J Digit Imaging (2013) 26:843–849
DOI 10.1007/s10278-013-9597-4

Content Analysis of Reporting Templates and Free-Text Radiology Reports

Yi Hong · Charles E. Kahn Jr.

Published online: 4 April 2013

Abstract The Radiological Society of North America (RSNA) has developed a set of templates for structured reporting of radiology results. To measure how much of the content of conventional narrative ("free-text") reports is covered by the concepts included in the RSNA reporting templates, we selected five reporting templates that represented a variety of imaging modalities and organ systems. From a sample of 8,275 consecutive, de-identified radiology reports from an academic medical center, we identified one corresponding imaging procedure code for each reporting template. The reports were annotated with RadLex and SNOMED CT terms using the BioPortal Annotator web service. The reporting templates we examined accounted for 17 to 49 % of the concepts that actually appeared in a sample of corresponding radiology reports. The findings suggest that the concepts that appear in the reporting templates occur frequently within free-text clinical reports; thus, the templates provide useful coverage of the "domain of discourse" in radiology reports. The techniques used in this study may be helpful to guide the development of reporting templates by identifying concepts that occur frequently in radiology reports, to evaluate the coverage of existing templates, and to establish global benchmarks for reporting templates.

Keywords Radiology · Reporting · Structured reports · Narrative (free-text) reports · Reporting templates · Biomedical ontologies · RadLex · SNOMED CT · BioPortal Annotator

Y. Hong
Department of Research and Development, Harris Corporation—Healthcare Solutions, Milwaukee, WI, USA

C. E. Kahn Jr. (✉)
Department of Radiology, Medical College of Wisconsin, 9200 W. Wisconsin Ave, Milwaukee, WI 53226, USA
e-mail: kahn@mcw.edu

Introduction

Despite remarkable advances in medical imaging technologies, the form and content of radiology reports has changed relatively little since the inception of radiology [1]. Unstructured ("free text") radiology reports remain the most common approach for radiology reporting. Structured radiology reports present information in a consistent format, employ standardized terminology, and allow reported information to be extracted efficiently for indexing and reuse [1]. Although some technological challenges have yet to be overcome [2], referring physicians have a strong preference for structured radiology reports [3–6]. In specialty areas such as cardiovascular imaging, policy statements have signaled a move to structured reporting [7].

To promote structured radiology reporting, the Radiological Society of North America (RSNA) has developed a large, freely accessible online library of radiology reporting templates (http://www.radreport.org) [8]. Radiologists and other users can browse, retrieve, and download templates in text format or encoded in the Extensible Markup Language. An application programming interface allows one to search template metadata and download reporting templates as a web service. Because information in a structured reporting template adheres to a consistent format and vocabulary, it is easier to integrate that information with generalized knowledge-based resources and incorporate the

structured reporting process with clinical guidelines and decision support.

As of January 2013, the RSNA report template library contained 200 reporting templates in English and 45 templates translated into several other languages. The templates are intended to serve as examples of "best practice" to guide radiologists in formulating reports [8]. Each reporting template has associated metadata, including information about the template's title, creator, subject, description, and date. The elements of the reporting templates have been mapped to corresponding terms in standardized biomedical ontologies such as the RadLex® radiology lexicon and the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT®) vocabulary [9].

RadLex® is a unified language of radiology terms for standardized indexing and retrieval of radiology information resources [10, 11]. RadLex® has more than 34,000 terms, including diseases, radiologically pertinent anatomy, and imaging observations. It is organized as an ontology and includes subsumption ("is-a") relationships to demarcate superclass–subclass relations among its terms. SNOMED CT®, an ontology of more than 310,000 terms, is considered to be the most comprehensive, multilingual clinical healthcare terminology in the world [12, 13]. SNOMED CT® is widely used in clinical information systems.

The National Center for Biomedical Ontology (NCBO) BioPortal web site (http://bioportal.bioontology.org) provides access to an open repository of biomedical ontologies via web services, and allows users to browse, search, and visualize ontologies. It contains ontologies that cover a broad range of biomedical topics, including anatomy, phenotypes, experimental conditions, imaging, chemistry, and health. BioPortal allows users to utilize ontologies for annotation of biomedical data on their sites in order to facilitate interoperability, search, and translational discoveries. Both RadLex® and SNOMED CT® can be accessed through NCBO BioPortal.

