Manuscript Number:

Title: Variation in the length and structure of reports written by reporting radiographers: a retrospective study

Article Type: Full Length Article

Keywords: Radiographer reporting; advanced practice; structured reporting; reporting variations.

Corresponding Author: Mr. Anthony S Manning-Stanley, MA(Cantab), BSc(Hons)

Corresponding Author's Institution: University of Liverpool

First Author: Anthony S Manning-Stanley, MA(Cantab), BSc(Hons)

Order of Authors: Anthony S Manning-Stanley, MA(Cantab), BSc(Hons); Laura Bonnett; Tara Mellett; Jennifer R Herreran; Rebecca Anforth

Suggested Reviewers:
15\textsuperscript{th} March 2018

Dear Prof. Nightingale,

**Manuscript for consideration: reporting variation in reporting radiographer reports**

Please find attached the relevant documents for the manuscript submission for the above article. I had the pleasure of your company during the afternoon session of UKRC 2017 where I presented my research findings; I have since written these findings up as a full article for consideration.

I can confirm that this article is not under consideration for publication elsewhere, and has not been previously submitted to another journal.

I accept responsibility for the integrity of the work as a whole and can confirm the listed authors have participated sufficiently to be listed as co-authors; all authors approve publication of the article in its current state.

As confirmed within the manuscript, no conflicts of interest exist or are anticipated to exist.

A funding statement has been included under ‘acknowledgements’ with Laura Bonnett identified as LB to preserve anonymity; I was unsure whether this is provided as part of the peer review process.

I look forward to receiving feedback on this article in due course, and thank you in advance for considering the article for publication in *Radiography*.

Kind regards,

Anthony S. Manning-Stanley
Variation in the length and structure of reports written by reporting radiographers: a retrospective study

Anthony S. Manning-Stanley\textsuperscript{a}; Laura Bonnett\textsuperscript{b}; Tara Mellett\textsuperscript{c}; Jennifer R. Herreran\textsuperscript{c}; Rebecca Anforth\textsuperscript{c}.

a. Directorate of Diagnostic Radiography, University of Liverpool, Johnston Building, Brownlow Hill, L69 3GB (corresponding author - antms@liv.ac.uk)

b. Department of Biostatistics, University of Liverpool, Waterhouse Building Block F, Brownlow Street, L69 3GL (ljbcmshe@liverpool.ac.uk)

c. Salford Royal NHS Foundation Trust, Stott Lane, Salford, M6 8HD (Tara.Mellett@srft.nhs.uk) (Jennifer.Herreran@srft.nhs.uk) (Rebecca.Anforth@srft.nhs.uk)
Introduction: The literature suggests that there is variation in various features of the written radiology report for a range of body areas and imaging modalities. The retrospective study presented here aims to determine if similar variation is demonstrated in a group of 5 reporting radiographers in a UK NHS Trust.

Methods: Full reports for 1,530 knee radiographic examinations performed from accident and emergency referrals were extracted for a 12-month period from a Radiology Information System (RIS) into Excel. Copied into Word, the word count function was used for each report and the number of words and characters (without spaces) was returned into Excel. Average word count and word length per report, by radiographer, were calculated for the following sections of the report: report title, main body and signature. SPSS was used to perform inferential statistical analysis.

Results: There was a wide range in the maximum and minimum average report lengths (60.88 v 17.83 words). Following log (base 10) transformation of the skewed average word count data, an ANOVA test demonstrated statistically significant differences (p<0.05) between all but one pair-wise comparison (Rad 2 v Rad 4; p=0.98). Average word length demonstrated less variation. 4 out of 5 radiographers always included a report title; 3 out of 5 never included a report signature.

Conclusion: Report length and structure for a group of 5 reporting radiographers demonstrates similar variation to that for radiologists described in the literature. Further research is required to determine if there is a clinically significant impact of this variation.
• Radiographer reports of knee radiographs demonstrate significant variation
• The variation manifests as differences in both length and structure
• The variation is comparable to that established for radiologist reports
Introduction

The written radiology report is the primary method of communicating the findings from imaging studies, and must accurately convey these findings to the referring physician in a timely manner \(^1\)[^2].

The report should contain no ambiguity, and clearly identify recommendations for further treatment or imaging, should they be required. Consistent language and nomenclature should be used, with the terminology contained within the report not interfering with the communicative process; as reported in the literature\(^3\) for some oncology staging reports, up to 16 different stakeholders may receive the report, so communicating results in an accurate and easily understood manner is essential. Radiographers are well placed in a team based approach, and through approved training, to fill the increased reporting demands placed upon imaging departments, as highlighted by the College of Radiographers (CoR), with the role of reporting radiographer now a common role-extension in the UK\(^4\).

