de kalkscore ongedaan gemaakt. De categorisering van het DL-netwerk was nauwkeurig in 79 ± 6% van de gevallen. Het percentage Agatston scores groter dan nul steeg van 65% tot 85% bij gebruik van het DL-netwerk. In vergelijking met de Agatston score zonder beweging, daalde de variatie in Agatston score van 38% naar 4% door de correctie van het DL-netwerk. Vervolgens werden drie DL-netwerken getraind met behulp van CT-data van de vier belangrijkste CT-fabrikanten, wederom voor het identificeren van bewegingsonscherpte van calcificaties. Hiervoor werd gebruik gemaakt van reconstructies van scans met de klinische referentie stralingsdosis, en 40% en 80% gereduceerde stralingsdosis. De reconstructies waren zowel conventionele gefilterde terug projectie, of FBP, en minimaal drie niveaus van iteratieve reconstructie (IR). Er zijn scans gemaakt van drie calcificatiedichtheden (laag, midden, hoog) en zeven calcificaties snelheden (0 tot 60 mm/s met stappen van 10 mm/s). De

18

netwerken categoriseerden de bewegingsonscherpte bij een specifieke combinatie van CT-fabrikant en calcificatiedichtheid. De nauwkeurigheid van het classificeren van de bewegingsonscherpte was 90% tot 94% voor de DL-netwerken, met de hoogste nauwkeurigheid voor de DenseNet architectuur. Alle netwerken konden de bewegingsonscherpte met hoge nauwkeurigheid classificeren, onafhankelijk van CT-fabrikant, calcificatie snelheid, stralingsdosis en reconstructie. Een directe klinische toepassing van deze techniek wordt belemmerd door de afwezigheid van een bibliotheek met bewegingsonscherpte van in-vivo calcificaties van verschillende dichtheden met verschillende snelheden of hartslagfrequenties. Wanneer deze bibliotheek wel gemaakt zou worden, zou het mogelijk zijn om meer nauwkeurige kalkscore scans te maken bij elke hartslag.

Deel 2 - Acquisitie and reconstructie parameters

In deel 2 van dit proefschrift werd de invloed van reconstructie en acquisitie parameters op detectie en kwantificatie van coronair calcium onderzocht.

In **hoofdstuk 5** werd een systematische review uitgevoerd over het gevolg van stralingsdosisreductie op de kwantificatie van coronair calcium en de daaropvolgende cardiovasculaire risico inschatting. In totaal werden 28 studies geïncludeerd, waarvan 17 patiënt studies (24 tot 200 patiënten per studie), 10 fantoom studies en 1 studie met zowel fantoom- als patiëntdata. In deze studies werden verschillende technieken gepresenteerd die reductie van stralingsdosis mogelijk maakten zonder gevolgen voor de resulterende kalkscores. Deze technieken waren reductie van de buisspanning, reductie van de buisstroom, IR, additionele filters in de röntgenbundel en verschillende combinaties van deze technieken. In 78% van de studies werd een stralingsdosis reductie van minimaal 50% gerapporteerd, met een risico op her-classificatie tussen 3% en 21%. Vooral specifieke combinaties van buisstroom reductie en IR, of de toepassing van een extra filter in de röntgenbundel kunnen gebruikt worden voor CT-kalkscore protocollen, of bij toekomstige screening-studies voor cardiovasculaire risico inschatting, zonder noemenswaardige impact op de gemeten kalkscore.

In **hoofdstuk 6** werd een technisch argument voor de optimalisatie van kalkscore scans voor nieuwe CT-technologie gepresenteerd. Voor de optimalisatie van detectie van kleine en/of lage densiteit calcificaties werden verlaagde buisspanning acquisities, met IR beeldreconstructie en dunne coupe-dikte voorgesteld. Dit maakt een meer nauwkeurige 3-dimensionale afbeelding van coronair calcium mogelijk, met een verbeterde reproduceerbaarheid voor kwantificatie van coronair

18

calcium. Verlaging van de buisspanning verlaagt ook de stralingsbelasting voor de patiënt. Voorafgaand aan klinische toepassing zou dit protocol geëvalueerd moeten worden binnen een grote observationele studie, met lange follow-up duur. Als alternatief zouden antropomorfe fantoom studies, samen met klinische studies uitgevoerd kunnen worden. Het doel van deze studies is om vast te stellen in welke mate de gevoeligheid voor detectie en kwantificatie van kleine en/of lage densiteit calcificaties zou verbeteren, zonder invloed op andere calcificaties. Onze verwachting is dat de geoptimaliseerde protocollen de reproduceerbaarheid zullen verhogen, met een verbetering van de detectie en kwantificatie van kalkdichtheid, en met mogelijk informatie over de vorm en verspreiding van coronair calcium binnen de kransslagaders.

De impact van IR, als alternatief voor beeldreconstructie met FBP, stond centraal in **hoofdstuk 7**. Hiervoor werden kalkscores bij verschillende hartfrequenties gebruikt, in dezelfde studieopzet als in hoofdstuk 2. De artificiële calcificaties werden bewogen met snelheden van 0, 10, 20 en 30 mm/s, wat correspondeerde met hartfrequenties van respectievelijk 0, <60, 60-75 en >75 bpm. Voor elke hartfrequentie werden de beelden gereconstrueerd met FBP en drie IR niveaus. De calcificaties werd gekwantificeerd als Agatston en massa scores. Bovendien werd een IR susceptibiliteitsindex (IRS) gedefinieerd, waarmee de kans op een verandering in kalkscores ten gevolge van het toepassen van IR kon worden vastgesteld. IR kalkscores werden vergeleken met de referentie FBP-kalkscores. Agatston scores waren gevoeliger voor veranderingen in beeldreconstructie algoritme dan massa scores. Er werd aangetoond dat het effect van IR op kalkscores niet alleen afhing van het CT-systeem of de kalkdichtheid, maar ook van de hartfrequentie. Daarom werd aanbevolen om opvolgende coronair calcium onderzoeken uit te voeren op gelijkwaardige CT-systemen, met hetzelfde reconstructie algoritme, en bij voorkeur met dezelfde hartfrequentie.