Standardized terminologies are used to reduce ambiguity and improve the clarity of radiology reports and image annotations, and provide a uniform means of indexing radiological materials in a variety of settings [9]. In this study, we sought to evaluate how well the RSNA reporting templates covered the "domain of discourse" of actual radiology reports. We measured how frequently terms from the reporting templates appeared in conventionally dictated, narrative ("free-text") radiology reports. Our hypothesis was that the reporting templates included the more frequently used terms in clinical radiology reports.

Material and Methods

Five reporting templates for frequently performed procedures (computed tomography (CT) brain, chest X-ray, magnetic resonance (MR) imaging spine, nuclear medicine (NM) bone scan, and ultrasonography (US) abdomen) were chosen from the RSNA reporting template library. The templates represented a variety of imaging modalities, such CT, MR imaging, radiography, NM, and US, and a variety of body areas and organ systems (thorax, brain, abdomen, and skeletal system). To identify the concepts that appeared in the templates, we extracted the reporting elements from each template. These reporting elements—terms such as "left kidney," "hydronephrosis," and "mild"—described potential report content that could appear in encoded form within a structured report based on the specified template.

A series of 8,275 consecutive, de-identified radiology reports from an academic medical center served as the test set for this investigation. The study protocol received Institutional Review Board approval and the study was performed in compliance with the Health Insurance Portability and Accountability Act of 1996. All of the reports were created by voice dictation, and were transcribed either manually or using a speech recognition system. The report text represented final, approved report content, and consisted of the procedure name, narrative ("findings") section, and report impression. The reports and the reporting templates were created independently. The reports were created about 2 years before the templates were developed; the reporting templates were developed by national committees without access to a specific set of radiology reports.

For each reporting template, we identified a single corresponding radiology procedure name from the institution's charge master. For example, the "Chest Xray" template was matched with the "DX CHEST XRAY PA-LAT" procedure code. Other chest radiographic procedures, such as single-view chest examinations, were not included in this analysis. The reporting templates and corresponding radiological procedure codes are shown in Table 1.

We applied the NCBO BioPortal Annotator [14] (http://bioportal.bioontology.org/annotator) to identify matching concepts from the RadLex and SNOMED CT vocabularies with terms in our sample of reporting templates and free-text reports. The Annotator processes text submitted by users through a RESTful web interface, uses string-matching to recognize terms from specified biomedical ontologies within the given text, and returns the annotations to the user [15].

Table 1 The five reporting templates and corresponding imaging procedure names, selected from the chargemaster of the participating institution

Report description	Template name	Procedure name
Two-view chest radiography	Chest Xray	DX CHEST PA-LAT
Non-contrast head CT	CT Brain	CT HEAD W/O CONT
Non-contrast spine MRI	MR Spine	MR L SPINE W/O CONT
Radionuclide bone scintigraphy	NM Bone Scan	NM BONE WHOLE BODY
Complete abdomen ultrasound examination	US Abdomen	US ABD COMPLETE

First, we applied BioPortal Annotator to identify RadLex and SNOMED CT concepts from each reporting template. Then, for each clinical radiology report, BioPortal Annotator was used to annotate the clinical reports with matching RadLex and SNOMED CT concepts. Report annotation was automated completely through the NCBO Annotator's web service, and results were stored in a database. In our analysis, the number of "Unique Concepts" represents the number of distinct concepts from these two ontologies that appear in at least one of the clinical reports for a specific radiology procedure. We defined "Concept Occurrences" as the sum of the number of reports in which each of the unique concepts occurs. For all of the reports of a specific procedure, we tallied the number of unique concepts identified by the annotation process and the total number of clinical reports in which each concept occurred. We compared the concepts that appeared in the report templates ("template-based concepts") with the concepts that appeared in free-text reports.

Results

The five reporting templates are shown in Table 2 with the number of elements and concepts for each template. The number of reporting elements indicates how many predefined terms such as section headings (e.g., "Findings"), anatomic sites (e.g., "Left kidney"), observation descriptors ("Hydronephrosis"), and predefined values (e.g., "Severe") appear in the reporting template. Each element may have mapped to zero, one, or more than one term in a vocabulary; the total number of annotations is shown in the rightmost column. For example, the 25 elements of the chest radiograph reporting template were mapped to 41 concepts ("Appendices" section).

The annotation results of the full-text reports are shown in Table 3. The 860 chest radiograph exam ("DX CHEST PA-LAT") reports, for example, contained 2,360 unique concepts, of which 33 (1.4 %) matched the 41 concepts generated from the corresponding reporting template ("Chest Xray"). As expected, this result indicates the reporting template contains far fewer terms than those found in actual radiology reports. Of the 53,624 concept occurrences for this procedure's reports, however, 9,931 (17.2 %) were related to concepts that appeared in the reporting template.

