Received:
05 September 2017

Revised:
09 November 2017

Accepted:
16 November 2017

<https://doi.org/10.1259/bjr.20170670>

Cite this article as:

Buckley BW, Daly L, Allen GN, Ridge CA. Recall of structured radiology reports is significantly superior to that of unstructured reports. *Br J Radiol* 2018; **90**: 20170670.

FULL PAPER

Recall of structured radiology reports is significantly superior to that of unstructured reports

¹BRYAN W BUCKLEY, MRCPI, ²LESLIE DALY, PhD, ²GRAINNE N ALLEN, BSc and ³CAROLE A RIDGE, MD, FFRCSI

¹St. Vincents University Hospital, Dublin, Ireland

²University College Dublin, Dublin, Ireland

³Royal Brompton and Harefield Trust, London, UK

Address correspondence to: Dr Bryan W Buckley
E-mail: bryan.buckley@ucdconnect.ie

Objective: To measure recall of structured compared with unstructured radiology reports.

Methods: Institutional review board approval was obtained. Four hypothetical radiology reports, two structured and two unstructured reports, were created for the purposes of this study by an experienced consultant radiologist. The reports, each followed immediately by a multiple-choice questionnaire listing possible diagnoses from the report, were distributed to the members of two national physician associations using a web-based survey tool. Based on the number of correct responses, correct critical findings and incorrect responses, rates per number of potential diagnoses were calculated for each individual and averaged. The paired sign test compared results between structured and unstructured reports.

Results: 148 respondents completed the survey, 126 (85.1%) of whom were physicians. The mean percentage of incorrect diagnoses was 4.5% for structured reports compared with 16.7% for unstructured reports ($p < 0.001$). The average rate of critical diagnosis recall was 82.7% for structured reports and 65.1% for unstructured reports ($p < 0.001$). The average percentage of all diagnoses detected for structured compared with unstructured reports was 64.3 and 59.0%, respectively ($p = 0.007$).

Conclusion: Recall of structured radiology reports is significantly superior to recall of unstructured reports immediately after reading the report.

Advances in knowledge: A structured radiology report format can positively impact the referring clinician's ability to recall the critical findings with statistical significance.

INTRODUCTION

The radiologist's report is the final product of any radiology department; it serves as the main means of communication between the radiologist and the referring clinician, and is a formal medicolegal document. Radiology reports can follow a "free-text" structure without headings and/or subheadings. An alternative to free-text reporting is the structured report, which may comprise a uniform and consistent report templates, headings, subheadings and occasionally, a standardized lexicon to create uniformity and potentially improve communication with referring clinicians. In practice, there is great variability among structured reports in terms of how rigidly they adopt a structured format.

The unstructured report allows the radiologist to use a personal reporting style, different in language and structure, which a radiologist may prefer as it is unique. The structured report may be construed as suppressing the reporter's autonomy. However, structured reporting has been demonstrated to confer a benefit in both surgical

and pathology reports, improving communication and consistency.¹⁻³

Much of the literature to date has been focussed on the effect report format has on radiologist performance, only one other study has examined the end users recall of structured compared with unstructured radiology reports.⁴ In this latter study, researchers recruited 16 senior medical students to review 12 radiology reports, half of which were structured and the other half of which followed a free-text format. The medical students then answered 10 multiple choice questions about the reports. Students were permitted to refer back to the report as needed without time constraints. Although no significant differences between test score, time to completion and efficiency were found when comparing structured and free-text reports, participants expressed a preference for the structured format. To our knowledge, no study has analysed a physician's recall of structured vs an unstructured format. In this study, therefore, we aim to measure recall of diagnoses after reading structured compared with unstructured radiology reports.

METHODS AND MATERIALS

The study received ethical board approval and respondent consent was obtained.

Radiology reports

A consultant radiologist with 10 years of experience in radiology devised four hypothetical radiology reports: two structured (MRI brain and CT thorax) and two unstructured (CT abdomen and CT angiogram). Structured reports followed template structured reports which are freely available on the Radiologic Society of North America radiology reporting website (www.radreport.org).⁵ Word counts for each report were similar for the structured MRI brain report (194 words), structured CT thorax report (135 words), unstructured CT abdomen report (152 words) and unstructured CT angiogram report (202 words). The radiologist selected typical radiologic findings in common clinical scenarios ensuring that each report had four positive findings. Balanced between structured and unstructured reports, there was either one critical and three non-critical findings or two critical and two non-critical findings. A critical finding was defined as a radiologic finding requiring diagnostic or therapeutic action. [The reports are provided in [Supplementary Material 1](#) (Supplementary material available online), which reproduces the entire survey instrument].

