Published in final edited form as:

J Am Coll Radiol. 2015 August ; 12(8): 783–787. doi:10.1016/j.jacr.2015.03.011.

A Default Normal Chest CT Structured Reporting Field for Coronary Calcifications does not Cause Excessive False Negative Reporting

William R. Walter, MD,

Montefiore Medical Center, Department of Radiology, Bronx, NY

Shlomit Goldberg-Stein, MD,

Montefiore Medical Center, Albert Einstein College of Medicine, Department of Radiology, Bronx, NY

Jeffrey M. Levsky, MD, PhD,

Montefiore Medical Center, Albert Einstein College of Medicine, Department of Radiology, Bronx, NY

Hillel W. Cohen, DrPH, MPH, and

Albert Einstein College of Medicine, Department of Epidemiology & Population Health, Bronx, NY

Meir H. Scheinfeld, MD, PhD

Montefiore Medical Center, Albert Einstein College of Medicine, Department of Radiology, Bronx, NY

William R. Walter: wwalter@montefiore.org; Shlomit Goldberg-Stein: sgoldberg@montefiore.org; Jeffrey M. Levsky: jlevsky@montefiore.org; Hillel W. Cohen: hillel.cohen@einstein.yu.edu; Meir H. Scheinfeld: mscheinf@montefiore.org

Abstract

Objective—To compare the accuracy of coronary atherosclerosis reporting before and after implementing a structured reporting chest CT template.

Materials and Methods—IRB exemption was obtained. A non-cardiac, non-contrast chest CT structured reporting template was developed and mandated for department-wide use at our large academic center. The template included the statement, “There are no coronary artery calcifications.” We retrospectively collected all non-cardiac, non-contrast chest CT exams reported over 3 days, one month after template implementation (structured template group), and from a 3-day period one year prior (control group). Final radiology reports were reviewed and designated positive or negative for coronary calcifications. CT images were reviewed in consensus

Corresponding author: Meir Scheinfeld MD PhD, Montefiore Medical Center, Department of Radiology, 111 East 210 Street, Bronx, NY 10467, mscheinf@montefiore.org, Phone: 917-865-3297, Fax: 718-920-4854.

Institutional review board (Albert Einstein College of Medicine) exemption was obtained for this retrospective HIPAA compliant study.

The authors have no conflict of interest.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

by two radiologists, who scored each case for the presence or absence of coronary calcifications, blinded to the original report. Statistical analysis was performed using Pearson's chi-square and Fisher's exact tests.

Results—65% (69/106) of structured template group and 58% (62/106) of control group cases had coronary calcifications. Reports from the structured template group were more likely to correctly state the presence or absence of coronary atherosclerosis compared to the control group (96.2% vs. 85.8%, OR 4.2, 95% CI 1.3, 13.1, $p=0.008$). Structured template group reports were less likely to be falsely negative compared to the control group (3.8% vs. 11.7%, OR 3.4, 95% CI 1.0, 10.8, $p=0.03$).

Conclusion—Implementing a structured reporting template improves reporting accuracy of coronary calcifications.

Introduction

Coronary artery disease (CAD) continues to be a leading cause of death in the United States and internationally [1]. CAD may be present in asymptomatic people, even early in life [2]. It has recently been shown that coronary artery calcification on non-cardiac CT is under-reported [3]. This is despite the validation of calcium scoring on non-cardiac gated exams [4] and when using visual (rather than automated) calcium scoring [5]. Considering that coronary calcium scoring is a powerful risk-stratification tool [6], if under-reported, it may represent a missed opportunity to trigger diagnosis and management of cardiovascular disease. Conversely, reporting of such an incidental finding may trigger evaluation for CAD, allowing for early lifestyle modification or pharmacotherapy to reduce the risk of cardiac events.

Structured radiology reporting templates have become more practical with progressive adoption of voice-recognition radiology-reporting software packages [7]. A structured report allows for elements in a potentially long report to be described in a specific order, and ensures that certain findings (whether present or absent) will be recorded consistently. Our department recently implemented a department-wide structured reporting initiative. We hypothesized that by incorporating a field for coronary calcifications into the new structured non-contrast chest CT template, we would increase accurate reporting of coronary artery calcifications.

Methods

Case Selection

Institutional review board exemption was obtained for this retrospective, HIPAA compliant study. A structured (non-cardiac) non-contrast chest CT template was developed collaboratively using a process similar to that previously described by Larson et al. [8]. This template was disseminated along with other chest templates to all radiologists in the department, without any focused training regarding its use. The structured chest CT template includes both normal, default standard terminology which auto-populates the report, and alternative standard picklist options for use when abnormal findings are present (terminology is listed in Table 1). All non-contrast chest CT cases from a three day period

(July 29-31, 2014) one month after adoption of the structured template were reviewed (structured template group). As a control group, all non-contrast chest CT cases from a three day period one year prior (when radiologists used variable dictation methods at their discretion such as free text, self-developed structured template, or structured template “cloned” from another radiologist) were used. We chose control studies from this time period because we aimed to match the imaging volume which varies with the time of year, and at an academic medical center there is potential considerable variation of report quality with the resident and fellow training cycle and we aimed to eliminate this potential difference.