Hoofdstuk 8 ging over de invloed van stralingsdosis reductie samen met IR op kalkscores voor alle grote CT-fabrikanten. Wederom werd gebruikt gemaakt van een dynamisch antropomorf fantoom, waarbij een calcificaties werden bewogen met 20 mm/s, overeenkomend met een hartfrequentie van 60-75 bpm. Klinische CT-kalkscore protocollen met high-end CT-systemen van de vier grote CT-fabrikanten werden gebruikt voor data acquisitie. Bovendien werd de stralingsdosis gereduceerd met 40% en 80% ten opzichte van de standaard dosis. Verlaging van de stralingsdosis resulteerde in verhoogde kalkscores. Onder invloed van verhoogde IR niveaus daalde deze kalkscore weer. In vergelijking met kalkscores

18

van de standaard protocollen, werd met alle vier de high-end CT-systemen een vergelijkbare kalkscore gevonden bij 40% stralingsdosisreductie. Voor sommige CT-systemen werd zelfs bij 80% stralingsdosisreductie in combinatie met IR geen verandering in kalkscore gevonden.

Hoofdstuk 9 en 10 beschreven detectie en kwantificatie van coronair calcium voor een nieuw generatie CT-systeem dat gebruik maakte van een nieuw type detector, een photon-counting detector. In hoofdstuk 9 werd als eerste de mogelijkheid van detectie en kwantificatie van coronair calcium voor deze spectrale photon-counting CT (SPCCT) onderzocht, door kalkscores op SPCCT te vergelijken met kalkscores op een conventioneel CT-systeem. Vervolgens werd de verbeterde detectie en kwantificatie van coronair calcium bij reconstructies met gereduceerde coupe-dikte onderzocht. Voor deze onderzoeken werden twee kalkhoudende cilinders in een antropomorf thorax fantoom geplaatst. Om de invloed van de patiënt-grootte te evalueren, werd het thorax fantoom met en zonder een extensiering gescand. Zonder extensiering representeerde het fantoom een kleine patiënt, en met extensiering een grote patiënt. De calcificaties werden gekwantificeerd met Agatston en volume scores. Voor klinische kalkscore protocollen waren de kalkscores vergelijkbaar voor beide CT-systemen. Voor kleine fantomen nam de detectie van coronair calcium toe met 142% en 169% bij gereduceerde coupe-dikte voor respectievelijk CT en SPCCT. Voor grote fantomen nam de detectie toe met 31% en 110%. In vergelijking met CT verbeterde de bepaling van het volume van coronair calcium met SPCCT. Concluderend gaf SPCCT vergelijkbare resultaten voor detectie en kwantificatie van coronair calcium als conventionele CT voor bestaande klinische kalkscore protocollen. Verder resulteerde de verbeterde spatiele resolutie van SPCCT in verbeterde detectie en verbeterde volume schatting bij gereduceerde coupe-dikte.

In het volgende hoofdstuk werd ingegaan op verbeterde detectie en kwantificatie van coronair calcium bij gereduceerde coupe-dikte en gereduceerde stralingsdosis. Hiervoor werden dezelfde opstelling en scaninstellingen als in het vorige hoofdstuk gebruikt. Hiernaast werd data bij 50% stralingsdosis verkregen, door het verlagen van de buisstroom. De calcificaties werden gekwantificeerd als volume en massa scores. Voor zowel gereduceerde stralingsdosis als gereduceerde coupe-dikte nam detectie van coronair calcium voor het kleine fantoom toe met 108% en 150% voor respectievelijk CT en SPCCT. Voor het grote fantoom interfereerde het ruisniveau met kalkdetectie. Voor klinische kalkprotocollen was de mediane afwijking tussen de volume score en het fysieke kalkvolume groot (tot 134%). Voor gereduceerde

18

stralingsdosis en gereduceerde coupe-dikte nam de overschatting van het fysieke volume af tot 96% en 72%, wederom voor respectievelijk CT en SPCCT. In vergelijking met de volume scores was de afwijking tussen de massa score en de fysieke massa kleiner. Concluderend was de SPCCT superieur in detectie van coronair calcium ten opzichte van CT, wat ook werd behouden bij gereduceerde stralingsdosis. Hiernaast was de benadering van het fysieke volume beter voor SPCCT dan voor CT.

Een speciale calcium-aware (Ca-aware) reconstructietechniek stond centraal in de volgende twee hoofdstukken. Voor de standaard kalkscore scans wordt gebruik gemaakt van een vaste buisspanning van 120 kVp. Een verlaging van deze buisspanning zou tot een lagere stralingsdosis kunnen leiden, maar zou ook een verandering van CT-getallen met zich meebrengen. De Ca-aware reconstructietechniek schaalt de CT-getallen van enkel calcium, zodat de waarde van de voxels met calcium overeenkomen met de waarde die verkregen zou zijn met de referentie buisstroom van 120 kVp. Hierdoor is het dus in potentie mogelijk om met patiënt-specifieke buisspanningen te scannen, met als voor een verlaging van de stralingsbelasting voor de patiënt, zonder invloed op de kalkscore.

In **hoofdstuk 11** werden de mogelijkheden tot verdere dosisreductie met deze nieuwe Ca-aware reconstructietechniek onderzocht met behulp van een statisch antropomorf fantoom in drie groottes (klein, middel, en groot). In het fantoom werden twee kalkhoudende cilinders geplaatst, met verschillende groottes en dichtheden van coronair calcium. De buisstroom varieerde tussen de 70 en 150 kVp en er werd een protocol met tin-filter (100Sn kVp) gebruikt. Bij alle scans werd gebruikt gemaakt van automatische buisstroom bepaling. In vergelijking met de referentie scan bij 120 kVp nam de stralingsdosis af met 22%, 15% en 12% voor respectievelijk het kleine, middelgrote en grote fantoom. Het tin-filter protocol gaf een reductie van de stralingsdosis van 55%, 55% en 60% voor deze fantoomgroottes. In vergelijking met de referentie weken de CT-getallen tot 64% en 11% af voor respectievelijk de standaard en Ca-aware reconstructietechniek. De Agatston scores weken tot 40% en 8% af voor deze reconstructietechnieken. Door toepassing van de Ca-aware reconstructietechniek op scans met variërende buisspanning bleven in het algemeen de CT-getallen consistent, en waren de kalkscores vergelijkbaar met de referentie. Dit ging samen met een verlaagde stralingsdosis (tot 22%). De spreiding in CT-getallen was groter voor calcificaties met een lage dichtheid.