Survey

The survey was conducted using an online web-based service, SurveyMonkey.com. The four radiology reports were presented to each respondent in random order each followed immediately by a seven-response multiple-choice questionnaire of different diagnoses, three of which were not in the report (dummy diagnoses) and four of which were in the report.

Respondents were asked to select any number of the seven diagnoses but could not return to the report or to their responses once the question webpage had been left and received no feedback regarding the correctness of their responses. Respondents could pause the questionnaire to return at a later time and were not restricted in the time allowed to read the report, however, there was no facility to return to a completed section of the survey. Finally, respondents were asked to identify their practice setting, level of experience of clinical practice and structured radiology reports and the geographic location of their hospital or institution. [Supplementary Material 2](#) describes the correct critical and non-critical findings.

In April 2016, a link to the survey questions was included within a monthly newsletter and delivered to the 6405 electronic addresses registered to receive a newsletter from the Royal Colleges of Surgeons and Physicians in Ireland, respectively. The survey was electronically mailed a second time, 1 month later, to enhance the response rate. To avoid multiple survey completion by the same respondent, only one survey could be completed per electronic device.

Statistical analysis

All accuracy measures were calculated for each individual for unstructured and structured reports separately. The number of correct diagnostic choices (including critical findings) made by an individual is expressed as a percentage of the eight true diagnoses that could have been chosen (four per report). This is a measure of overall sensitivity to recall any of the diagnoses. The number of correct critical findings is expressed as a percentage of the three true critical findings (one or two per report). The number of choices of incorrect (dummy) diagnoses was expressed as a percentage of the six dummy diagnoses (three per report). Thus, for example, an individual who, in the two unstructured reports, ticked two of the critical diagnoses, three of the other true diagnoses and three dummy diagnoses would have a percentage of critical diagnoses detected of $2/3 = 66.7\%$, a percentage of all diagnoses detected of $5/8 = 62.5\%$ and a percentage of incorrect diagnoses detected of $3/6 = 50.0\%$. Each of the measures was then averaged over the number of relevant respondents involved. For the analysis of single reports, a similar approach was used.

The average proportion of correct critical diagnoses, correct positive findings and dummy diagnoses for all four reports is calculated using the number of the selected responses for all individuals taken as a percentage of all possible responses, illustrated in the "Results" section, the averaging of all responses by each individual respondent has the advantage that estimates are available the individual respondent and therefore, statistical significance tests are possible.

Although in many cases, the distributions were skewed, mean values are presented because the median was not sensitive to differences close to 100%. Comparisons between structured and unstructured reports were made using the exact paired sign test, and between groups using the independent Kruskal–Wallis test. All tests were as implemented in International Business Machines Corporation (New York, NY) Statistical Package for the Social Sciences (SPSS) Statistics v. 20. *p*-values of less than 0.05 were considered to indicate a statistically significant difference. 95% confidence intervals (CI) for a percentage were based on Wilson's method.

RESULTS

Respondents

207 recipients viewed the survey and 148 respondents completed it (71.4% completion rate) [95% CI (65.0 to 77.2)]. All respondents consented to participate. Respondents took an average of 21 min to complete the survey (range: 2 min–4.06 h). The 148 complete surveys were accepted for further analysis.

The slight majority of respondents were female ($n = 89$; 60.1%) and aged 25–34 years ($n = 88$; 59.5%). English was the primary spoken language of 98.0% ($n = 145$) of respondents. 19 respondents (12.8%) were medical students and 126 (85.1%) were physicians. Respondents selected Australia, Canada, India, Ireland, UK and USA as their country of current residence. [Table 1](#) provides the career stage of the physicians and [Table 2](#) provides the specialties of those who provided information.

Table 1. Career stage of respondents

Career stage	N	%
Intern	18	14.4
Senior house officer/junior resident	19	15.2
Registrar/senior resident	11	8.8
Specialist registrar/fellow	42	33.6
Consultant/attending	35	28
Total	125	100%

Of the 125 physicians answering this portion of the survey, 48 (38%) described their staff grade as either “intern”, “senior house officer” or “registrar”, 42 (33%) were “specialist registrar” or “fellow” 35 (28%) were consultants. In terms of exposure to radiology reports, 63 (50.4%) read radiology reports over 10 times per week. 49 (39.2%) stated that less than a quarter of the radiology reports they read in a given week followed a structured format.

