CrossMark

ORIGINAL ARTICLE

Structured report compliance: effect on audio dictation time, report length, and total radiologist study time

Tarek N. Hanna[1] · Haris Shekhani[1] · Kiran Maddu[1] · Chao Zhang[2] · Zhengjia Chen[2] · Jamlik-Omari Johnson[1]

Abstract The purpose of this study was to examine structured template use among emergency radiologists, and if this influences audio dictation time, radiology report length, or total radiologist study time. Retrospective data collection of consecutive occurrences of seven common imaging examinations interpreted by a dedicated emergency radiology division over a 2-month period yielded 3449 reports. Templates had been in place for >3 years. For each examination, we documented the individual audio dictation time (ADT), total words, and total time the radiologist spent on a study from report creation until final signing. In 81.2 % ($n = 2772$) of all cases, a basic template was used. In 2.8 % ($n = 78$) of these template-use cases, the radiologist removed key elements from the structured template. Of the 3417 reports with complete data, mean ADT was 37.3 s, mean word length was 132.3 (of which, on average, 64 were dictated), and total radiologist time per study (TRT) was 349.7 s. Study type was significantly associated with ADT, total words, and TRT ($p < 0.001$). Template usage decreased ADT ($p < 0.001$) by 47 %, but did not affect total word length or TRT. Parameters varied by individual attending ($p < 0.001$): 20 % (2/10) of attendings had differences in report length when using versus not using templates ($p < 0.001$). With long-term template usage, compliance with structured templates is high, and few radiologists significantly alter the templates. Template use decreases ADT and for a small fraction of radiologists impacts total word length and has a mixed impact on TRT.

Keywords Structured report · Radiology · Informatics · Speech recognition · Template · Turnaround time

Introduction

Communication of diagnostic radiology results is heavily, although not exclusively, reliant on the written radiology report [1]. As such, it is not surprising that scholarship regarding radiology reports has existed for decades [2–5]. Following the widespread implementation of speech recognition beginning in the late 1990s, structured reporting templates gained supporters, as the clarity and comprehensiveness implied in their itemized format was preferred among both radiologists and referring physicians [6–9]. Speech recognition has myriad benefits, however, 4.8–22 % of reports created via speech recognition may have errors [10–13] and 1.9 % of these errors may alter report interpretation [13]. When capturing any error, no matter how minor, up to 60 % of reports may contain speech recognition errors [14]. Although there are few existing studies that correlate use of structured reports to quality [6, 14], usage of pre-created templates should decrease speech recognition time and thus may reduce error rate. The written radiology report also serves as a billing document, where the comprehensive and standardized nature of structured reports can add value.

In the emergency department, as well as in other settings, clinicians prefer itemized structured reports [8, 15, 16], presumably because of their readability and clarity. For these reasons, our academic emergency radiology practice

✉ Tarek N. Hanna
    tarek.hanna@emory.edu

[1] Division of Emergency Radiology, Department of Radiology and Imaging Sciences, Emory University School of Medicine, Atlanta, GA 30322, USA

[2] Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA 30322, USA

Springer

transitioned to itemized structured reporting approximately 3 years prior to this study, and all radiologists were encouraged to use pre-created reporting templates. Previous work in subspecialty divisions has shown that (with incentives) radiologist compliance with structured reports can approach 100 % [16]. This study was undertaken to examine template compliance in our emergency radiology division, the frequency in which an itemized radiology report was substantially altered, and the effect that structured reporting template use had on radiologist-specific parameters such as audio dictation time, report length, and total radiologist study time. We hypothesized that template driven reports would have shorter audio duration times leading to a decrease in speech recognition errors and decreased overall radiologist study time resulting in increased radiologist efficiency.

Materials and methods

This retrospective study was approved by our institutional review board. Data was acquired from our departmental Powerscribe 360 (Nuance Communications, Burlington, MA) database. All examinations were ordered from the Emergency Departments (ED) of three university-affiliated hospitals. Interpretations were provided by a dedicated 24-7-365 Emergency Radiology Division consisting of ten fellowship-trained attending radiologists during the study period.