As some authors have noted, historically, the process of writing a radiology report has been perceived more as an art\(^5\), with the decision as to the format and length of the report the result of preference and experience; attempts at standardisation may be resisted. The result is that different reporting practices may develop through variations in local preferences and experiences of the reporting community. For example, research investigating the structure and format of computed tomography (CT) reports of the abdomen across a number of sites in two countries with a common language found a wide variation in report structure\(^6\). Significant differences were found in word count, with averages higher in trainee versus qualified radiologists, University versus Community hospitals and in Flanders versus The Netherlands. Another study\(^7\) found significant differences in the inclusion of a conclusion, with the overall length of chest radiograph reports dependent upon whether they were written by specialist or generalist radiologists. A review of available literature to elucidate the important elements of a high-quality radiology written report identified a number of important components as targets for optimisation, including report length, language and format\(^8\).
Further research has also identified that the reporting styles, or the ‘report signature’ of individual radiologists is not only distinctive, but can be learnt through training a neural network; a machine learning approach led to 100% accuracy in identifying the report author in 60 reports, based on 12 lexical parameters of the report \cite{9}.

The retrospective review presented here aims to determine if the report length and structure for a cohort of five reporting radiographers from a single Trust in the North-West of England demonstrates similar variation to that described above for other professional groups which report radiological studies.
Method

Ethical consideration was given to the study. No randomisation to groups or patient care/treatment changes were required. It was also assumed that local practice at the Trust would not necessarily be generalizable to other Trusts given variation in local protocols, and variations in local practice identified above for radiologists. On this basis, use of the NHS Health Research Authority online decision tool\textsuperscript{[10]} subsequently confirmed this retrospective study did not require ethical approval.

The radiology information system (CRIS) for a major trauma centre in the North-West of England was interrogated for a 12-month period (1\textsuperscript{st} October 2015 to 30\textsuperscript{th} September 2016). Data including report text for every radiographic knee examination performed via Accident and Emergency was obtained through this query and exported into a spreadsheet (Microsoft Excel, 2013). The data was further filtered in the spreadsheet, and the report text for the five reporting radiographers was identified to give a total of 1,530 reports confirmed as authored by a reporting radiographer. For each report, the report text was copied and pasted into Microsoft Word (2013), with word and character counts (without spaces) for the overall report, the report title, the main body (findings) of the report and the report signature recorded. Average word length was calculated. Note was made of inclusion of a separate conclusion or impression section.

Data analysis was performed using Excel (descriptive statistics) and IBM Statistics SPSS (Version 24) (inferential statistics). For the inferential statistics, normality of the distribution of each variable was determined through inspection of the histograms for each data set to establish the degree of kurtosis; where data was not normally distributed an appropriate transformation was performed. To test for statistical significance the ANOVA with a Tukey add-on was performed, with a p-value of less than 0.05 considered statistically significant. ANOVA is considered a robust statistical test that can be used for normally distributed data or transformed data; simulation studies have demonstrated that ANOVA is not sensitive to moderate deviations from normality\textsuperscript{[11, 12, 13]}. 
To determine the consistency of the data analysis methodology a second researcher performed the analysis for a stratified sample of over 13% of the reports.

Results

Descriptive statistics

The breakdown of the total number of reports (1,530) for each reporting radiographer is shown in Table 1. The greatest number were reported by radiographer 4 (496), with the least by radiographer 1 (81).

Table 1.

As Figure 1 demonstrates, the two radiographers with the longest mean report length, in terms of the number of words used, were very similar, with radiographer 2 (mean 60.77, SD 29.05) slightly higher than radiographer 4 (mean 60.02, SD 23.18). In contrast radiographer 5 had the lowest mean word count (mean 17.83, SD 10.11), with radiographer 3 (mean 18.76, SD 17.5) slightly higher. The mean report length for radiographer 1 (mean 51.71, SD 29.39) was closer to that of radiographers 2 and 4.

Figure 1.