As shown in Table 3, the template-based concepts appeared significantly more frequently. The 33 concepts in the "Chest Xray" template appeared 14.7 times more frequently in actual reports than concepts that did not appear in the reporting template. For all five of the procedures studied here, the template-based concepts appeared in actual reports at least 2.5 times more frequently than non-template-based concepts. The chi-squared test for each report type showed a significant difference at a threshold of $p<0.00001$.

Discussion

The RSNA reporting templates have been created to represent "best practice" in radiology reporting [8], rather than as a normative standard. In general, the templates were crafted

Table 2 The number of reporting elements and associated RadLex® and SNOMED CT® concepts for the five selected reporting templates

Template name	Template ID	No. of reporting elements	No. of concepts
Chest Xray	0000102	25	41
CT Brain	0000004	61	152
MR Spine	0000071	200	254
NM Bone Scan	0000079	53	130
US Abdomen	0000087	97	222

Table 3 For each procedure, the table indicates the number of reports analyzed and their number of unique concepts

Procedure name	No. of reports	All concepts			Template-based concepts				
		No. of concepts	No. of occurrences	Mean occurrences per concept	No. of concepts	No. of occurrences	Mean occurrences per concept	Coverage, %	Relative frequency
DX CHEST PA-LAT	860	2,360	53,624	22.7	33	9,231	279.7	17.2	14.7
CT HEAD W/O CONT	323	2,041	39,314	19.3	127	9,971	78.5	25.4	5.1
MR L SPINE W/O CONT	35	766	4,407	5.8	155	1,808	11.7	41.0	2.7
NM BONE WHOLE BODY	26	505	2,571	5.1	50	586	11.7	22.8	2.7
US ABD COMPLETE	57	757	7,552	10.0	146	3,708	25.4	49.1	4.0

The number of concept occurrences is the sum of number of reports in which each concept appears

"Coverage" indicates the percentage of concept occurrences related to template-based concepts. The "Relative Frequency" is the ratio of mean occurrences per concept for template-based and non-template-based concepts

by national committees of subspecialty experts or as "time-tested" examples of reporting templates used at individual institutions. That the reporting templates adequately capture the salient aspects of corresponding radiology reports is an untested hypothesis. The experiment described here sought to evaluate the extent to which the RSNA reporting templates covered the content of corresponding free-text reports.

The RSNA reporting templates that we examined accounted for no fewer than 17 % and up to 49 % of the concept occurrences in a sample of corresponding radiology reports. Although the reporting templates contained a small number of unique concepts, their concepts appeared with high frequency in radiology reports. For all reports in this study, template-based concepts appeared in actual reports at least 2.5 times more frequently than non-template-based concepts.

This study had several limitations. We examined a small number of reporting templates, and explored reports of only one procedure type for each template. The reports were obtained over a relatively brief period (1 week) from a single institution, and hence may reflect individual biases. The NCBO Annotator often identified multiple concepts for a specific term. For example, the phrase "right kidney" was mapped to annotations for "right," "kidney,", "right kidney," "entire kidney," "kidney structure," and "right kidney structure." Such redundancy may artificially increase the percentage of matching terms.

Despite these limitations, we believe that our results provide useful estimates of how well the reporting templates capture the concepts that appear frequently in radiology reports. A more "complete" template may be desirable, but it is likely to be more complex and possibly more difficult to use. Even a relatively simple template, such as "Chest Xray," addressed almost one-fifth of the concepts that appeared in two-view chest radiograph reports. The techniques used here may be helpful to determine the appropriate complexity of radiology reporting templates, and to identify those concepts that appear most frequently and should be considered for inclusion in the templates. Such techniques may be incorporated into automated approaches to construct reporting templates that optimally model the content of clinical radiology reports.

Conclusion

The reporting templates analyzed in this study yielded 17 to 49 % of the concept occurrences in actual radiology reports, and contained concepts that appeared significantly more frequently than others. This finding suggests that the RSNA reporting templates provide useful coverage of the "domain of discussion" in radiology reports. The techniques used in this study can guide the development of reporting templates by identifying concepts that occur frequently in radiology reports. These techniques also can help evaluate the coverage of existing templates and establish global benchmarks for reporting templates.

Acknowledgments This research was supported in part by the National Institute of Biomedical Imaging and Bioengineering (NIBIB). We thank the RSNA Radiology Informatics Committee for leading and supporting the radiology reporting initiative, and we acknowledge the many RSNA volunteers who helped develop the reporting templates.