Data collection

For this analysis, seven common imaging studies performed in the Emergency Department were selected: computed tomography (CT) of the chest with intravenous (IV) contrast, CT head without IV contrast, CT abdomen and pelvis with IV contrast, CT abdomen and pelvis without IV contrast, one view chest radiographs, two view (posteroanterior and lateral) chest radiographs, and right upper quadrant ultrasound. Consecutive ED occurrences of these examinations during a 2-month period (July 1, 2014 through August 31, 2014) identified 3449 diagnostic radiology reports. Corresponding patient and examination data were extracted into a database.

Using the Powerscribe 360 database for each examination, we documented the individual audio dictation time, dictated words, total words, and total time the radiologist spent on a study from the time the report was created until when it was completed and signed. We excluded examinations with resident involvement or when greater than one study was linked and dictated in a single radiology report.

Template analysis

For every radiology report, we recorded whether a basic template was used by comparing the final radiology report to base templates in our system. If a basic template was used, we assessed if there were missing elements from that template or if the template was complete.

We used the following guidelines for the template qualifications:

1. Every standard template in our system has capitalized *findings* and *impression*. If a report did not have *findings* and *impression* sections capitalized, then no basic template was used.
2. Our template reports all contain itemized subheadings under the *findings* section. If some of these subheadings were missing the report was considered to be a *basic template with missing elements*.
3. Even if criteria 1 was fulfilled, if the *findings* section in the template was entirely erased and compiled freelance then it was considered that a basic template was not used.

Statistical methods

For numeric covariates, the mean and standard deviation of the outcomes are calculated and presented. For categorical variables, frequency and percentage are calculated and presented. Fisher's exact test is employed to test if there is any association between the providers and questions. For univariate analysis of study type and template usage, ANOVA, Chi-square test, and Kruskal-Wallis test were employed based on the characteristics of the data set, as detailed in the result section. The significance level is set at 0.05. SAS 9.4 is used for data analyses and management.

Results

Among the 3449 cases, 32 cases were excluded from analysis due to missing data. In 81.2 % ($n = 2772$) of all cases, a basic template was used. In 2.8 % ($n = 78$) of these cases with template usage, the radiologist removed key elements from the structured template. Descriptive statistics of the 3417 reports with complete data are presented in Table 1.

Table 2 presents the reports by study type. For each study type, reports are displayed by template usage, as well as audio duration, total words, and total radiologist time. For all covariates, there was a significant association with study type

Table 1 Descriptive variables of 3417 radiology reports

	Mean ± SD
Audio duration (s)	37.3 ± 45.5
Total words	132.3 ± 97.4
Dictated words	64.1 ± 86.7
Total radiologist time (s)	349.7 ± 1263

Table 2 Analysis of template use by study type

Covariate	Statistics	CT chest with contrast (N = 179)	CT head without IV contrast (N = 810)	CT abdomen + pelvis with IV contrast (N = 343)	CT abdomen + pelvis without IV contrast (N = 237)	XR chest 1 view portable (N = 708)	XR 2 view pa lateral (N = 1029)	US RUQ (N = 111)	p value
No template	N (Col %)	14 (7.8)	100 (12.4)	94 (27.4)	62 (26.22)	164 (23.2)	197 (19.2)	13 (11.7)	<0.001*
Template with missing elements	N (Col %)	20 (11.2)	7 (0.9)	30 (8.8)	19 (8.0)	1 (0.1)	1 (0.1)	0 (0)	
Template with no missing elements	N (Col %)	145 (81.0)	702 (86.8)	219 (63.9)	156 (65.8)	543 (76.7)	830 (80.7)	98 (88.3)	
Audio duration	N	179	810	343	237	708	1029	111	<0.001**
	Mean (±SD)	80.26 (±61.3)	27.58 (±30.6)	98.21 (±62.9)	89.96 (±56.2)	21.01 (±16.9)	16.37 (±16.6)	35.37 (±28.5)	
Total words	N	179	810	343	237	708	1029	111	<0.001**
	Mean	233.76 (±103.6)	129.89 (±67.1)	283.88 (±96.3)	251.63 (±89.7)	77.72 (±33.9)	73.27 (±28.0)	159.16 (±108.2)	
Total radiologist (s)	N	179	810	343	237	708	1029	111	<0.001*
	Mean	753.5 (±646.5)	412.7 (±2313.1)	826.65 (±809.2)	723.02 (±1157.3)	149.5 (±525.5)	120.7 (±211.3)	366.4 (±606.1)	

($p < 0.001$). That these parameters (audio duration, word usage, and radiologist time) should be associated with study type is not surprising—as more complex cross-sectional imaging exams take longer to dictate and use more total words than radiographs. That template usage is significantly associated with study type implies that in certain cases (such as the right upper quadrant ultrasound with 88.3 % template usage with no missing elements) radiologists find it easier to deploy templates.