As tables 2, 3 and 4 demonstrate, there was variation in the use of a title and signature in the reports. Table 2 demonstrates the use of a title and signature across all reports (n=1,530), where 4 out of the 5 reporting radiographers always used a report title, while radiographer 3 only used a report title in 23.9% of reports. The use of a title was seen in 83.1% of all reports. In contrast, the use of a signature was much more variable, with Radiographer 1, 3 and 5 never using a signature, and
radiographers 2 and 4 using a signature in 86.3% and 91.9% of all reports. Overall, a signature was present in 55.4% of the reports.

**Table 2.**

Table 3 demonstrates that, where the report for the knee radiographs was the only examination reported for that attendance for that particular patient (n=1,230), the use of report titles and signatures demonstrated further variation. As above, the use of a report title was seen in 100% of reports for all radiographers except radiographer 3; now the use of a title was much lower (1.5% of reports) for radiographer 3, which reduced the overall use of a report title to 79%, from 83.1%.

Again, radiographers 1, 3 and 5 never used a report signature, but now the use of a report signature by radiographers 2 and 4 demonstrated an increase to 97.2% and 100% respectively. Subsequently, the overall use of a signature increased to 62.6% of reports, from 55.4%.

**Table 3.**

Table 4 demonstrates the use of report titles and signatures where the knee examination reported was part of 2 or more examinations performed on that patient for that attendance (n=300). For all radiographers, 100% of these reports included a title. In contrast, the inclusion of a signature for radiographers 2 (45.3%) and 4 (46.7%) was lower again, with the overall use of a signature decreasing to 26% of reports.

**Table 4.**

As demonstrated in Figure 2, radiographer 3 had the shortest mean word length (mean 5.54, SD 0.59), whilst radiographer 4 (mean 6.21, SD 0.38) was the longest, slightly longer than radiographer 2 (mean 6.19, SD 0.34).

**Figure 2.**

Reanalysis of reports
Following analysis of 205 (13.4%) of the reports by a second investigator, it was found that 5 out of 205 reports (2.4%) demonstrated differing values for either the length and/or identification of a report title or signature, compared to the original results. Closer inspection revealed there were 3 simple transcription errors in the re-analysis, with the values entered into the spreadsheet differing to the original analysis. For the remaining 2 discrepancies, these were due to the omission to include report text as a report title for one of the reporting radiographers; a full stop was not present after the initial statement of the laterality of the examination and the full body of text was considered to represent the body of the report without an explicit report title. For the remaining reports, where a full stop was present, the reanalysis was identical to the original analysis.

**Inferential statistics – overall word count**

Visual inspection of the histograms for the distribution of word counts, for each radiographer, reveals overlaid distribution curves which are skewed towards lower values, i.e. the data is not normally distributed; as an example, Figure 3 demonstrates the distribution of the overall word count per report for radiographer 3.

![Figure 3](image_url)

Given the skewed nature of the overall word count data, a log transformation (base 10) was applied and the ANOVA test was performed on this transformed data to determine if any statistically significant differences in the mean overall word count were seen between radiographers. As Table 7 demonstrates, statistically significant differences (p<0.05) between the mean overall word count for all but one pair of radiographers were found. The p-values for radiographer 2 versus 4 (p=0.98) indicates that there is no statistically significant difference between the mean word count for these pairings.

![Table 5](image_url)

**Inferential statistics – overall word length**
As figure 4 demonstrates, inspection of the histograms for the average word length for each radiographer reveals that only the values for radiographer 1 demonstrate variation from a true normal distribution. As such, an overall transformation of the data was not performed and the ANOVA test was again applied to the data.

**Figure 4.**

Table 8 demonstrates ANOVA analysis of the average word length, with statistically significant differences ($p<0.05$) seen between the mean word count for all but two pairs of radiographers. The $p$-values for radiographer 1 versus 5 ($p=0.778$) and for radiographer 2 versus 4 ($p=0.972$) indicate that there is no statistically significant difference between the overall mean word length for these pairings.

**Table 6.**
Discussion

As the results demonstrate, there are statistically significant differences in mean overall report length for a group of 5 reporting radiographers from a major trauma centre in the North-West of England, when the radiographic reports for knee examinations were considered over a 12-month period. This can be attributed to variations in various elements of the report; variation was seen in the overall number of words used, inclusion of a report title and signature, as well as the mean word length per radiographer.

This variation was similar to other published results for radiologists. For CT reporting of the abdomen\[6\], variation in the length of report was found to be statistically significant when country of report and the level of training were compared (resident v staff member: 160.5 v 122.9 mean words per report; p<0.05). The longest and shortest reports were 366 and 7 words respectively. A study investigating the quality of chest radiograph reports\[7\] found a statistically significant difference in mean report length for specialists versus general radiologists, with the specialist reports over three times the length (91 v 29 words; p=0.000). However, both long and short reports were considered clear by referrers, suggesting that the report length is not major determinant in the readability of that report.