Appendix A

Table 4 Reporting elements from the "Chest Xray" template, shown in order of appearance. Indentation is added to show the elements' hierarchy

Report
Procedure
View
PA
AP
Lateral
Clinical information
Cough
Fever
Shortness of breath
Preoperative exam
Comparison
None
Findings
Heart
Normal
Lungs
Normal
No acute disease
Bones
Normal
Degenerative changes
Impression
Normal
No acute disease

Appendix B

Table 5 The 41 concepts derived from the "Chest Xray" template's reporting elements

Concept name	Ontology	Concept ID
Acute	SNOMED CT	53737009
acute	RadLex	RID5718
Acute disease	SNOMED CT	2704003
anteroposterior view	RadLex	RID28784
Breath	SNOMED CT	11891009
Clinical	SNOMED CT	58147004
clinical information	RadLex	RID13166
comparison	RadLex	RID28483
Cough	SNOMED CT	49727002
cough	RadLex	RID39051
Disease	SNOMED CT	64572001
Dyspnea	SNOMED CT	267036007
Dyspnea	SNOMED CT	49233005
Entire bony skeleton	SNOMED CT	128530007
Entire heart	SNOMED CT	302509004
Fever	SNOMED CT	386661006
fever	RadLex	RID39083
heart	RadLex	RID1385
Heart structure	SNOMED CT	80891009
impression section	RadLex	RID13170
Increased body temperature	SNOMED CT	64882008
Lateral	SNOMED CT	49370004
lateral	RadLex	RID39121
lateral view	RadLex	RID5821
lungs	RadLex	RID13437
none	RadLex	RID28454
Normal	SNOMED CT	17621005
normal	RadLex	RID13173
observations section	RadLex	RID28486
Preoperative	RadLex	RID28815
posteroanterior view	RadLex	RID28625
Procedure	SNOMED CT	71388002
procedure	RadLex	RID1559
Pyrexia	SNOMED CT	248425001
Report	SNOMED CT	229059009
report	RadLex	RID28487
Report procedure	SNOMED CT	308561006
set of bones	RadLex	RID28569
shortness of breath	RadLex	RID39265
View	SNOMED CT	246516004
view	RadLex	RID12243

Appendix C

Sample narrative (free-text) chest radiography report

Narrative Chest. Comparison: 03/06/07. AP upright and left lateral upright views of the chest reveal a transverse cardiac diameter that is within normal limits. There is mild tortuosity and ectasia of the thoracic aorta which is unchanged. Mediastinal width and pulmonary vasculature is normal. The lung fields are free of infiltrate, consolidation, or effusion. There is evidence of hyperinflation with increased AP chest dimension. One questions if patient has an element of obstructive pulmonary disease. Again noted is a sending device overlying the left midlung field. An electrode lead extends cephalad into the cervical area on the left. This is essentially unchanged from the previous films.

Impression (1) Aortic tortuosity and ectasia with no acute cardiopulmonary disease. (2) Lung field changes suggestive of obstructive pulmonary disease.

Appendix D

Table 6 Concepts identified by NCBO Annotator for the example report in Appendix C, listed alphabetically by concept name. Of the 72 concepts identified in this report, 10 appear in the corresponding report template