Table 2 Analysis of template use by study type: univariate association with study type and whether or not a template was used. *The parametric p value is calculated by ANOVA for numerical covariates. For variable template and missing, using, p value is calculated by Chi-square test for categorical covariate. **The non-parametric p value is calculated by the Kruskal-Wallis test because of non-homogeneity of variance.

Subsequently, the audio duration, total words, and total radiologist time were analyzed by individual attending to see if these reporting properties varied by attending. These covariates all were significantly associated with the individual attending ($p < 0.001$). For total words, the mean among attendings varied from 86.2 for the most succinct attending to 170.5 for the most verbose, while mean audio duration ranged from 18 to 71.6 s. Interestingly, the mean longest reports did not correspond to the audio time, indicating that some radiologists either typed a significant among of additional material, used pre-created macros which populated sentences, or dictated at a much faster word-per-second rate (getting more words into a shorter audio duration period).

Returning to the overall data set, Table 3 examines the impact of template usage across all reports. Template use results in a significant decrease in audio duration time ($p < 0.001$), but does not impact total number of words or total radiologist time spent on exams.

Table 3: Association of report parameters with template use: values are displayed as mean ± SD. *The parametric p value is calculated by ANOVA. **The non-parametric p value is calculated by Kruskal-Wallis test.

Given that template use did not have an impact on total words or total radiologist study time in all reports—the data was grouped by attending and reanalyzed. We sought to determine if the total word length and time per study for an individual attending was impacted if they used a template or not. In 20 % (2/10) of cases, attendings had significant differences in total word length when templates were used. Ten percent of cases ($n = 1/10$) were higher word count with template usage ($p < 0.001$), and 10 % of cases ($n = 1/10$) were lower word count with template usage ($p < 0.001$). These adjusted p values were calculated by the Satterthwaite test because of non-homogeneity of variance. For total study time, there was significance in only 10 % of cases ($n = 1/10$), with longer total study time for radiologists in the template usage group ($p < 0.05$).

Table 3 Association of report parameters with template use

Covariate	Template ($N = 2772$)	No template ($N = 644$)	p value
Audio duration (s)	31.97 (±39.9)	60.14 (±59.1)	<0.001**
Total words	129.16 (±92.5)	146.02 (±115.4)	0.925**
Total radiologist time (s)	347.96 (±1372.9)	357.15 (±593.8)	0.868*

A regression analysis was performed to see what variables predicted longer audio duration time or longer total radiologist time. The type of study and basic template usage were significant predictors of audio duration ($p < 0.001$). Total radiologist time per study was not significantly associated with any variables.

Discussion

More than 3 years after implementation, template compliance in our practice is strong, with 81 % use of structured reporting templates. However, our compliance does not approach the 100 % reported by Larson et al., perhaps due to the constant feedback and financial incentives in that organization [16]. This comparison in itself may be a defense for radiologist incentivization if 100 % template use is desired. When radiologists deploy templates, in only 2.8 % of cases do they delete major template elements. Thus, the readability, completeness, clarity, and billing utility which are the hallmark of structured itemized reports [8] are unaffected in the majority of cases. Template use significantly decreased audio duration time, with mean dictation time of 32 s in cases of template use, and 60 s without template use, representing a 47 % reduction. In the era of speech recognition, the radiologist replaced the transcriptionist as report editor, resulting in increased errors of syntax, grammar, and semantics [17, 18]. Given that speech recognition results in errors in up to 9.7 % of reports [13], we hypothesize that decreasing audio duration would proportionally decrease speech recognition errors. Although a logical extension of our data, we did not specifically examine for decreased errors with template use.

