Factors Impacting the Use of Terminology to Convey Diagnostic Certainty in Radiology Reports

Ronilda Lacson*, Victor Babatunde, Atul Shinagare and Ramin Khorasani

Department of Radiology, Brigham and Women's Hospital/Harvard Medical School, Boston, USA

Received: September 2, 2018 Revised: October 31, 2018 Accepted: November 6, 2018

Abstract:

Background:
Variable use of phrases expressing diagnostic uncertainty can lead to ambiguous radiology reports, a concern for information processing.

Objective:
This study aimed to quantify the usage of phrases conveying diagnostic certainty for abdominal imaging findings and assess factors that impact use of phrases with “good agreement” between radiologists and referring providers.

Methods:
This retrospective, Institutional Review Board-Approved study included all diagnostic reports generated by the Abdominal Radiology Division at an academic medical center July-September 2016. We assessed the use of 16 diagnostic certainty phrases using information retrieval from the Impression section of radiology reports. Phrases with good provider agreement for conveying the level of certainty are defined as “good agreement” phrases - including “diagnostic of”, “represents” and “unlikely.” We assessed the impact of imaging modality, trainee contribution to report generation, and individual radiologists.

Results:
In 5,598 radiology reports, 2,071 (37%) contained diagnostic certainty phrases, 119 (6%) of which were “good agreement” phrases. There was a significant difference between how frequently “good agreement” phrases were used in Magnetic Resonance Imaging (MRI), Computed Tomography (CT) and X-ray reports (p=0.0003). There was a significant variation among attending radiologists on the use of “good agreement” phrases (p<0.0019). There was no difference in the use of “good agreement” phrases in reports generated by attending radiologists alone compared to reports with trainees.

Conclusion:
Although phrases to convey diagnostic certainty were commonly used in radiology reports, the use of phrases with a good agreement was uncommon. Standardizing terminology to convey diagnostic certainty may reduce ambiguity in radiology reports and generate more accurate information processing tools.

Keywords: Radiology, Terminology, Uncertainty, MRI, CT, X-ray.

1. INTRODUCTION
Radiology reports serve as the primary method of communication between radiologists and other healthcare providers [1 - 3]. Many studies and interventions have been implemented to increase timely report generation, decrease
error rates in reporting, and provide comprehensive documentation of findings [4 - 10]. Information extraction and natural language processing tools have integrated certainty modifiers when extracting disease concepts [11]. However, despite multiple studies reporting radiologists’ usage of vague and uncertain phrases in their reports [3, 12 - 14], there has been little effort to evaluate the use of terminology conveying diagnostic certainty in radiology reports, and whether providers agree on the level of certainty they convey.

While ideally, radiology reports are concise and unambiguous [15, 16], in practice, there is wide variability in the terminology used by radiologists in reporting findings, depending on their degree of certainty and judgment [3, 9]. This variability in expressing certainty has been significantly associated with reports that are more ambiguous [12]. Standardized reporting has been shown to decrease ambiguity for particular radiology procedures [17 - 20]. Breast Imaging Reporting and Data System [BI-RADS] for breast imaging reporting has been utilized to help understand and communicate the probability of a pathologic finding [19]. Similarly, Prostate Imaging Reporting and Data System (PI-RADS) and Lung Reporting and Data System (Lung-RADS) have been used to communicate the probability of prostate and lung cancer, respectively [17, 18]. However, expressing the probability of other imaging findings, other than these few malignancies, is a difficult task. Moreover, the standardized terminology for conveying diagnostic certainty in abdominal radiology reports is not yet established.