The lack of the use of a conclusion or impression section for the reporting radiographers is also similar to other studies. In CT reporting of the abdomen\[6\], 13.5% of all reports did not contain a conclusion or impression section. A statistically significant difference was found\[7\] in the presence of a conclusion in a chest radiograph report written by a specialist compared to a general radiologist (22% v7%; p=0.000). The longest report of 228 words written by the radiographers studied here is comparable in length to a CT or magnetic resonance imaging (MRI) report. Whilst knee radiograph reports are based on less detailed imaging, compared to MRI or CT examinations, such radiographic reports may benefit from the inclusion of a conclusion to highlight the salient findings in cases where the report is over a certain length, or there are multiple paragraphs.
The inclusion of a report title in 100% of cases where the knee examination was part of a series of studies is considered good practice, as this delineates the report for this body area and correctly identifies the relevant body area with the correct report text. In other cases the variable use of a report title, and report signature, is considered to have developed through individual preference. It is worth noting that the viewing environment for the software employed by the Trust does prepopulate some title and signature detail, which may explain the variations in this practice.

The implications of the variability in the structure of the report, indicated here by the variable inclusion of report features and significant differences in mean word length, has also been discussed in the literature. Structured reporting (SR) has been offered as a solution, in terms of standardisation, with multiple proposed formats. Standardisation of terminology, recommendations and reporting has been adopted for various regions of anatomy through ‘reporting and data systems’ (RADS), including breast imaging (BI-RADS). Developed by the American College of Radiology (ACR), BI-RADS has led to an overall quality improvement\cite{14}. Given the success of these initiatives, further systems have been suggested, with a recent consensus\cite{15} developed for coronary artery disease (CAD-RADS).

SR was found\cite{7} to be significantly more complete and more effective in reporting chest radiographs. In a study investigating preferences of radiologists and ICU practitioners in relation to portable chest radiographs the referrers were found to prefer more complete, itemised structured reports describing support devices in detail\cite{16}. In the emergency setting, physicians preferred itemised reporting over point-and-click and basic structured reporting\cite{17}. However, preferences for SR amongst radiologists and referring physicians have not been universally established\cite{5, 16, 18}, with the potential benefits on patient outcomes still open to debate\cite{5}. Whilst some institutions have adopted SR across the whole range of Radiology examinations\cite{19}, this has not been a widespread phenomenon\cite{20}. One argument is that an itemised or checklist format interferes with the cognitive processes of reporting, taking the reporters gaze away from the images. In a bid to mitigate against
this, one study adopted a vocalised checklist, where the reporter does not have to turn away from
screen; this did not prevent the satisfaction of search effect in reporting chest radiographs as there
was a continued reluctance to report other abnormalities\(^{[21]}\). However, another study\(^{[22]}\) found that a
checklist-formatted SR reduced missed non-fracture findings on c-spine CT examinations. The mixed
results suggest that the effect of having a checklist style approach on the reporting process may be
modality and/or body region specific.

Several limitations to this study are acknowledged. Whilst a significant number of reports (n=1,530)
were analysed compared to other similar studies, only one radiographic examination was
considered; further analysis extended to other body areas is warranted. Also, the results are only
applicable to a local population of reporting radiographers from a single Trust and it is clear from the
variable inclusion of report titles and signatures that individual practices are not based on an agreed
local strategy. Again, further investigation of similar trends in other reporting radiographer
populations is recommended, as well as qualitative studies to elicit further information on the
drivers of the observed variation in report length and structure observed. Such studies should also
consider the minor variations that were found when a second investigator analysed a sub-set of the
images, as definitions of what constitutes a report title, the body of the report and the report
signature need to be explicitly defined in any data analysis protocol to prevent conflicting results.

The impact on patient outcome from variation in report structure and length also warrants further
investigation. For example, four cohort studies\(^{[3,23,24,25]}\) have demonstrated improved clinical
decision making in cancer staging with reports implementing structured reporting templates.
However, none extended this analysis to consider the impact upon diagnostic accuracy or patient
outcomes. Such considerations should be attended to in future research investigating variations in
radiographer reporting.
Conclusions

Statistically significant differences in the mean length of, and the mean word length in, radiographer reports demonstrate variation similar to other published studies. Whilst such reports may be improved through standardization, such as offered by the use of structured reporting (SR), it is noted that SR is not universally accepted, nor a definitive quality improvement strategy. Further quantitative studies investigating the impact of the observed variation in report structure and length on patient outcomes are required, as well as the application of qualitative methods to investigate the drivers for the differences observed in this small group of reporting radiographers.