Total radiologist time spent on examinations was not significantly associated with template use. To be clear, the use of speech recognition software (SRS) improves report turnaround time [19, 20]; we show here that within the context of SRS, template deployment does not further improve radiologist time per report. It stands to reason that in the same overall time per study, the non-template using radiologist is performing different tasks given the marked difference in audio duration time among these two groups. Specifically, the template users dictate for 32 s out of 348 s, which means they are dictating for 9.2 % of the time they have a study open. For non-template users, this is higher at 16.8 %. How does this difference in action impact study interpretation and error rate? Our study does not answer this, but raises interesting questions about the impact of multi-tasking on the diagnostic radiologist. The authors note that all users in our system were familiar with our structured templates, which had

been around for just over 3 years, and these numbers likely represent a plateau for our group. When individual radiologists were examined, one radiologist (10 % of the total pool) showed a statistically significant increase in total study time in cases of template use. This amounted to a mean increase of 60 s when a template was used; examining more granular data did not readily explain why this was the case for this individual, and this remains unclear to the authors.

Total report word length was not significantly associated with template use. Yet, when we looked at radiologists on an individual level to see how templating affected their dictation length, 10 % (1/10) of radiologists had significantly longer dictations with templates ($p < 0.001$), 10 % (1/10) had significantly shorter dictations ($p < 0.001$), and the remaining 80 % of radiologists showed no affect. However, we note that in our study the internal contents of the unstructured, non-template reports were not examined for completeness. Additionally, there may be some degree of selection bias on the part of the radiologist: the decision to deploy or not deploy a template may be based on study complexity. We know the structured itemized reports are comprehensive; it could be that the unstructured reports, while the same overall length, do not contain all necessary elements in all cases. Notably, the ultimate in standardized language is point-and-click structured reporting, which has not been shown to improve report accuracy or completeness [21, 22].

There is significant individual variability among attendings in all parameters [23], with audio duration, total report length, and total radiologist study time (all with $p < 0.001$). Many radiology groups, including ours, report turnaround time metrics to individual radiologists, in a quality effort to track and improve efficient and timely report availability. Our analysis shows cumulative radiologist time per study to be an overall basket of many reporting tasks. As a specific example: individual radiologist time per study ranged from 86 to 701 s on average. The radiologist with the longest time per study also had the most total average words at 170, versus a low of 86 among all radiologists. However, a separate radiologist had the longest audio duration time. The radiologist with the shortest time per study was not the radiologist with the shortest reports or the shortest audio duration time. This type of in-depth, individualized reporting analysis could be helpful to coach both radiologists and trainees to improve speed through efficiency in dictation and report creation by identifying specific factors in which they are taking more time then peers.

Limitations

Our study has limitations. First, our data arise from a single academic institution among radiologists who were familiar with our templates and applied to an Emergency Department (ED) examination set. Thus, generalizability must be done with care. Specifically, the familiarity of our attendings with our templates likely made them easier to use and increased compliance. Although we had a large number of reports, there were only ten radiologists, and individual factors among these radiologists could conceivably affect the data. Finally, our attending radiologists range in age and training background, with some attendings who used templates in training, while others had significant practice or training exposure in transcription or freelance speech recognition use. These affects are difficult to ascertain.

Conclusions

Standard itemized template usage is accepted by radiologists with greater than 80 % compliance, even in an organization without incentives for template use. Template use does not affect radiologist time per study and does not affect report length. However, template use significantly decreases audio dictation time, which is the major cause of errors in the era of speech recognition software.