18

In **hoofdstuk 12** werd een dynamisch fantoom (0 en 60-75 bpm) gebruikt voor het analyseren van de invloed van de Ca-aware reconstructietechniek op kalkscores van bewegende calcificaties. De bewegende calcificaties waren van verschillende dichtheden en omgeven door verschillende patiënt-equivalente materialen. Ruwe data werden verzameld met drie protocollen: 1. standaard kalkprotocol met 120 kVp, 2. gereduceerde buisspanning van 100 kVp, en 3. een tin-filter protocol met 150Sn kVp. Ruwe data werden gereconstrueerd met zowel de standaard als de Ca-aware reconstructietechniek. Voor elke scan werd het centrale fantoom compartiment gevuld met oplossingen met CT-getallen van vet, water en zacht weefsel. Omdat CT-getallen afhankelijk zijn van de buisspanning, moest het compartiment bij elke verandering van de buisspanning opnieuw gevuld worden. Voor alle hartslagen waren de Agatston scores van de calcificaties in vet vergelijkbaar (<10% afwijking) met de referentie (120 kVp + standaard reconstructie) voor het gereduceerde buisspanning protocol met de Ca-aware reconstructietechniek. Voor lage densiteit calcificaties waren de afwijkingen klinisch relevant (>10% afwijking) wanneer het compartiment gevuld was met water of zacht weefsel. Uit deze studie konden we concluderen dat de Ca-aware reconstructietechniek toegepast kan worden voor patiënt afhankelijke buisspanningen voor calcificaties in een vettige omgeving, zoals in-vivo het geval is, onafhankelijk van de kalkdichtheid.

Hoofdstuk 13 richtte zich op het in kaart brengen van de invloed van vier generaties reconstructietechnieken van één CT-fabrikant op de detectie en kwantificatie van coronair calcium. In het eerste deel van de studie werd met behulp van zowel een statisch als een dynamisch fantoom de invloed van FBP, hybride iteratieve reconstructie (HIR), model-based iteratieve reconstructie (MBIR) en deep-learning reconstructie (DLR) onderzocht. Een statisch fantoom met 100 kleine calcificaties werd gebruikt voor de evaluatie van de detectie van coronair calcium. Een dynamisch fantoom (hartfrequentie 60 – 75 bpm), met lage, gemiddelde en hoge densiteit calcificaties, werd gebruikt voor het evalueren van de Agatston score van coronair calcium. Naast acquisitie met het standaard kalkscore protocol, werd er ook gescand met een gereduceerde stralingsdosis door middel van een verlaging van de buisstroom. In het tweede deel van de studie werden vijftig patiënten gescand met het standaard kalkscore protocol, waarna de data met alle vier de reconstructietechnieken werd gereconstrueerd. Voor alle reconstructietechnieken nam de kalkdetectie in vergelijking met FBP af, behalve voor HIR. De grootste vermindering in detectie van kalk was 22% voor MBIR in vergelijking met FBP bij de standaard stralingsdosis. Ook nam de kalkdetectie af bij een lagere stralingsdosis. Kalkscores van het dynamische fantoom lieten geen klinisch

18

relevante verschillen zien tussen de verschillende reconstructietechnieken bij de standaarddosis. Het onderzoek met patiënten liet een uitstekende overeenkomst zien in Agatston scores tussen FBP, HIR, MBIR en DLR, met slechts een klein aantal her-classificaties van het risico op cardiovasculaire gebeurtenissen. Concluderend liet deze studie zien dat Agatston scores reproduceerbaar zijn voor verschillende generaties reconstructietechnieken, maar dat een andere reconstructietechniek tot een andere kalkdetectie kan leiden.

In **hoofdstuk 14** werd een nieuw scanprotocol voor coronair calcium bestudeerd met als doel verbeterde reproduceerbaarheid van kalkscores bij lagere stralingsdosis voor CT-systemen van de vier grote CT-fabrikanten. Wederom werd gebruik gemaakt van een antropomorf fantoom, met en zonder extensiering, waarin een kalkhoudende cilinder werd geplaatst. Het fantoom werd met klinische kalkscore protocollen gescand op de nieuwste CT-systemen van de vier CT-fabrikanten. Vervolgens werden scans gemaakt met variaties in buisspanning (80 – 120 kVp), buisstroom (100% tot 25% stralingsdosis), coupe-dikte (1 – 3 mm) en reconstructietechniek (FBP en IR). Coronair calcium werd gekwantificeerd met Agatston scores. Inter- en intrascanner variabiliteit werd geëvalueerd met behulp van het inter-kwartiel afstand (interquartile range (IQR)). In vergelijking met de klinische referentie protocollen gaf een protocol met 100 kVp, 75% stralingsdosis, coupe-dikte van 1 of 1.25 mm en hogere IR sterkte een IQR-verlaging van gemiddelde 36%. Het nieuwe protocol liet een toename zien in de detectie van kalk van 6.2±0.4 naar 7.0±0.4 calcificaties. Een paarsgewijze vergelijking tussen CT-systemen van de Agatston score bij dezelfde fantoomgrootte liet drie significante verschillen zien voor het standaard protocol ($p < 0.05$), waarbij het nieuwe protocol geen significant ($p > 0.05$) verschillende kalkscores gaf. We stellen daarom voor om het standard scanprotocol voor coronair calcium te updaten naar een protocol met 100 kVp, een 25% lagere stralingsdosis en dunnere coupes. Voordat dit protocol klinisch in gebruik genomen kan worden is een uitvoerige klinische validatie noodzakelijk. Nadat het protocol ook gevalideerd is op klinische data, zal de klinische inter- en intrascanner reproduceerbaarheid toenemen met een meer consistente risico inschatting en bijbehorend behandeltraject.

18

Part 3 Post-processing parameters

In deel 3 van dit proefschrift werd ingegaan op de invloed van post-processing parameters op de detectie en kwantificatie van coronair calcium.