A previous study identified 15 commonly-used phrases utilized by radiologists for expressing certainty [3]. In that study, the two phrases which had the highest agreement among radiologists in conveying the level certainty were “diagnostic of” and “unlikely”, based on a questionnaire wherein radiologists voluntarily ranked 15 phrases in order of diagnostic certainty. The phrase “diagnostic for” and “represents” were ranked with the highest likelihood in conveying malignancy by both radiologists and primary care physicians in a separate study [13]. The purpose of this study was to assess the usage of 16 phrases to convey diagnostic certainty for imaging findings in the Abdominal Division of an academic radiology practice [3, 12], specifically to quantify the usage of phrases with “good agreement” between radiologists and referring providers that are expected to express less ambiguity. Understanding the level of variability and factors that impacted the usage of phrases with “good agreement” will help in developing standard terminology to convey diagnostic certainty.

2. MATERIALS AND METHODS

2.1. Study Setting and Cohort

This retrospective, Health Insurance Portability and Accountability Act-compliant study was approved by the Institutional Review Board, and performed at an urban, quaternary care, academic medical center. All abdominal X-ray, Computed Tomography (CT) scan and Magnetic Resonance Imaging (MRI) radiology reports generated by the Abdominal Radiology Division in the third quarter of 2016 (July-September 2016) were included in the analysis. Exclusion criteria included reports from other modalities (e.g., Upper Gastro-Intestinal series, Hysterosalpingogram) and abdominal interventional procedures (e.g., paracentesis). The radiology department has 13 abdominal radiology subspecialty-trained attending physicians, 5 abdominal radiology subspecialty fellows and 39 diagnostic radiology residents. Every finalized radiology report is signed by one of the attending physicians, whether or not it was first reviewed by a trainee.

2.2. Phrases Conveying Diagnostic Certainty

We assessed radiology reports for the use of 16 previously-identified phrases that convey diagnostic certainty [3]. In particular, we evaluated the use of three phrases for which radiologists had “good agreement” with referring providers in conveying diagnostic certainty - “diagnostic of/for”, “represents” and “unlikely” - as well as 13 phrases for which there was variable agreement among radiologists and non-radiologists about conveying diagnostic certainty [3].

2.3. Data Collection

We utilized information retrieval and query matching to extract phrases conveying diagnostic certainty from within the impression sections of radiology reports. We utilized a publicly available toolkit, Information from Searching Content with an Ontology-Utilizing Toolkit (iSCOUT) for this task [21]. The Header Extractor component in iSCOUT enables selective retrieval of sections of reports by restricting a search to content within various headers of a report.
Factors Impacting the Use of Terminology to Convey The Open Medical Informatics Journal, 2018, Volume 12

(e.g., Impression, Findings). From the Impression section, we subsequently performed query matching for each individual diagnostic certainty phrase using Microsoft Access (Microsoft Corporation, Redmond, Washington). Reports that included at least one of the 16 phrases conveying diagnostic certainty were retrieved.

2.4. Outcomes and Statistical Analysis

We assessed the proportion of reports with phrases conveying diagnostic certainty, as well as the frequency and distribution of those phrases. In addition, we calculated the proportion of reports with “good agreement” phrases out of all reports with phrases that convey diagnostic certainty. Each report that contains one or more phrases that convey diagnostic certainty is counted once in the denominator. Each report is counted once in the numerator (i.e. reports with “good agreement” phrases) if it contains at least one “good agreement” phrase. We assessed the impact of imaging modality, trainee contribution to report generation and individual radiologists in the usage of “good agreement” phrases. Chi-square test was performed to evaluate the difference between the proportion of reports with “good agreement” phrases and reports from X-ray, CT scan and MRI examinations, between attending-only reports or joint attending and trainee reports, and between the individual attendings. For comparing individual attendings, those who contributed fewer than 100 reports were excluded from statistical analysis. All statistical analyses were performed using commercially available software (R 3.4.1 software, Vienna, Austria). A p-value of <0.05 was considered statistically significant.

3. RESULTS

3.1. Cohort Characteristics

During the study period, a total of 5,598 unique radiology reports were generated by the Abdominal Radiology Section; 1,226 (22%) abdominal X-ray reports, 3,010 (54%) CT scan reports and 1,362 (24%) MRI reports. There were 1,734 attending-only reports and 3,864 joint attending and trainee reports. The 13 radiologist attendings accounted for an average of 430 reports each. Among 5,598 reports, 2,071 (37%) contained at least one diagnostic certainty phrase.