Case Review and Data Extraction

All chest CT final radiology reports were reviewed; those indicating any coronary artery calcifications were deemed positive while those indicating no coronary artery calcifications and those without mention of coronary calcifications were deemed negative. To establish a gold standard, the 2-3 mm axial soft tissue algorithm images available in PACS (picture archiving and communication system) were reviewed in consensus by two of the authors (WRW and MHS) who were blinded to the final radiology report and clinical data. Where 2-3 mm axial images were not available in PACS to the original reader, the thinnest available axial images were utilized in our consensus review (5mm axial images, three template group cases and nine control group cases, and 0.625 mm axial images one template group case and two control group cases).

Uncertain cases were judged as negative because in routine practice we would not expect the radiologist to diagnose coronary calcifications if there was considerable uncertainty. For each case the following parameters were recorded: age and gender of the patient, number of detector rows of the scanner, low dose versus standard dose technique, whether or not the case was interpreted by a fellowship trained cardiothoracic radiologist, and whether or not a trainee (resident or fellow) initially dictated the case. Note that all reports were ultimately edited as necessary and signed by an attending radiologist. All data were tabulated in a Microsoft (Redmond, WA) Excel spreadsheet.

Statistical Analysis

The authors' consensus review of CT images was considered the gold standard interpretation for each case. The final radiology report was marked as correct if it agreed with the consensus review and incorrect if it disagreed. Results were reported as proportions and associations assessed with Pearson's chi-square or Fisher's exact tests if expected values in > 20% of cells were <5. Magnitude of association was estimated with odds ratios (OR) and 95% confidence intervals (CI). P values < 0.05 were considered statistically significant. Statistical analyses were performed with SPSS for Windows (Version 20).

Results

For the 2014 time period (structured report group), 107 cases were identified, and for the 2013 time period, 106 cases were identified (control group). One of the 2014 cases was eliminated due to an unintelligible statement in the coronary artery calcification field, likely

due to a voice recognition error. (For all of the department's templates, radiologists are encouraged to use choices from picklists, but are able to populate the field with their own text if they feel it is more appropriate.)

The two groups were significantly different with respect to number of MDCT detector rows and the proportion of cases with reports drafted by trainees (Table 2). The two groups did not differ significantly with respect to the proportion with normal versus low dose scan technique or whether the case was read by a fellowship-trained cardiothoracic radiologist versus other radiologist (Table 2).

Based on the gold standard consensus scoring, 65% (69/106) of structured reporting patients and 58% (62/106) of control patients had coronary artery calcifications. Characteristics of correct/incorrect official reports are listed in Table 3. The main analyses compared cases with coronary calcifications with the original final report. The structured template group was more likely to correctly report the presence or absence of coronary atherosclerosis compared to the control group (96.2% vs. 85.8%, OR 4.2, 95% CI 1.3, 13.1, $p=0.008$).

Incorrect reports included false negative (FN) and false positive (FP) cases (Table 4). All of the four incorrect reports of the structured report group were FN. For the control group, 12 of the 15 incorrect reports were false negatives and three were FP. To gauge the performance of the structured template to assist the reading radiologist in reducing the number of cases where coronary calcifications were present but not reported, we compared FN reports to correct reports in the two groups. The template group was less likely, to report a false negative as compared to the control group (3.8% vs. 11.7%, OR 3.4, 95% CI 1.0, 10.8, $p=0.03$).

Discussion

Structured radiology reporting promised to improve report quality as structures are described in a regular and predictable order and there is a decrease of missense-omission errors (words omitted from dictations which ultimately change the intended meaning) [11,9]. Indeed, a recent study has demonstrated that structured chest radiograph reports are more complete and effective than unstructured reports [10]. Furthermore, fields for entry of specific data within a structured report have the potential to direct radiologists' search patterns and facilitate consistent reporting of findings that would not otherwise have elicited specific comment if the radiologist was using free-style dictation [8,11,12,13]. Radiologists tasked with creation of standardized reports for an organization may incorporate the latest thinking on value-adding imaging data and suggest items for incorporation. For coronary calcifications as an example, non-cardiothoracic radiologists may not be aware that qualitative reporting of coronary calcifications has been validated for its prognostic value [5] and may not report on this finding if not prompted by a data entry field.