Concept name	Ontology	Concept ID	Appears in template
Abnormally hard consistency	SNOMED CT	19730000	
Acute	SNOMED CT	53737009	X
acute	RadLex	RID5718	X
anteroposterior view	RadLex	RID28784	X
aorta	RadLex	RID480	
Aortic	SNOMED CT	261051005	
Aortic structure	SNOMED CT	15825003	
Area	SNOMED CT	42798000	
Cephalic	SNOMED CT	66787007	
Cervical	SNOMED CT	261064006	
Chemical element	SNOMED CT	57795002	
comparison	RadLex	RID28483	X
Consolidation	SNOMED CT	9656002	
Device	SNOMED CT	49062001	
Diameter	SNOMED CT	81827009	
diameter	RadLex	RID13432	
Dilatation	SNOMED CT	25322007	
dilation	RadLex	RID4743	
Disease	SNOMED CT	64572001	X
Disorder of lung	SNOMED CT	19829001	
Effusion	SNOMED CT	41699000	
Effusion	SNOMED CT	430869004	
effusion	RadLex	RID4872	
Electrode	SNOMED CT	16470007	
electrode	RadLex	RID5456	
Entire aorta	SNOMED CT	181298001	
Entire lung	SNOMED CT	181216001	
Entire thoracic aorta	SNOMED CT	302510009	
Evidence of	SNOMED CT	18669006	
Free of	SNOMED CT	37837009	
Hyperdistention	SNOMED CT	73578008	
impression section	RadLex	RID13170	X
Increased	SNOMED CT	35105006	
increased	RadLex	RID36043	
Infiltration	SNOMED CT	47351003	
Is a	SNOMED CT	116680003	
Lateral	SNOMED CT	49370004	X
Lead	SNOMED CT	88488004	
lead	RadLex	RID11924	
Left	SNOMED CT	7771000	
left	RadLex	RID5824	
lung	RadLex	RID1301	
Lung field	SNOMED CT	34922002	
Lung structure	SNOMED CT	39607008	
Mediastinal	SNOMED CT	264099006	
Mild	SNOMED CT	18647004	
Mild	SNOMED CT	255604002	
mild	RadLex	RID5671	
Morphology within normal limits	SNOMED CT	125112009	
No status change	SNOMED CT	260388006	
Normal	SNOMED CT	17621005	X
normal	RadLex	RID13173	X
Normal limits	SNOMED CT	260394003	
observations section	RadLex	RID28486	X
Over	SNOMED CT	21481007	
Overlying behavior	SNOMED CT	32102004	
Patient	SNOMED CT	116154003	
Previous	SNOMED CT	9130008	
previous	RadLex	RID5726	
Pulmonary	SNOMED CT	264164005	
Suggestive of	SNOMED CT	7196007	
Thoracic	SNOMED CT	261179002	
thoracic aorta	RadLex	RID879	
Thoracic aorta structure	SNOMED CT	113262008	

Table 6 (continued)

Concept name	Ontology	Concept ID	Appears in template
Thoracic structure	SNOMED CT	51185008	
thorax	RadLex	RID1243	
Tortuosity	SNOMED CT	15690004	
Transverse	SNOMED CT	62824007	
transverse	RadLex	RID5854	
upright position	RadLex	RID10455	
vasculature	RadLex	RID15989	
Width	SNOMED CT	103355008	

References

1. Weiss DL, Langlotz CP: Structured reporting: patient care enhancement or productivity nightmare? Radiology 249:739–747, 2008
2. Langlotz CP: Structured radiology reporting: are we there yet? Radiology 253:23–25, 2009
3. Bosmans JM, Weyler JJ, De Schepper AM, Parizel PM: The radiology report as seen by radiologists and referring clinicians: results of the COVER and ROVER surveys. Radiology 259:184–195, 2011
4. Schwartz LH, Panicek DM, Berk AR, Li Y, Hricak H: Improving communication of diagnostic radiology findings through structured reporting. Radiology 260:174–181, 2011
5. Sistrom CL, Honeyman-Buck J: Free text versus structured format: information transfer efficiency of radiology reports. AJR Am J Roentgenol 185:804–812, 2005
6. Grieve FM, Plumb AA, Khan SH: Radiology reporting: a general practitioner's perspective. Br J Radiol 83(985):17–22, 2010
7. Douglas PS, et al: ACCF/ACR/AHA/ASE/ASNC/HRS/NASCI/RSNA/SAIP/SCAI/SCCT/SCMR 2008 Health policy statement on structured reporting in cardiovascular imaging. Journal of the American College of Cardiology 53:76–90, 2009
8. Kahn, CE Jr, et al: Toward best practices in radiology reporting. Radiology 252:852–856, 2009
9. Hong Y, Zhang J, Heilbrun ME, Kahn, CE Jr: Analysis of RadLex coverage and term co-occurrence in radiology reporting templates. J Digit Imaging 25:56–62, 2012
10. Langlotz CP: RadLex: a new method for indexing online educational materials. RadioGraphics 26:1595–1597, 2006
11. Rubin DL: Creating and curating a terminology for radiology: ontology modeling and analysis. J Digit Imaging 21:355–362, 2008
12. Elkin PL, et al: Evaluation of the content coverage of SNOMED CT: ability of SNOMED clinical terms to represent clinical problem lists. Mayo Clin Proc 81:741–748, 2006
13. SNOMED Clinical Terms. 2013. http://www.ihtsdo.org/snomed-ct/. Accessed 16 Jan 2013.
14. Jonquet C, Shah NH, Musen MA: The open biomedical annotator. Summit on Translational Bioinformatics 2009:56–60, 2009
15. Whetzel PL, et al: BioPortal: enhanced functionality via new Web services from the National Center for Biomedical Ontology to access and use ontologies in software applications. Nucleic Acids Res 39:W541–545, 2011