In **hoofdstuk 15** werd een vendor-neutrale Agatston score (vnAS) gepresenteerd en gevalideerd. Het doel van de vnAS was om de variatie in kalkscores van verschillende CT systemen te verminderen, en daarmee de risico-inschatting voor patiënten met coronaire hartziekte te verbeteren. Bij zeven verschillende CT systemen werd de vnAS bepaald op basis van twee statische antropomorfe kalkfantomen. Hierbij werd de CT kalkscore van alle calcificaties in het fantoom vergeleken met de EBT kalkscore als gouden standaard. Met behulp van een lineaire regressieanalyse werd vervolgens de vnAS bepaald. Voor de validatie van de vnAS werd gebruik gemaakt van de data van 3181 deelnemers aan de Multi-Etnische Studie naar Atherosclerose (MESA). De Chi-kwadraat methode werd gebruikt om een vergelijking te kunnen maken van de cardiovasculaire gebeurtenissen bij de deelnemers van de studie in de lage calcium (vnAS < 100) en de hoge calcium groep (vnAS ≥ 100). Multivariabele Cox proportionele regressiemodellen werden gebruik om het risico op basis van de vnAS te bepalen. Vervolgens werd voor een subgroep van de MESA deelnemers (n = 890) bepaald hoeveel deelnemers met statines behandeld zouden moeten worden (Number Needed to Treat (NNT)). Deze subgroep bestond enkel uit deelnemers van de MESA met een gemiddeld risico op atherosclerotische cardiovasculaire ziekte. Deze deelnemers werden onderverdeeld in twee groepen: groep 1 met een vnAS ≥ 100 en een Agatston score lager dan 100 en groep 2 met een vnAS ≥ 300 en een Agatston score lager dan 300. Voor alle CT systemen werd een sterke correlatie gevonden tussen de CT en EBT Agatston scores ($R^2 > 0.932$). Voor patiënten met een body-mass-index van meer dan 25 kg/m^2 lag de CT Agatston score tussen 88 en 148, en was de vnAS 100. Voor de CT systemen die bij de MESA studie werden gebruikt nam de vnAS toe ten opzichte van de niet-aangepaste Agatston score, wat voor 85 deelnemers resulteerde in een herclassificatie naar de hogere calcium groep (vnAS ≥ 100) en nam het relatieve aantal cardiovasculaire gebeurtenissen significant toe van 7% naar 15% (p = 0.008), met een relatief risico van 3.39 (95% betrouwbaarheidsinterval 1.82 – 6.35, p = 0.001). Binnen de subgroep nam voor groep 1 de NNT af van 12 naar 7 en voor groep 2 van 15 naar 2. Concluderend kunnen we stellen dat het aantal cardiovasculaire gebeurtenissen toenam voor de deelnemers aan de MESA die door de herclassificatie in een hogere calcium groep terecht kwamen. Hiernaast nam het mogelijke voordeel van statine therapie toe voor deelnemers met een gemiddeld risico op atherosclerotische cardiovasculaire

18

ziekte, die door de vnAS in een andere risicogroep waren ingedeeld. Toepassing van een vnAS leidt tot een sterke verbetering in de risicoschatting voor patiënten met mogelijk coronaire hartziekte en kan de behandeling en mogelijk de klinische uitkomst voor patiënten verbeteren.

Tenslotte werd in **hoofdstuk 16** een volledig automatische open source kwantificatie methode (FQM) voor coronair calcium in een internationaal gestandaardiseerd fantoom ontwikkeld en gevalideerd. Zowel standaard parameters, als CT fabrikant-specifieke parameters voor kalkscores werden geïmplementeerd in FQM. Hiernaast werden verschillende beeldkwaliteitsmetingen toegevoegd, waaronder noise-power-spectrum (NPS), task-transfer-function (TTF), contrast-ruis-verhouding (CNR) en signaal-ruis-verhouding (SNR). Bij de implementatie van de fabrikant-specifieke parameters werden verschillen in drie kalkscore parameters aangetoond: de CT-getal drempelwaarde voor de detectie van coronair calcium, het minimale oppervlak van een groep voxels dat werd aangemerkt als coronair calcium, en het gebruik van isotrope voxels voor de volume score. Al deze parameters werden in FQM geïmplementeerd. Omdat FQM en de software van de fabrikanten dezelfde kalkscores lieten zien is niet langer nodig om de kalkscores handmatig te bepalen, maar kan onze volledig automatische methode gebruikt worden. Bovendien geeft FQM informatie over de beeldkwaliteit zodat een directe evaluatie van de invloed van verschillende acquisitie en reconstructie parameters op de resulterende CT-beeldkwaliteit eenvoudig mogelijk wordt.

Concluderend heeft dit proefschrift laten zien dat er substantiële technische vooruitgang is voor CT, waarvan gebruik gemaakt moet worden bij de detectie en kwantificatie van coronair calcium. Toepassing van deze bevindingen in de praktijk zal leiden tot een meer reproduceerbare kalkscore, verkregen bij een lagere stralingsdosis en tot een individueel geoptimaliseerde behandeling.

18

CHAPTER 19

Acknowledgements (*Dankwoord*)
Curriculum Vitae
List of publications

Wetenschap is teamwork. In de afgelopen jaren heb ik dan ook veel mensen leren kennen, die allemaal op hun eigen manier een bijdrage hebben geleverd aan de realisatie van dit proefschrift. Iedereen die zich hierdoor aangesproken voelt, wil ik dan ook hartelijk bedanken. Een aantal mensen wil ik graag in het bijzonder noemen.

Prof. dr. T. Leiner, beste Tim, bedankt voor het zijn van de perfecte promotor voor mijn promotieonderzoek. Jouw positieve energie heeft mijn enthousiasme voor wetenschap in de afgelopen jaren nog verder versterkt. Mijn onderzoek speelde zich af gedurende slechts één dag per week, waardoor de doorlooptijd van projecten soms opliep. Toch wist jij direct tot de kern van de onderzoeken te komen in onze gesprekken, met daardoor altijd nuttige opmerkingen en vragen. Ik verliet onze gesprekken ook steevast met een grote glimlach. En ondanks mijn beperkte tijd heb je me echt onderdeel van het UMCU-team gemaakt. Bijvoorbeeld tijdens de vele congressen, waarvan ik vooral de congres-diners, natuurlijk met bijpassende wijn, nooit zal vergeten.

Dr. M.J.W. Greuter, beste Marcel, dank voor het zijn van veel meer dan mijn copromotor. Jouw aanstekelijke enthousiasme zorgt ervoor dat ik, al sinds ons eerste project in 2012, met heel veel plezier met jou samenwerk. Dat dit project – een simpele stageopdracht van 10 weken – zou uitgroeien tot eerst mijn masteronderzoek en later zelfs dit promotieonderzoek had ik toen niet kunnen bedenken. In deze tijd hebben we niet alleen heel veel contact gehad over kalk-onderzoek, maar ook over van alles buiten dit werk om. Leuke en minder leuke zaken. Eerst ging ons contact nog keurig via reguliere afspraken of e-mails. Inmiddels whatsappen / beeldbellen we wat af, ook in het weekend, 's avonds en vanaf onze vakantiebestemmingen (lees: Ameland).