In this study, we compared reports generated in a 3 day period one month following dissemination of the non-contrast chest CT structured template, to reports generated one year previously (when radiologists used variable dictation methods). Our main finding was that following the implementation of structured reporting and introduction of a coronary

calcification field, accuracy of reporting coronary calcification increased compared to the control time period.

An error of retention occurs when a normal statement in a structured template is erroneously “retained” in the final report because it appeared by default, rather than due to active assessment by the radiologist. In our development of this template, there was concern for accidental retention of the default statement of “No coronary artery calcifications,” when there were indeed coronary calcifications present. Our data suggests that although there were some false negative cases (either due to an error of retention or misdiagnosis), they were statistically less common with the structured template when compared to the control group. Our findings were different than the report of Hawkins et. al. which noted that a default normal appendix statement led to “errors of commission” (an unnoticed template statement which results in internally contradictory statements within the final report) which necessitated changes in their template [11]. A possible explanation for the difference is that in our template there was a completely separate sentence for coronary calcifications, which was easily and consistently seen by users and removed if necessary. In the appendix scenario however, Hawkins et al. describe that the appendix was mentioned as “part of a large grouping of pertinent negatives,” suggesting that its presence was less noticeable.

There were no FP cases in the structured reporting group but there were three in the control group. Due to the retrospective nature of the study it is not possible to determine with certainty the cause of these FP cases in the control group. As calcifications were not present even on the thinnest images available on PACS, we do not believe that this was due to different images being reviewed by the original radiologists and the readers in our study. We have two possible explanations. One possibility is that these FP cases were due to voice recognition errors, which strengthens the contention that a structured reporting template with easily selectable common pick-list options helps to eliminate dictation errors. Alternatively, the original radiologists may have misinterpreted mitral annular calcifications, aortic root calcifications or surgical material as representing coronary artery calcifications.

There are several limitations of our study. First, we scored non-mention of coronary calcifications in official reports of both the template and control group as the radiologist stating that there were no coronary calcifications. It may be argued that such scoring would be inappropriate since not reporting an “incidental” abnormality would not be equivalent to actively stating that it was not there. Furthermore, perhaps not mentioning coronary calcifications would not be tantamount to giving the patient a clean coronary bill-of-health so such non-mentions would be acceptable, whereas an error of retention, even when rare would be less acceptable as it is misleading the clinician into thinking that the patient had no evidence of coronary atherosclerosis. Still, considering that the point of our study was that coronary calcifications are an important prognostic finding rather than an incidental finding (perhaps more important than other incidental findings such as small lung nodules or non-specific mediastinal lymph nodes) we considered it acceptable to consider a non-mention as a missed opportunity for detection of coronary calcifications.

Second, there was a significantly larger number of studies in the structured reporting group performed on 64 detector CT scanners and a significantly larger number of studies in the

control group that had trainee involvement (Table 2). Still, since these characteristics were not significantly associated with both the group (exposure) and the proportion of correctly reported calcification (outcome), it is unlikely that these differences confound our results (Table 3). The observation that low radiation dose technique exams were significantly associated with inaccurate reporting of coronary calcification is an interesting finding, perhaps grounds for follow up research. Finally, we only looked at a sample of 3 days of reports one year prior to our intervention. While there is no reason to suspect that these calendar days were not representative of any days in the year, we cannot be sure that they were. Since the 3 days for the structured template group were only a month after implementation, it is not unreasonable to suspect that the number of incorrect reports may have decreased in later months, as radiologists became more familiar with the new templates.

In conclusion, our study demonstrates that a coronary artery calcification statement in a structured report template improves consistency of reporting of the presence or absence of coronary calcifications. Furthermore, such a statement increases accuracy of reporting and does not result in more false negative errors as compared with free-style or non-structured dictation methods.

Acknowledgments

This publication was supported (for H.W.C.) in part by the CTSA Grant 1 UL1 TR001073-01, 1 TL1 TR001072-01, 1 KL2 TR001071-01 from the National Center for Advancing Translational Sciences (NCATS), a component of the National Institutes of Health (NIH), Bethesda, MD.