3.2. Frequency and Distribution of Diagnostic Certainty Phrases

Fig. (1) demonstrates the proportion of reports containing phrases used to convey diagnostic certainty out of all reports, by imaging modality (examples of phrase usage are provided in Table 1). Only 119 (6%) of phrases were “good agreement” phrases. The most commonly used diagnostic certainty phrases in abdominal imaging were “likely” (including with modifiers), and “may represent”, accounting for 15% and 5%, respectively, of all reports with phrases that convey diagnostic certainty. Specifically, in abdominal x-ray reports, the three most common phrases were “likely”, “may represent”, and “consistent with”; in CT scans, the most common words were “likely”, “may represent” and “may be”; in MRI imaging, the most frequently used phrases were “likely”, “likely (with modifiers)”, and “consistent with”.

Table 1. Examples of diagnostic certainty phrase usage in the Impression section of the radiology report.

<table>
<thead>
<tr>
<th>Diagnostic Certainty Phrase</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likely</td>
<td>New mild compression of superior endplates of T12 and L3 likely degenerative</td>
</tr>
<tr>
<td>Probably</td>
<td>Two tiny cystic lesions in segment 6 are probably benign</td>
</tr>
<tr>
<td>Suspicious of</td>
<td>PI-RADS* 4 lesion within the left mid gland posterior lateral peripheral zone, with contour suspicious of left extracapsular extension</td>
</tr>
<tr>
<td>Diagnostic of</td>
<td>Multiple hepatic masses measuring up to 2.5 cm most with signal characteristics diagnostic of hemangiomas</td>
</tr>
<tr>
<td>Unlikely</td>
<td>Increase in the size of a sclerotic lesion in L1. In the absence of the known malignancy this is unlikely to be metastasis</td>
</tr>
<tr>
<td>Most likely</td>
<td>Bilateral hypodense renal lesions, most likely cysts</td>
</tr>
</tbody>
</table>

*PI-RADS: Prostate Imaging Reporting and Data System
3.3. Diagnostic Certainty Phrases by Imaging Modality

Table 2 shows the results of bivariate analyses comparing use of phrases with “good agreement” between imaging modalities. There was significant difference in how frequently phrases with “good agreement” were used as compared with phrases with variable agreement between imaging modalities. Reports from MRI and CT exams had a higher usage of “good agreement” phrases compared to X-ray reports (p=0.0003; chi-square analysis).

Table 2. Bivariate analysis comparing frequency of diagnostic certainty phrases by imaging modality in reports that contain diagnostic certainty phrases (n=2,071/5,598).

<table>
<thead>
<tr>
<th>Imaging Modality</th>
<th>Reports that Contain Phrases with Good Agreement / Reports that Contain Diagnostic Certainty Phrases</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-ray (n=1,226)</td>
<td>6/331</td>
<td>2%</td>
</tr>
<tr>
<td>CT (n=3,010)</td>
<td>67/1,159</td>
<td>6%</td>
</tr>
<tr>
<td>MRI (n=1,362)</td>
<td>48/581</td>
<td>8%</td>
</tr>
</tbody>
</table>

RESULTS

Chi-square Statistic (comparing X-ray with CT+MRI) 16.00

P-value 0.0003

3.4. Diagnostic Certainty Phrases by Generating Physician

Table 3 shows the results of bivariate analyses comparing use of phrases with “good agreement” between reports generated by attending physicians, and trainees with attending physicians. No significant difference was noted in usage of phrases with “good agreement” between these two groups (p<0.5169). Fig. (2) shows the proportion of reports containing phrases used to convey diagnostic certainty out of all reports generated by attending-only and trainees with attending physicians. Reports that were generated by attending physicians had “likely”, “may be” and “may represent” as most commonly used phrases, whereas joint reports from both attending and trainee physicians most commonly contained “likely”, “may represent” and “consistent with”.