Survey responses

Table 2 presents the mean percentages of incorrect responses, of critical diagnoses detected and of all diagnoses detected in the four individual reports and by structured vs unstructured reports. Supplementary Material 2 provides the number of times each answer was selected for individual reports. The correct answers are in *italics*. Respondents selected an average of 2.7 and

Table 2. Specialty of respondents

Specialty	Number of responses
Anaesthesia	3
Cardiothoracic surgery	2
Colorectal surgery	1
Emergency medicine	4
Ear, Nose and Throat	1
General internal medicine	26
General practice	8
General surgery	7
Geriatric medicine	3
Obstetrics and gynaecology	13
Ophthalmology	3
Other	28
Paediatric medicine or surgery	8
Plastic surgery	2
Psychiatry	5
Radiology	12
Trauma and orthopaedics	2
Urology	2
Total responses	130

Figure 1. Respondent selection for Report 1 of a structured MRI brain report. Correct diagnoses (striped bars) and dummy diagnoses (black bars) were selected by respondents. *denotes a critical finding.

2.9 of seven possible answers, for structured and unstructured reports, respectively ($p = 0.038$).

Structured MRI brain report

Figure 1 presents the results of the structured MRI brain report. 140 responses of 148 correctly identified the single critical finding of “acute left middle cerebral artery infarct” (94.6%). The mean percentage of all correct diagnoses detected was 66.0, 2% of responses were incorrect.

Structured CT Thorax report

Figure 2 presents the results of the structured CT Thorax. 76.7% of the possible responses correctly identified the critical findings of “suspicious right upper lobe nodule” and “left atrial appendage thrombus”. The percentage of all correct diagnoses detected was 62.5, 7.0% of responses were incorrect.

Unstructured CT abdomen report

Figure 3 shows the results for the unstructured CT abdomen report, 136 of 148 possible responses correctly identified the single critical finding “acute sigmoid diverticulitis” (91.9%). The mean percentage of all diagnoses detected was 60.5% and only 9.7% of responses were incorrect.

Unstructured cardiac CTA report

Figure 4 shows the results for the unstructured cardiac CTA report, 153 of 296 possible responses correctly identified the two critical findings of “right upper lobe spiculated nodule” and “left

Figure 2. Respondent selection for Report 1 of a structured CT thorax report. Correct diagnoses (striped bars) and dummy diagnoses (black bars) were selected by respondents. *denotes a critical finding.

ventricular apex infarct" (51.7%). Overall, the percentage of all diagnoses detected was 57.4, 23.7% of responses were incorrect. Of note, is the 55.4% who incorrectly selected "severe left main stem stenosis owing to calcified plaque" when in fact the report reported that the "left main stem is mildly stenosed owing to mixed plaque".

The accuracy of recall of the structured reports was superior to that of the unstructured reports (Table 3). The percentage of all diagnoses detected was 64.3 and 59.0% for structured and unstructured reports, respectively ($p < 0.007$). Again 82.7 and 65.1% of critical findings were correctly identified after reading the structured and unstructured reports, respectively ($p < 0.001$). The percentage of incorrect responses was 16.7% for unstructured reports compared with only 4.5% for the structured reports ($p < 0.001$).

Neither gender, specialty, geographic location, time taken to complete the survey, experience of structured reporting, age nor career stage showed a trend or correlation with recall of radiology report findings.

Interobserver agreement between the 128 respondents was calculated for each of the four reports using the κ statistic for multiple raters of a series of binary outcomes. Within each report, the seven possible diagnoses were rated based on whether they had been ticked or not.

Figure 3. Respondent selection for Report 1 of an unstructured CT abdomen report. Correct diagnoses (striped bars) and dummy diagnoses (black bars) were selected by respondents. *denotes a critical finding.

The κ -values for the two unstructured reports (Questions 2 and 3) were 0.36 and 0.28. The values for the two structured reports (Questions 1 and 4) were 0.53 and 0.38. Although these values are not high, there was more agreement within the structured reports than the non-structured ones

DISCUSSION

The decrease in face-to-face time between referring clinicians and radiologists and the inherent lost opportunity for radiologists to add value to clinical care is becoming an increasing challenge.^{6,7} Aside from multidisciplinary team meetings, the written word is increasingly the main method of radiology report communication during the working day. While attention in the published literature has been given to the effect of report format on radiologist performance, little has been reported regarding the effect of report format on physician recall of radiology reports.