Dr. M.J. Willemink, beste Martin, dankzij jouw kennis, wetenschappelijke drive en enthousiasme bevat ook mijn proefschrift een groot aantal hoofdstukken (natuurlijk wel 1 minder dan dat van jou). Mijn studies waren vaak groots van opzet, met vele parameters en resultaten. Schrijven is schrappen, en dankzij jouw opmerkingen kwamen we dan ook vaak snel tot de meest interessante kern van het onderzoek. Ook jij had, ondanks alle andere werkzaamheden, altijd tijd voor een goed gesprek over het onderzoek (of randzaken). Je had altijd de behoefte om even op te scheppen over het weer in jouw home-state, Californië. Gelukkig hadden Marcel en ik de kans om dit ook zelf te ervaren, tijdens een briljante road-trip met jou, onder het genot van vele Frappuccinos.

19

Mijn promotieteam (Tim, Marcel, Martin), ik heb met ontzettend veel plezier het onderzoek onder jullie begeleiding mogen uitvoeren. Ik hoop dan ook op nog vele samenwerkingen in de toekomst!

Leden van de beoordelingscommissie, prof. dr. M.L. Bots, prof. dr. N. Buls, prof. dr. ir. H.W.A.M. de Jong, prof. dr. A. van der Lugt en prof. dr. B.K. Velthuis, hartelijk dank voor het lezen en beoordelen van mijn proefschrift. Prof. dr. van der Lugt, beste Aad, ik wil jou graag in het bijzonder bedanken voor alle mogelijkheden en gesprekken tijdens mijn tijd in het Erasmus MC. Jouw bevlogen leiding van de PCCT-meeting heeft geleid tot plaatsing van een van de eerste klinische photon-counting CT systemen ter wereld. En nog veel belangrijker, een sfeer in het ziekenhuis waarin we (radiologen, klinisch fysici, laboranten) gezamenlijk deze nieuwe technologie konden doorgronden. Ik hoop dat onze paden zich nog vaak zullen kruisen.

Mijn paranimfen, dr. Robbert van Hamersvelt en dr. Ronald Booij, zonder jullie klinische kennis en pure gezelligheid had mijn proefschrift er heel anders uitgezien. Robbert, dankzij jou ben ik echt onderdeel geworden van de UMCU-promovendi, en dan vooral van de Vissenkom. Volgens mij is er geen enkele interesse die we niet delen. En dit geldt ook voor onze werkethos: urenlang 100% gefocust werken, afgewisseld met tafeltennis, day-traden of elkaar nieuwe gadgets laten kopen. Inmiddels is er een nieuwe passie bijgekomen, namelijk onze jonge kinderen, die aanleiding geven tot gezellige gezamenlijke play-dates. Ronald, vanaf het moment dat ik begon in het Erasmus MC hebben wij een enorme klik. Natuurlijk hebben wij met CT-onderzoek dezelfde hobby, maar vooral alle laagdrempelige gesprekken waardeer ik enorm. Jouw klinische blik is onmisbaar voor goed en nuttig onderzoek. Gezamenlijk de introductie van de photon-counting CT meemaken in Rotterdam was voor mij dan ook geweldig leerzaam. Ronald en Robbert, ik hoop nog heel vaak met jullie te mogen samenwerken!

Gijs van Praagh, in 2019 zijn wij samen gestart met het volledig automatisch analyseren van fantoomscans in zowel Python als Matlab. Dit project bleek een ware zoektocht, waarin we beide veel aan onze programmeer-skills hebben kunnen werken. Uiteindelijk is het ons gelukt om van alle fabrikanten de kalkscore methodiek te ontrafelen, en hebben we ons ook nog verdiept in verschillende geavanceerde beeldkwaliteitsmaten. Dit was niet gelukt zonder jouw kritische blik en de vele ad-hoc overleggen. Het was een plezier om met je samen te werken!

19

Magdalena Dobrolinska, thank you so much for all your hard work! At our peak, we were working together on more than 10 projects at one time. This resulted – as noted many times by Gijs – in daily virtual meetings. And considering your schedule, this meant 7 days a week, and not limited by 'normal' working hours. In these meetings, we could ask any question to each other, which gave each of us the opportunity to learn more about both the technical and the clinical perspective of coronary calcium assessments. What I will remember most about our cooperation is your persistent optimism, which resulted in many interesting papers!

I would also like to thank the SPCCT-team from Lyon: prof. Philippe Douek, dr. Salim Si-Mohamed, Pierre-Antoine Rodesch, dr. Sara Boccalini, dr. Loïc Boussel and all other colleagues and supporting staff. The opportunity to perform scans on your photon-counting CT was one of the highlights of my PhD. You also immediately included me in the group, allowing me to join in very interesting scientific sessions. Salim, thank you for all the discussions we had. Without your technical and clinical insight, our papers would not have been the same.

Dan zijn er ook verschillende mensen uit het ziekenhuis waar het voor mij begonnen is, het UMCG, die ik graag zou willen bedanken: dr. Marleen Vonder, dr. Tineke Willems, prof. Riemer Slart, prof. Rozemarijn Vliegenthart, dr. Hildebrand Dijkstra, Margo van Gent, Joost Hop, Arjen van Hulzen en vele andere collega's. Dankzij de prettige werkomgeving in het UMCG is mijn stageonderzoek zo van de grond gekomen, en heeft het kunnen uitgroeien tot een volledig promotieonderzoek. Marleen, de systematische review die wij samen hebben geschreven was ontzettend veel werk, maar ook zeer interessant. Ik denk met plezier terug aan ons gezamenlijke bezoek aan Lyon, op de prototype photon-counting CT.

Bronislaw Abramiuc, when I started my internship in 2012, you were involved with coronary calcium scoring for your masters project. As you were already a few weeks into your project, you were able and willing to introduce me to this subject. Taking part in your phantom scans at the UMCU, meant spending 4 hours driving back and forth from Groningen. Our discussion was continuous for the entire drive, which was the first step towards my ability to write this thesis. Thank you for this!