Fig. (1). Proportion of reports containing phrases (n=16) used to convey diagnostic certainty out of all reports, by modality.
Table 3. Bivariate analysis comparing frequency of diagnostic certainty phrases in reports generated by attending physicians, and trainees with attending physicians that contain diagnostic certainty phrases.

<table>
<thead>
<tr>
<th>Reports</th>
<th>Reports that Contain Phrases With Good Agreement/ Reports that Contain Diagnostic Certainty Phrases</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attending-only reports (n=1,734)</td>
<td>29/559</td>
<td>5%</td>
</tr>
<tr>
<td>Trainee &amp; Attending joint reports (n=3,864)</td>
<td>90/1,512</td>
<td>6%</td>
</tr>
<tr>
<td>TOTAL (n=5,598)</td>
<td>119/2,071</td>
<td>6%</td>
</tr>
</tbody>
</table>

RESULTS

Chi-square Statistic

0.44

P-value

<0.5169*

* statistically significant

Fig. (2). Proportion of reports containing phrases (n=16) used to convey diagnostic certainty in attending-only and trainees with attending physician reports.

3.5. Diagnostic Certainty Phrases by Attending Physician

Individual attending physicians used diagnostic certainty phrases with “good agreement” in 0-14% of their reports that contain diagnostic certainty phrases. Table 4 demonstrates a significant difference in use of phrases with “good agreement” in reports generated between individual attending physicians (p<0.0019).

Table 4. Bivariate analysis comparing frequency of diagnostic certainty phrases in reports generated by individual attending physicians that contain diagnostic certainty phrases.

<table>
<thead>
<tr>
<th>Attendings</th>
<th>Reports that Contain Phrases With “Good Agreement”/ Reports that Contain Diagnostic Certainty Phrases</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attending 1</td>
<td>13/156</td>
<td>8%</td>
</tr>
<tr>
<td>Attending 2</td>
<td>11/221</td>
<td>5%</td>
</tr>
<tr>
<td>Attending 3</td>
<td>18/133</td>
<td>14%</td>
</tr>
<tr>
<td>Attending 4</td>
<td>19/378</td>
<td>5%</td>
</tr>
<tr>
<td>Attending 5**</td>
<td>4/73</td>
<td>5%</td>
</tr>
<tr>
<td>Attending 6**</td>
<td>5/74</td>
<td>7%</td>
</tr>
<tr>
<td>Attending 7</td>
<td>3/133</td>
<td>2%</td>
</tr>
</tbody>
</table>
Attendings | Reports that Contain Phrases With “Good Agreement”/Reports that Contain Diagnostic Certainty Phrases | Percent
---|---|---
Attending 8 | 8/161 | 5%
Attending 9** | 0/40 | 0%
Attending 10 | 16/256 | 6%
Attending 11 | 19/284 | 7%
Attending 12** | 1/36 | 3%
Attending 13 | 2/126 | 2%
TOTALS | 119/2,071 | 6%
RESULTS
Chi-square Statistic | 24.52
P-value | <0.0019*

* statistically significant ** excluded from statistical analysis (phrase usage <100)

4. DISCUSSION

In a previous study, surveyed radiologists and non-radiologists ranked the phrases “diagnostic of” and “represents” as the phrases which conveyed the highest degree of certainty for expressing diagnostic findings in radiology reports [3]. Phrases such as “likely” and “consistent with” were generally considered to have a variable agreement in conveying diagnostic certainty [3]. However, although radiologists perceive that there is a variable agreement in conveying certainty in using these phrases, in this study, these were frequently encountered phrases in radiology reports.