The preference among referring clinicians may be for structured radiology reports^{8,9} with a reported 84.5% preference among clinicians for structured reports with templates and headings.¹⁰ It has also been demonstrated that satisfaction ratings and clarity ratings among both referring clinicians and radiologists is higher for structured reports when compared with free-text reports, particularly in the case of CT scans, suggesting that structured reporting lends itself well to complex imaging studies.¹¹

Much like the initial resistance to the widespread use of voice recognition software for improved cost and turnaround

Figure 4. Respondent selection for Report 1 of an unstructured cardiac CTA report. Correct diagnoses (striped bars) and dummy diagnoses (black bars) were selected by respondents. *denotes a critical finding.

efficiencies,^{12,13} there has been reluctance to abandoning the free-text form of reporting familiar to so many radiologists owing to concerns about the risk of excessive simplification, loss of autonomy and poor user compliance.¹⁴

In our study, respondents exhibited statistically significant inferior recall of all positive findings and critical findings, and made more incorrect diagnoses when reading unstructured or free-text reports compared with structured reports. The structured report format conferred a significant advantage in the transfer of information to the reader, when compared with the unstructured format.

The clinical implication of these findings is that referring clinicians are less likely to forget a critical finding and more likely able to recall all positive findings in a structured radiology report. The implications for healthcare institutions using structured reporting over unstructured may offer further benefits. With a low cost of implementation and potential for data mining for outcome and remuneration purposes, widespread structured reporting in radiology is also likely to improve operational efficiency.¹⁵

It is worth noting that in our survey, respondents performed worst after reading the unstructured cardiac CTA report. This echoes the findings of Ghoshhajra and colleagues who describe the impact that a final impression with bullet points has on report recall of cardiac CTA reports.¹⁶ Ghoshhajra and colleagues report κ -values for reader recall compared with the report's final impression which were 53% for an unstructured final impression compared with 68% for a final impression with bullet point format. These findings and the findings of our study suggest that the greater the complexity of the study, the greater the benefit there is from a structured report.

Our survey has some limitations. The greatest limitation is the fact that the reports used different modalities and areas of investigation between structured and unstructured reports. This may limit conclusions drawn from our comparisons as we were not comparing like with like. Comparing the same reports in both a structured and unstructured format would involve showing a respondent the same report in two different formats, separated in time. We expected this would lead to a respondent high drop out rate and weaken the study greater than having a comparison among a variety of modalities and imaged areas.

Our survey uses hypothetical radiology reports and, therefore, is a further limitation when extrapolating our findings to real world practice.

Our study was also limited by a small number of reports to compare against one another. As our survey was online, sent by email, our priority was to keep the survey short for respondents and achieve a high completion rate. Our 71.4% completion rate is good compared with rates for other published surveys.^{17,18}

Table 3. Average measures of recall accuracy in individuals for structured and unstructured radiology reports

	Mean % incorrect responses	Mean % critical diagnoses recalled	Mean % all diagnoses recalled
Structured MRI brain report	2.0%	94.6%	66.0%
Structured CT thorax report	7.0%	76.7%	62.8%
Unstructured abdominal CT report	9.7%	91.9%	60.5%
Unstructured cardiac CTA report	23.7%	51.7%	57.6%
All structured reports	4.5%	82.7%	64.3%
All unstructured reports	16.7%	65.1%	59.0%
Significance comparing structured with unstructured reports	$p < 0.001$	$p < 0.001$	$p = 0.007$

A final limitation is that data were obtained outside of a clinical setting, and thus the findings are not necessarily universally applicable and merit verification in clinical practice.

Future directions of study include¹ the time taken to read a radiology report relative to recall,² report recall in clinical practice and³ the benefit of a structured report in modalities other than CT and MRI.

CONCLUSION

Recall of structured radiology reports is significantly superior to recall of unstructured reports. Respondents exhibited better

recall of all findings, recalled less incorrect diagnoses and missed less critical findings after reading a structured radiology report. Structured radiology reports conferred a benefit when used in cross-sectional radiology reports.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the contributions of Dr Lucia Prihodova, Royal College of Physicians in Ireland and Professor Sean Tierney, Royal College of Surgeons in Ireland for assistance in study design and survey recruitment and Steve Drew and the RSNA for permission to replicate radiology report templates.