References

1. Kushner DC, Lucey LL, American College of R (2005) Diagnostic radiology reporting and communication: the ACR guideline. J Am Coll Radiol 2(1):15–21. doi:10.1016/j.jacr.2004.08.005
2. Friedman PJ (1983) Radiologic reporting: structure. AJR Am J Roentgenol 140(1):171–172. doi:10.2214/ajr.140.1.171
3. Friedman PJ (1983) Radiologic reporting: the hierarchy of terms. AJR Am J Roentgenol 140(2):402–403. doi:10.2214/ajr.140.2.402
4. Martin LF (1982) Opinion: is this your report? J Can Assoc Radiol 33(4):255–256
5. Orrison WW, Nord TE, Kinard RE, Juhl JH (1985) The language of certainty: proper terminology for the ending of the radiologic report. AJR Am J Roentgenol 145(5):1093–1095. doi:10.2214/ajr.145.5.1093
6. Pool F, Goergen S (2010) Quality of the written radiology report: a review of the literature. J Am Coll Radiol 7(8):634–643. doi:10.1016/j.jacr.2010.03.016
7. Schwartz LH, Panicek DM, Berk AR, Li Y, Hricak H (2011) Improving communication of diagnostic radiology findings through structured reporting. Radiology 260(1):174–181. doi:10.1148/radiol.11101913
8. Naik SS, Hanbidge A, Wilson SR (2001) Radiology reports: examining radiologist and clinician preferences regarding style and content. AJR Am J Roentgenol 176(3):591–598. doi:10.2214/ajr.176.3.1760591
9. Sistrom CL, Langlotz CP (2005) A framework for improving radiology reporting. J Am Coll Radiol 2(2):159–167. doi:10.1016/j.jacr.2004.06.015
10. Kanal KM, Hangiandreou NJ, Sykes AM, Eklund HE, Araoz PA, Leon JA, Erickson BJ (2000) Evaluation of the accuracy of a continuous speech recognition software system in radiology. J Digit Imaging 13(2 Suppl 1):211–212
11. McGurk S, Brauer K, Macfarlane TV, Duncan KA (2008) The effect of voice recognition software on comparative error rates in radiology reports. Br J Radiol 81(970):767–770. doi:10.1259/bjr/20698753
12. Quint LE, Quint DJ, Myles JD (2008) Frequency and spectrum of errors in final radiology reports generated with automatic speech recognition technology. J Am Coll Radiol 5(12):1196–1199. doi:10.1016/j.jacr.2008.07.005
13. Ringler MD, Goss BC, Bartholmai BJ (2015) Syntactic and semantic errors in radiology reports associated with speech recognition software. Health Inform J. doi:10.1177/1460458215613614
14. Hawkins CM, Hall S, Hardin J, Salisbury S, Towbin AJ (2012) Prepopulated radiology report templates: a prospective analysis of error rate and turnaround time. J Digit Imaging 25(4):504–511. doi:10.1007/s10278-012-9455-9
15. Manoonchai N, Kaewlai R, Wibulpolprasert A, Boonpramarn U, Tohmee A, Phongkitkarun S (2015) Satisfaction of imaging report rendered in emergency setting: a survey of radiology and referring physicians. Acad Radiol 22(6):760–770. doi:10.1016/j.acra.2015.01.006
16. Larson DB, Towbin AJ, Pryor RM, Donnelly LF (2013) Improving consistency in radiology reporting through the use of department-wide standardized structured reporting. Radiology 267(1):240–250. doi:10.1148/radiol.12121502
17. Langer S (2002) Radiology speech recognition: workflow, integration, and productivity issues. Curr Probl Diagn Radiol 31(3):95–104
18. Pezzullo JA, Tung GA, Rogg JM, Davis LM, Brody JM, Mayo-Smith WW (2008) Voice recognition dictation: radiologist as transcriptionist. J Digit Imaging 21(4):384–389. doi:10.1007/s10278-007-9039-2
19. Prevedello LM, Ledbetter S, Farkas C, Khorasani R (2014) Implementation of speech recognition in a community-based radiology practice: effect on report turnaround times. J Am Coll Radiol 11(4):402–406. doi:10.1016/j.jacr.2013.07.008
20. Rana DS, Hurst G, Shepstone L, Pilling J, Cockburn J, Crawford M (2005) Voice recognition for radiology reporting: is it good enough? Clin Radiol 60(11):1205–1212. doi:10.1016/j.crad.2005.07.002
21. Johnson AJ, Chen MY, Swan JS, Applegate KE, Littenberg B (2009) Cohort study of structured reporting compared with conventional dictation. Radiology 253(1):74–80. doi:10.1148/radiol.2531090138
22. Lin E, Powell DK, Kagetsu NJ (2014) Efficacy of a checklist-style structured radiology reporting template in reducing resident misses on cervical spine computed tomography examinations. J Digit Imaging 27(5):588–593. doi:10.1007/s10278-014-9703-2
23. Rosenkrantz AB, Bonavita JA, Foran MP, Matza BW, McMenamy JM (2014) Is there an association between radiologist turnaround time of emergency department abdominal CT studies and radiologic report quality? Emerg Radiol 21(1):5–10. doi:10.1007/s10140-013-1164-2