These results confirm previous findings that radiologists frequently utilize phrases which convey diagnostic uncertainty. Multiple publications express the need for minimizing uncertainty and ambiguity when reporting findings [12, 22]; nevertheless, there continues to be the widespread utilization of these phrases. Uncertainty in diagnostic testing is unavoidable because the sensitivity and specificity of tests is rarely 100%. MRI and CT, for instance, are more accurate for diagnosing metastatic tumors in the peritoneum compared to US. MRI is more accurate for detection of lymph node metastases from ovarian cancer than CT [23]. Thus, rather than ignoring uncertainty, it is critical to focus on factors that impact the use of diagnostic certainty phrases, and potentially standardize how certainty is conveyed in reports.

Factors that were significantly associated with utilizing diagnostic certainty phrases with a variable agreement for conveying diagnostic certainty included imaging modality (i.e., X-ray), generating reports without trainees, and individual attending physicians. As expected, there is more frequent use of “good agreement” phrases in MRI and CT reports, compared to X-ray, as these tests typically provide more diagnostic detail than X-rays. There may also be the concurrent use of “good” and variable agreement terms in the same report. However, we focused on comparing usage of “good agreement” terms in each report. Reports generated with trainees compared to those generated by attendings alone did not result in more frequent use of phrases with good agreement. In addition, there was variation in use of “good agreement” phrases between attending physicians.

Communication of uncertainty has been referred to as hedging, with unnecessarily ambiguous expressions [24]. Communicating ambiguity, however, has more often resulted in mismanagement and delayed care [16]. Thus, radiologists have recognized that this practice does not protect from litigation [24]. Rather, it is driven by the uncertainty that clinicians face [25].

It appears that individual attending physicians need education in order to minimize the use of phrases that do not have “good agreement”. However, educating physicians alone may not be an optimal approach for changing behavior without adequately described standards for communicating diagnostic certainty [26, 27]. As stated previously, uncertainty is unavoidable. Thus, having the means to convey uncertainty in a way that expresses the degree of certainty, like diagnostic test statistics (i.e. 99% sensitivity is greater than 80%), is necessary. Standardizing phrases for conveying certainty might aid in enhancing communication and minimizing ambiguity [22, 28]. This is especially critical for natural language processing and information extraction tools so that findings and disease concepts are described with appropriate certainty modifiers.

4.1. Limitations

This study is limited to phrases that were commonly encountered in our institution and have been previously published. These phrases may vary by geographic location, practice, and institution. We only analyzed reports from a...
single academic medical center, and our findings may not reflect terminology usage that conveys diagnostic certainty in other institutions. We expect that there is greater variability in terminology usage than identified within one institution alone, highlighting the importance of assessing terminology usage. In addition, the use of diagnostic certainty phrases was not assessed in the context of the findings being described. For example, physicians may be more likely to use phrases with “good agreement” for more benign findings as opposed to findings suggesting malignancy. However, the use of uncertain phrases has been previously shown to lead to ambiguous reports [12].

Further studies to enhance communication between providers by providing terminology to express their degree of certainty for findings are needed. In some cases, developing more standard terminology based on diagnostic confidence in specific imaging findings or patient outcomes may be useful. Standardizing phrases that convey certainty, and limiting the use of phrases with variable agreement among physicians may improve communication of radiology results.

CONCLUSION

In abdominal imaging reports, more than a third of reports contained at least one diagnostic certainty phrase, however, phrases with good agreement between physicians were infrequently used. Standardization of terminology used to convey diagnostic certainty may help create less ambiguous reports and improve quality of care. Furthermore, it can generate more accurate and reliable information processing tools for research and quality initiatives.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Health Insurance Portability and Accountability Act-compliant study was approved by the Institutional Review Board.

HUMAN AND ANIMAL RIGHTS

No Animals were used in this research. All human research procedures followed were in accordance with the ethical standards of the committee responsible for human experimentation (institutional and national), and with the Helsinki Declaration of 1975, as revised in 2013.

CONSENT FOR PUBLICATION

Not applicable.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

The authors would like to thank Laura Peterson for reviewing the manuscript.

REFERENCES


© 2018 Lacson et al.
This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Public License (CC-BY 4.0), a copy of which is available at: https://creativecommons.org/licenses/by/4.0/legalcode. This license permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.