REFERENCES

1. Leslie KO, Rosai J. Standardization of the surgical pathology report: formats, templates, and synoptic reports. *Semin Diagn Pathol* 1994; **11**: 253–7.
2. Markel SF, Hirsch SD. Synoptic surgical pathology reporting. *Hum Pathol* 1991; **22**: 807–10. doi: [https://doi.org/10.1016/0046-8177\(91\)90209-8](https://doi.org/10.1016/0046-8177(91)90209-8)
3. Korman LY, Delvaux M, Bidgood D. Structured reporting in gastrointestinal endoscopy: integration with DICOM and minimal standard terminology. *Int J Med Inform* 1998; **48**: 201–6.
4. Sistrom CL, Honeyman-Buck J. Free text versus structured format: information transfer efficiency of radiology reports. *AJR Am J Roentgenol* 2005; **185**: 804–12. doi: <https://doi.org/10.2214/ajr.185.3.01850804>
5. Morgan TA, Helibrun ME, Kahn CE. Reporting initiative of the Radiological Society of North America: progress and new directions. *Radiology* 2014; **273**: 642–5. doi: <https://doi.org/10.1148/radiol.14141227>
6. Knechtges PM, Carlos RC. The evolving role of radiologists within the health care system. *J Am Coll Radiol* 2007; **4**: 626–35. doi: <https://doi.org/10.1016/j.jacr.2007.05.014>
7. Blackmore CC. Defining quality in radiology. *J Am Coll Radiol* 2007; **4**: 217–23. doi: <https://doi.org/10.1016/j.jacr.2006.11.014>
8. Grieve FM, Plumb AA, Khan SH. Radiology reporting: a general practitioner's perspective. *Br J Radiol* 2010; **83**: 17–22. doi: <https://doi.org/10.1259/bjr/16360063>
9. Plumb AA, Grieve FM, Khan SH. Survey of hospital clinicians' preferences regarding the format of radiology reports. *Clin Radiol* 2009; **64**: 386–94. doi: <https://doi.org/10.1016/j.crad.2008.11.009>
10. 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* 2011; **259**: 184–95. doi: <https://doi.org/10.1148/radiol.10101045>
11. Schwartz LH, Panicek DM, Berk AR, Li Y, Hricak H. Improving communication of diagnostic radiology findings through structured reporting. *Radiology* 2011; **260**: 174–81. doi: <https://doi.org/10.1148/radiol.11101913>
12. Quint DJ. Voice recognition: ready for prime time? *J Am Coll Radiol* 2007; **4**: 667–9. doi: <https://doi.org/10.1016/j.jacr.2007.03.020>
13. Boland GW, Guimaraes AS, Mueller PR. Radiology report turnaround: expectations and solutions. *Eur Radiol* 2008; **18**: 1326–8. doi: <https://doi.org/10.1007/s00330-008-0905-1>
14. Faggioni L, Coppola F, Ferrari R, Neri E, Regge D. Usage of structured reporting in radiological practice: results from an Italian online survey. *Eur Radiol* 2017; **27**: 1934–43. doi: <https://doi.org/10.1007/s00330-016-4553-6>
15. Towbin AJ, Hawkins CM. Use of a web-based calculator and a structured report generator to improve efficiency, accuracy, and consistency of radiology reporting. *J Digit Imaging* 2017; **30**: 584–8. doi: <https://doi.org/10.1007/s10278-017-9967-4>
16. Ghoshhajra BB, Lee AM, Ferencik M, Elmariah S, Margey RJ, Onuma O, et al. Interpreting the interpretations: the use of structured reporting improves referring clinicians' comprehension of coronary CT angiography reports. *J Am Coll Radiol* 2013; **10**: 432–8. doi: <https://doi.org/10.1016/j.jacr.2012.11.012>
17. Richardson CM, Pointon KS, Manhire AR, Macfarlane JT. Percutaneous lung biopsies: a survey of UK practice based on 5444 biopsies. *Br J Radiol* 2002; **75**: 731–5. doi: <https://doi.org/10.1259/bjr.75.897.750731>
18. Eisenberg RL, Bankier AA, Boieselle PM. Compliance with Fleischner Society guidelines for management of small lung nodules: a survey of 834 radiologists. *Radiology* 2010; **255**: 218–24. doi: <https://doi.org/10.1148/radiol.09091556>