Conflict of interest statement

No conflicts of interest are noted.
References:


radiology and referring physicians’, *Academic Radiology*, 22(6), pp. 760-70. [Available online. DOI: 10.1016/j.acra.2015.01.006; accessed 5th April 2017]


Funding

LB is funded by a Post-Doctoral Fellowship from the National Institute of Health Research (PDF-2015-08-044). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.
Figure 1. Mean report length per radiographer
Figure 3
Click here to download high resolution image
Figure 1. Number of reports per radiographer.

Figure 2. Mean report length per radiographer.

Figure 3. Distribution of the overall word count per report for radiographer 3.

Figure 4. Distribution of the average word length per report for each radiographer.
Table 1. Summary of report number per reporting radiographer

<table>
<thead>
<tr>
<th>Radiographer</th>
<th>Rad 1</th>
<th>Rad 2</th>
<th>Rad 3</th>
<th>Rad 4</th>
<th>Rad 5</th>
<th>% of total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Report Number</td>
<td>160</td>
<td>454</td>
<td>339</td>
<td>496</td>
<td>81</td>
<td>10.5%</td>
</tr>
<tr>
<td>% of total</td>
<td>10.5%</td>
<td>29.7%</td>
<td>22.2%</td>
<td>32.4%</td>
<td>5.3%</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Percentage use of Title and Signature for all examinations

<table>
<thead>
<tr>
<th></th>
<th>Rad 1</th>
<th>Rad 2</th>
<th>Rad 3</th>
<th>Rad 4</th>
<th>Rad 5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>142</td>
<td>359</td>
<td>262</td>
<td>421</td>
<td>46</td>
<td>1230</td>
</tr>
<tr>
<td>Title</td>
<td>100.0%</td>
<td>100.0%</td>
<td>1.5%</td>
<td>100.0%</td>
<td>100.0%</td>
<td>79.0%</td>
</tr>
<tr>
<td>Signature</td>
<td>0.0%</td>
<td>97.2%</td>
<td>0.0%</td>
<td>100.0%</td>
<td>0.0%</td>
<td>62.6%</td>
</tr>
</tbody>
</table>

Table 3. Percentage use of Title and Signature where only 1 examination

<table>
<thead>
<tr>
<th></th>
<th>Rad 1</th>
<th>Rad 2</th>
<th>Rad 3</th>
<th>Rad 4</th>
<th>Rad 5</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>18</td>
<td>95</td>
<td>77</td>
<td>75</td>
<td>35</td>
<td>300</td>
</tr>
<tr>
<td>Title</td>
<td>100.0%</td>
<td>100.0%</td>
<td>100.0%</td>
<td>100.0%</td>
<td>100.0%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Signature</td>
<td>0.0%</td>
<td>45.3%</td>
<td>0.0%</td>
<td>46.7%</td>
<td>0.0%</td>
<td>26.0%</td>
</tr>
</tbody>
</table>

Table 4. Percentage use of Title and Signature where more than 1 examination

<table>
<thead>
<tr>
<th></th>
<th>Rad 2</th>
<th>Rad 3</th>
<th>Rad 4</th>
<th>Rad 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rad 1</td>
<td>0.011</td>
<td>0.000</td>
<td>0.002</td>
<td>0.000</td>
</tr>
<tr>
<td>Rad 2</td>
<td>0.000</td>
<td>0.980</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Rad 3</td>
<td>0.000</td>
<td>0.007</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rad 4</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5. ANOVA test: p-values for radiographer pairings for log10(overall word count). Figures highlighted in bold are p-values > 0.05
<table>
<thead>
<tr>
<th></th>
<th>Rad 2</th>
<th>Rad 3</th>
<th>Rad 4</th>
<th>Rad 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rad 1</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.778</td>
</tr>
<tr>
<td>Rad 2</td>
<td>0.000</td>
<td>0.972</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>Rad 3</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Rad 4</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6. ANOVA test: p-values for radiographer pairings for overall average word length. Figures highlighted and in bold are p-values > 0.05.
Data Statement

Click here to download Data Statement: dataprofile.xml