References

1. Mozaffarian, D.; Benjamin, EJ.; Go, AS., et al. Circulation. 2014 Dec 17. Heart Disease and Stroke Statistics-2015 Update: A Report From the American Heart Association. Epub ahead of print
2. Tuzcu EM, Kapadia SR, Tutar E, Ziada KM, Hobbs RE, McCarthy PM, Young JB, Nissen SE. High prevalence of coronary atherosclerosis in asymptomatic teenagers and young adults: evidence from intravascular ultrasound. Circulation. 2001 Jun 5; 103(22):2705–10. [PubMed: 11390341]
3. Williams KA, Kim JT, Holohan KM. Frequency of unrecognized, unreported, or underreported coronary artery and cardiovascular calcification on noncardiac chest CT. J Cardiovasc Comput Tomogr. 2013 May-Jun;7(3):167–72. [PubMed: 23849489]
4. Xie X, Zhao Y, de Bock GH, de Jong PA, Mali WP, Oudkerk M, Vliegenthart R. Validation and prognosis of coronary artery calcium scoring in nontriggered thoracic computed tomography: systematic review and meta-analysis. Circ Cardiovasc Imaging. 2013 Jul; 6(4):514–21. [PubMed: 23756678]
5. Jairam PM, Gondrie MJ, Grobbee DE, Mali WP, Jacobs PC, van der Graaf Y. PROVIDI Study Group. Incidental imaging findings from routine chest CT used to identify subjects at high risk of future cardiovascular events. Radiology. 2014 Sep; 272(3):700–8. [PubMed: 24865309]
6. Yeboah J, McClelland RL, Polonsky TS, Burke GL, Sibley CT, O'Leary D, Carr JJ, Goff DC, Greenland P, Herrington DM. Comparison of novel risk markers for improvement in cardiovascular risk assessment in intermediate-risk individuals. JAMA. 2012 Aug 22; 308(8):788–95. [PubMed: 22910756]
7. Boland GW. Voice recognition technology for radiology reporting: transforming the radiologist's value proposition. J Am Coll Radiol. 2007 Dec; 4(12):865–7. [PubMed: 18047977]
8. Larson DB, Towbin AJ, Pryor RM, Donnelly LF. Improving consistency in radiology reporting through the use of department-wide standardized structured reporting. Radiology. 2013; 267(1): 240–250. [PubMed: 23329657]

9. Schwartz LH, Panicek DM, Berk AR, Li Y, Hricak H. Improving communication of diagnostic radiology findings through structured reporting. *Radiology*. 2011 Jul; 260(1):174–81. [PubMed: 21518775]
10. Marcovici PA, Taylor GA. Journal Club: Structured radiology reports are more complete and more effective than unstructured reports. *AJR Am J Roentgenol*. 2014 Dec; 203(6):1265–71. [PubMed: 25415704]
11. Hawkins CM, Hall S, Zhang B, Towbin AJ. Creation and Implementation of Department-Wide Structured Reports: An Analysis of the Impact on Error Rate in Radiology Reports. *J Digit Imaging*. 2014 Oct; 27(5):581–7. [PubMed: 24859725]
12. Hawkins CM, Hall S, Hardin J, Salisbury S, Towbin AJ. Pre-populated radiology report templates: a prospective analysis of error rate and turnaround time. *J Digit Imaging*. 2012 Aug; 25(4):504–11. [PubMed: 22270786]
13. Johnson AJ, Chen MY, Zapadka ME, Lyders EM, Littenberg B. Radiology report clarity: a cohort study of structured reporting compared with conventional dictation. *J Am Coll Radiol*. 2010; 7:501–506. [PubMed: 20630384]

Table 1
Default and picklist statements for the heart and vessels

Default heart and vessel line:

HEART AND VESSELS: The heart is normal in size. There are no coronary artery calcifications. There is no pericardial effusion.

Picklist options for the second sentence:

There are no coronary artery calcifications. (default)

There are coronary artery calcifications.

There are aortic calcifications.

There are aortic and coronary artery calcifications.

Table 2
Characteristics of Structured and Control Groups

	Structured Group (Post template implementation)	Control Group (Pre template implementation)	P value
Cases	106	106	
Gender (males)	50%	51%	
Age (years, mean±SD)	63±16	64±17	
Number of scanner detector rows			0.01
64	74%	87%	
16	26%	12%	
4	0%	1%	
Low dose technique	36%	36%	>0.99
Read by cardiothoracic radiologist	68%	56%	0.07
Trainee involved	24%	39%	0.02

Table 3
Characteristics of correct/incorrect official reports

	Correct*	Incorrect*	P value
Cases	193	19	
Number of scanner detector rows			0.94
64	80%	79%	
16	19%	21%	
4	1%	0%	
Low-dose technique	34%	58%	0.04
Read by cardiothoracic radiologist	62%	58%	0.71
Trainee involved	31%	37%	0.57

* "Correct" is defined as the official report matching the consensus coronary calcification determination (true positives and true negatives).

"Incorrect" includes both false positives and false negatives for official reports compared to consensus coronary calcification determination. For example, 31% of correct reports had a trainee involved and 37% of incorrect reports had a trainee involved, with a non-statistically significant difference between as indicated by a p value of 0.57.

Table 4
Accuracy of Structured Reporting Group vs. Control Group for Reporting of Coronary Atherosclerosis on non-cardiac Chest CT

	Structured Report Group	Control Group
All reports	106	106
Accurate report	102 (96%)	91 (86%)*
FN	4 (4%)	12 (11%)
FP	0	3 (3%)

* Fisher's Exact Test p= 0.02