Na het UMCG kwam er voor mij een werkplek in het UMCU, en wel in het leukste kantoor: de (wal)vissenkom! Het kantoor zelf is niet heel inspirerend, maar de mensen die er werken zijn dat des te meer! Robbert, Anouk, Jonas, Wouter, Floor, Nienke, Tim, Maarten, Natascha, dank voor alle adviezen, gesprekken over jonge kinderen, tafeltennismomenten, wandelingen tijdens de lunch en gezelligheid.

19

Aangezien niet iedereen in de vissenkom kon komen werken, wil ik ook graag een groot aantal andere collega's van het UMCU willen bedanken: Fasco van Ommen, Carlo Lucci, Annemarie den Harder, de klinisch fysici, Marja Kool, Carin Marcus – van Geloven en het secretariaat. Naast de inhoud is ook gezelligheid en een fijne werkomgeving essentieel. Jullie hebben hier allemaal aan bijgedragen. Een goed voorbeeld van een geweldige sfeer is de ECR 2019 samen met Fasco, Carlo en Robbert. Tijdens dit congres hebben we, naast alle wetenschappelijke sessies en diners, ook ruimte gemaakt voor een schaatsavontuur op de Rathausplatz. Ook heb ik door jullie onmisbare kennis opgedaan, waardoor ik bijvoorbeeld nog altijd (voorzichtig..) tegen de onderkant van mijn bord met daarop een dampende risotto tik.

Zonder uitvoerige kennis van de verschillende CT systemen, en de inzet voor het scannen van al mijn fantomen, zou mijn onderzoek onmogelijk zijn geweest. Deze kennis en inzet komt meestal van de CT laboranten, met wie ik vaak tot laat in de avond heb gescand. Dank daarom aan iedereen die geholpen heeft bij het maken van de CT scans voor mijn promotieonderzoek in het UMCG, UMCU, Radboud UMC, Erasmus MC, MUMC, LUMC, OLVG, UZ Brussel, en Hôpital Louis Pradel. In het bijzonder wil ik graag Karin Thijn bedanken, voor haar geduld tijdens de eerste scans van mijn dynamische fantoom in het UMCU.

Roy Sanders, dank voor het maken van een schitterend visueel geheel van mijn promotieonderzoek.

Tijdens mijn promotieonderzoek heb ik, naast het UMCU, in vier ziekenhuizen gewerkt, namelijk het Albert Schweitzer ziekenhuis, het Erasmus MC, Holland PTC en het IJsselland ziekenhuis. Dank aan iedereen in deze ziekenhuizen die gedurende de jaren interesse heeft getoond in mijn promotieonderzoek.

Steffie Peters, na ons eerste gezamenlijk project (WAD 2.0 op Schiermonnikoog) spraken wij elkaar pas echt voor het DAQSPECT onderzoek. De zeer prettige samenwerking in dit onderzoek leidde tot veelvuldig contact. Naast fotografie, reizen, of werk gerelateerde zaken ging het vaak over onze gezamenlijk hobby: promotieonderzoek. Onze inhoudelijke onderwerpen lagen ver uit elkaar, maar over alle procedures en randzaken vonden we elkaar vaak. Dank voor alle gegeven adviezen en het zijn van een luisterend oor.

As stated in the first line of this chapter in Dutch: science is teamwork. Without the input from all co-authors on our published papers, these papers would not be the same.

19

Mathijs Lip, Robin Quax, Robin Doddema, Ivar Blomsma, Maikel Wiegertjes, Frank Westerhuis, al sinds de basisschool of middelbare school kennen wij elkaar. Dank voor jullie interesse in mijn onderzoek in de afgelopen jaren!

Leden van de Grasbroekgroep, Rients-Germ, Margot, Martin, Marcel, Martin, Nicole, Maud, Michiel, Peter, Ellenoor, Ronald, Silvia, Maarten en Anouck, ook jullie bedankt voor interesse in mijn werk! De Grasbroek-weekenden zijn altijd een mooie afleiding geweest, ik hoop dat we hier nog lang mee door gaan!

Opa en oma, het eindproduct van mijn promotieonderzoek maken jullie helaas niet meer mee. Maar jullie waren betrokken en geïnteresseerd toen het nog kon. Bedankt!

Marja, Jeroen, Sabine, Jasper en Eline, dank voor jullie steun en interesse tijdens mijn onderzoek. De 'donderdagen' waren nog hard nodig in de laatste jaren om tot afronding te kunnen komen, zonder oppas was dit niet gelukt.

Papa, mama, Laura en Mayjean, zonder jullie had ik dit moment nooit kunnen bereiken, en daar ben ik jullie ontzettend dankbaar voor. Wat ik precies aan het doen was, was voor jullie niet altijd duidelijk, maar jullie steun was onvoorwaardelijk. Bedankt!

En tot slot mijn lieve gezin en thuis: Paulien, Zoë, Joep en Willem. Lieve Paulien, hoewel mijn promotieonderzoek is gestart als hobby, kwam het in de laatste jaren steeds meer neer op avond- en weekendwerk. Zonder offers van jou was afronding dan ook onmogelijk geweest. Dank voor al jouw steun, vertrouwen en liefde. Ik zie enorm uit naar nog meer tijd met jou! Lieve Zoë, hoe leuk ik onderzoek ook vind, meer tijd met jou was de belangrijkste motivatie om mijn onderzoek zo snel mogelijk af te ronden. De donderdagen zijn nu helemaal voor ons (en jouw kleine broertje of zusje)!

19

CHAPTER 19

Acknowledgements (*Dankwoord*)

Curriculum Vitae

List of publications

Niels Roderik van der Werf was born on January 22, 1988 in Groningen. He completed high school at the Dr. Aletta Jacobs college in Hoogezand-Sappemeer in 2006. After one year of Veterinary Medicine at the University of Antwerp, he started his bachelor of Applied Physics at the University of Groningen. In 2013 he continued with his studies with a master of Biomedical Engineering, after which he started as a medical physics resident at the Albert Schweitzer hospital in Dordrecht. Alongside this residency, he started as a PhD student at the Radiology department of the University Medical Center Utrecht in 2017 under supervision of Prof. dr. T. Leiner, Dr. M.J. Willemink, and Dr. M.J.W. Greuter. His research was focused on the diagnostic performance of coronary artery calcium detection and quantification with computed tomography. After his residency, he continued as a Medical Physicist at the Erasmus University Medical Center (Erasmus MC) in 2019. In 2022, he started as a clinical scientist CT at Philips.

19

CHAPTER 19

Acknowledgements (*Dankwoord*)
Curriculum Vitae
List of publications

PEER-REVIEWED PUBLICATIONS

van der Werf N.R.*, Dobrolinska M.M.*, Greuter M.J.W., Willemink M.J., Fleischmann D., Bos D., Slart R.H.J.A., Budoff M. Leiner T. Vendor independent coronary calcium scoring improves individual risk assessment - the Multi-Ethnic Study of Atherosclerosis (MESA). Submitted

Dobrolinska M.M., van Praagh G.D., Oostveen L.J., Poelhekken K., Greuter M.J.W., Fleishmann D., Willemink M.J., de Lange F., Slart R.H.J.A, Leiner T., **van der Werf N.R.** Systematic assessment of coronary calcium detectability and quantification on four generations of CT reconstruction techniques: a patient and phantom study. The International Journal of Cardiovascular Imaging 2022 – in press.

van der Werf N.R.*, Greuter M.J.W.*, Booij R, van der Lugt A, Budde R.P.J., van Straten M. Coronary calcium scores on dual-source Photon Counting Computed Tomography: an adapted Agatston methodology aimed at radiation dose reduction. European Radiology 2022. DOI: 10.1007/s00330-022-08642-5

van der Werf N.R.*, Booij R*, Greuter M.J.W, Bos D, van der Lugt A, Budde R.P.J., van Straten M. Reproducibility of coronary artery calcium quantification on dual-source CT and dual-source Photon-Counting CT: a dynamic phantom study. The International Journal of Cardiovascular Imaging 2022. DOI: 10.1007/s10554-022-02540-z

van der Werf N.R.*, van Gent M.*, Booij R, Bos D, van der Lugt A, Budde R.P.J., Greuter M.J.W, van Straten M. Dose Reduction in Coronary Artery Calcium Scoring Using Mono-Energetic Images from Reduced Tube Voltage Dual-Source Photon-Counting CT Data: A Dynamic Phantom Study. Diagnostics 2021. DOI: 10.3390/diagnostics11122192

van der Werf N.R.*, Rodesch P.A.*, Si-Mohamed S., van Hamersvelt R.W., Greuter M.J.W., Leiner T., Boussel L., Willemink M.J., Douek P. Improved coronary calcium detection and quantification with low-dose full field-of-view photon-counting CT: a phantom study. European Radiology 2021. DOI: 10.1007/s00330-021-08421-8

Dobrolinska M., **van der Werf N.R.**, Greuter M.J.W., Jiang, B., Xie X. Classification of moving coronary calcified plaques based on motion artifacts using convolutional neural networks: a robotic simulating study on influential factors. BMC Medical Imaging 2021. DOI: 10.1186/s12880-021-00680-7

19

van der Werf N.R.*, Si-Mohamed S.*, Rodesch P.A., van Hamersvelt R.W., Greuter M.J.W., Boccalini S., Greffier J., Leiner T., Boussel L., Willemink M.J., Douek P. Coronary calcium scoring potential of Spectral Photon Counting CT: a phantom study. European Radiology 2021. DOI: 10.1007/s00330-021-08152-w

van Praagh G.D., Wang J., **van der Werf N.R.**, Greuter M.J.W., Mastrodicasa D., Nieman K., van Hamersvelt R.W., Oostveen L.J., de Lange F., Slart R.H.J.A., Leiner T., Fleischmann D., Willemink M.J. Coronary Artery Calcium Scoring: an Update of the International Standard. Investigative Radiology 2021. DOI: 10.1097/RLI.0000000000000808

Si-Mohamed S., Greffier J., Miailhes J., Boccalini S., Rodesch P.A., Vuillod A., **van der Werf N.R.**, Dabli D., Racine D., Rotzinger D., Becce F., Yagil Y., Coulon P., Vlassenbroek A., Boussel L., Beregi J.P., Douek P. Comparison of image quality between spectral photon-counting CT and dual-layer CT for the evaluation of lung nodules: a phantom study. European Radiology 2021. DOI: 10.1007/s00330-021-08103-5

Van Praagh G.*, **van der Werf N.R.***, Wang J., van Ommen F., Poelhekken K., Slart R.H.J.A., Fleischmann D., Greuter M.J.W., Leiner T., Willemink M.J.. Fully Automated Quantification Method (FQM) of Coronary Calcium in an Anthropomorphic Phantom. Medical Physics 2021. DOI: 10.1002/mp.14912

van der Werf N.R., Booij R., Schmidt B., Flohr T., Leiner T., de Groen J.J., Bos D., Budde R.P.J., Willemink M.J., Greuter M.J.W. Evaluating a calcium-aware kernel for CT CAC-scoring with varying surrounding materials and heart rates: a dynamic phantom study. European Radiology 2021. DOI: 10.1007/s00330-021-08076-5

Booij R., **van der Werf N.R.**, Budde R.P.J., Bos D., van Straten M. Dose Reduction for CT Coronary Calcium Scoring with a Calcium Aware Image Reconstruction Technique: a Phantom Study. European Radiology 2020. DOI: 10.1007/s00330-020-06709-9

Peters S.M.B.*, Meyer Viol S.L.*, **van der Werf N.R.***, de Jong N., van Velden F.H.P., Meeuwis A., Konijnenberg M.W., Gotthardt M., de Jong H.W.A.M., Segbers M. Variability in lutetium-177 SPECT quantification between different state-of-the-art SPECT/CT systems. EJNMMI Physics 2020. DOI: 10.1186/s40658-020-0278-3

19

Peters S.M.B.*, **van der Werf N.R.***, Segbers, M., van Velden F.H.P., Wierts R., Blokland K.A.K., Konijnenberg M.W., Lazarenko S.V., Visser E.P., Gotthardt M. Towards standardization of absolute SPECT/ CT quantification: a multi-center and multi- vendor phantom study. EJNMMI Physics 2019. DOI: 10.1186/s40658-019-0268-5

Zhang Y., **van der Werf N.R.**, Jiang B., van Hamersvelt R.W., Greuter M.J.W., Xie X. Motion Corrected Coronary Calcium Scores by a Convolutional Neural Network: A Robotic Simulating Study. European Radiology 2019. DOI: 10.1007/s00330-019-06447-7

Vonder M., **van der Werf N.R.**, Leiner, T., Greuter M.J.W., Fleischmann D., Vliegenthart R., Oudkerk M., Willemink M.J. The impact of dose reduction on the quantification of coronary artery calcifications and risk categorization: a systematic review. Journal of Cardiovascular Computed Tomography 2018. DOI: 10.1016/j.jcct.2018.06.001

Willemink M.J., **van der Werf N.R.**, Nieman K., Greuter M.J.W., Koweek L.M., Fleischmann D. Coronary artery calcium: A technical argument for a new scoring method. Journal of Cardiovascular Computed Tomography 2018. DOI: 10.1016/j.jcct.2018.10.014

van der Werf N.R., Willemink M.J., Willems T.P., Greuter M.J.W., Leiner, T. Influence of heart rate on coronary calcium scores: a multi-manufacturer phantom study. The International Journal of Cardiovascular Imaging 2017. DOI: 10.1007/s10554-017-1293-x

van der Werf N.R., Willemink M.J., Willems T.P., Greuter M.J.W., Leiner, T. Influence of iterative reconstruction on coronary calcium scores at multiple heart rates: a multivendor phantom study on state-of-the-art CT systems. The International Journal of Cardiovascular Imaging 2017. DOI: 10.1007/s10554-017-1292-y

van der Werf N.R., Willemink M.J., Willems T.P., Greuter M.J.W., Leiner, T. Influence of dose reduction and iterative reconstruction on CT calcium scores: a multi-manufacturer dynamic phantom study. The International Journal of Cardiovascular Imaging 2017. DOI: 10.1007/s10554-017-1061-y

van Horssen P., Schilham S., Dickerscheid D., **van der Werf N.R.**, Keijzers H., van Almere R., Kuijer J., Peters R., Hofman M. Automated quality control of ultrasound based on in-air reverberation patterns. Ultrasound 2017. DOI: 10.1177/1742271X17733145

19

Willemink M.J., Abramiuc B., den Harder A.M., **van der Werf N.R.**, de Jong P.A., Budde R.P.J., Wildberger J.E., Vliegenthart R., Willems T.P., Greuter M.J.W., Leiner T. Coronary calcium scores are systematically underestimated at a large chest size: A multivendor phantom study. Journal of Cardiovascular Computed Tomography 2015. DOI: 10.1016/j.jcct.2015.03.010
* Authors contributed equally

ACCEPTED CONFERENCE ABSTRACTS

van der Werf N.R., Booij R., Budde R.P.J., Leiner T., Willemink M.J., Greuter M.J.W. CT tube voltage independent reconstruction of coronary calcium scores at varying heart rates: a dynamic phantom study. Society of Cardiovascular CT (SCCT) 2020 – virtual oral presentation

van der Werf N.R., van Praagh G.D., Voet M., Wang J., Fleischmann D., Greuter M.J.W., Leiner T., Willemink M.J. Fully automatic quantification of coronary artery calcium scores: an anthropomorphic phantom analysis. Society of Cardiovascular CT (SCCT) 2020 – virtual oral presentation

van der Werf N.R., Vonder M., van Hamersvelt R.W., Greuter M.J.W., Si-Mohamed S., Douek P., Boussel L., Leiner T., Willemink M.J. Detectability of small coronary calcifications: a comparison between dual-layer and spectral photon-counting CT. ECR 2019 – oral presentation

van der Werf N.R., Vonder M., van Hamersvelt R.W., Willemink M.J., Willems T.P., Greuter M.J.W., Leiner T. Effect of dose and IR on CCS at different heart rates for dual-layer CT. ECR 2019 – oral presentation

van der Werf N.R., Vonder M., van Hamersvelt R.W., Willemink M.J., Willems T.P., Greuter M.J.W., Leiner T. Coronary calcium quantification with novel dual-layer CT: a dynamic phantom study. International Society for CT (ISCT) 2018 – poster presentation

van der Werf N.R., Vonder M., van Hamersvelt R.W., Greuter M.J.W., Si-Mohamed S., Douek P., Boussel L., Leiner T., Willemink M.J. Detectability and quantification of small coronary calcifications: a comparison between DLCT and SPCCT. International Society for CT (ISCT) 2018 – poster presentation

19

van der Werf N.R., Willemink M.J., Willems T.P., Greuter M.J.W., Leiner T. Influence of iterative reconstruction on coronary calcium scores at multiple heart rates. Society of Cardiovascular CT (SCCT) 2017 – Poster presentation

van der Werf N.R., Greuter M.J.W., Willemink M.J., Abramiuc B., Willems T.P., Leiner T. Dose reduction in iterative reconstructed coronary artery CT: a multivendor dynamic phantom study. EFOMP 2014 – Oral presentation

van der Werf N.R., Willemink M.J., Abramiuc B., Vliegenthart R., Willems T.P., Leiner T., Greuter M.J.W. Influence of motion and dose on calcium scores in IMR reconstructed CT: a dynamic phantom study. ESCR 2014 – Oral presentation

van der Werf N.R., Willemink M.J., Abramiuc B., Vliegenthart R., Willems T.P., Greuter M.J.W., Leiner T. Influence of motion on calcium scores in iterative reconstructed CT: a multivendor dynamic phantom study. ESCR 2014 – Oral presentation

van der Werf N.R., Willemink M.J., Abramiuc B., Vliegenthart R., Willems T.P., Greuter M.J.W., Leiner T. Overestimation of calcium scores in low dose CT: influence of iterative reconstruction. ESCR 2014 – Oral presentation

van der Werf N.R., Willemink M.J., Abramiuc B., Vliegenthart R., Willems T.P., Leiner T., Greuter M.J.W. Influence of motion and dose on calcium scores in IMR reconstructed CT: a dynamic phantom study. Radiological Society of North America (RSNA) 2014 – Oral presentation

van der Werf N.R., Willemink M.J., Abramiuc B., Willems T.P., Leiner T., Greuter M.J.W. Influence of motion on calcium scores in iterative reconstructed CT: a multivendor dynamic phantom study. Radiological Society of North America (RSNA) 2014 – Oral presentation

van der Werf N.R., Willemink M.J., Abramiuc B., Willems T.P., Leiner T., Greuter M.J.W. Influence of dose reduction on calcium scores in iterative reconstructed CT: a multivendor dynamic phantom study. Radiological Society of North America (RSNA) 2014 – Oral presentation

19