

Using Analysis of Speech and Linguistics to Characterize Uncertainty in Radiology Reporting

Bruce I. Reiner

Published online: 2 October 2012
© Society for Imaging Informatics in Medicine 2012

Introduction

It has been stated that the only certainty in medicine is the pervasiveness of uncertainty [1, 2]. While medicine is often viewed as a science driven by objective data, it must also be considered an art, driven by subjective and often imprecise thought. This imprecise thought often takes the form of uncertainty, which has been described as “a state of having *limited knowledge* where it is improbable to exactly describe existing state or future outcome” [3]. Ultimately, the underlying goal of medicine (and radiology reporting) is to impart knowledge through accurate and confident decision-making and analysis, leading to improved clinical outcomes.

While one would hope that medical decision-making is entirely based upon sound and rational use of scientific data (i.e., evidence-based medicine), this is not always the case in everyday practice. With the increasing volume and complexity of medical data and heightened productivity expectations, healthcare providers are increasingly challenged in rendering accurate and timely decisions [4, 5]. At the same time, the rapid growth of technology and scientific knowledge has paradoxically created an increase in uncertainty in medical decision-making [6]. This is especially true for radiology providers, who are often tasked with interpreting complex imaging exams in the absence of comprehensive clinical data, which in part is due to the lack of clinical and imaging information systems integration [7]. The net result is often contrary to the desired goal. Indecisive decisions are often rendered in the radiology report in an attempt to fulfill workflow expectations, which in turn can lead to increased occupational stress and additional error [8].

While this would imply that uncertainty in medicine decision-making is inherently bad, this is not necessarily true. Uncertainty does not always imply limitations on the part of the physician, but can instead be the result of insufficient, poor quality, or inconclusive data (i.e., external uncertainty). The alternative form of uncertainty is internal uncertainty, which is due to an individual's lack of knowledge or decisiveness and implies deficiency on the part of the individual decision-maker. Both forms of uncertainty permeate medical practice and play an important role in clinical outcomes. As a result, it is important to understand the causative factors associated with uncertainty, develop effective methods for quantifying and characterizing uncertainty, and create intervention strategies to combat uncertainty related to deficient knowledge and/or decisiveness.

Uncertainty in Radiology Reporting

Uncertainty has been described as the Achilles heel of the radiology report [9], which is the single most important basis on which radiologists are judged by their clinical colleagues [10]. A number of diverse external factors contribute to radiology report uncertainty including technical (e.g., poor image quality), clinical (e.g., insufficient clinical data), medico legal (e.g., increased risk of litigation), anatomic (e.g., anatomic variation), and societal (e.g., lack of established standards). The interpreting radiologist is tasked with rendering an accurate diagnosis with the available data in a clear, concise, and definitive fashion. Often times however, ambiguity and uncertainty is communicated in the report, which may lead to confusion and/or lack of confidence on the part of the referring clinician. When a radiologist and referring clinician have a longstanding and close working relationship with one another, this report uncertainty can often be properly placed in context and clarified

B. I. Reiner (✉)
Department of Radiology,
Veterans Affairs Maryland Healthcare System,
10 North Greene Street,
Baltimore, MD 21201, USA
e-mail: breiner1@comcast.net

through direct communication. The changing dynamics of radiology practice has however minimized this sense of familiarity through radiology outsourcing (e.g., teleradiology) while also reducing direct communication through the widespread implementation of PACS [11] and increased workload demands [12]. The resulting “disengagement” of radiologists and clinicians has the potential to increase the clinical consequences of radiology report uncertainty and heightens the necessity for proactive intervention.

Analysis of Language to Assess Uncertainty

The degree of uncertainty (i.e., decreased diagnostic confidence) associated with a judgment or decision plays a major role in how that judgment will be subsequently used, as well as the perceived quality [13, 14]. While confidence and certainty are commonly viewed positively, they do not necessarily correlate with accuracy and can lead to a false sense of security and potentially adverse consequences. The ultimate goal therefore is to accurately determine the relationship between uncertainty and outcome measures in order to maximize the quality of task performance.

In most forms of communication, people prefer the use of verbal expressions rather than quantitative numerical data in expressing uncertainty [15]. The specific terminology used to express uncertainty falls into a continuum (i.e., confidence–uncertainty continuum), with varying degrees and usage of qualifiers and intensifiers. Qualifiers are words used to strengthen or weaken the observation or judgment being communicated in a manner which can either strengthen (e.g., certainly) or weaken (e.g., probably) the observation of record. An illustration of this confidence–uncertainty continuum can be found in the observation of a hepatic cavernous hemangioma on a CT report.

- (a) An enhancing 3.5 cm liver lesion is present, *consistent with* a cavernous hemangioma.
- (b) An enhancing 3.5 cm liver lesion is present, which *probably represents* a cavernous hemangioma.
- (c) An enhancing 3.5 cm liver lesion is present, which is *possibly* a cavernous hemangioma.

In addition to the fairly straightforward use of terms of uncertainty, a more insidious method of introducing uncertainty in reporting is through the use of recommendations and suggestions. By including a follow-up recommendation to the observation or judgment the reader can introduce uncertainty without explicitly stating it, as illustrated in the example of the cavernous hemangioma.

- (a) An enhancing 3.5 cm liver lesion is present, *consistent with* a cavernous hemangioma.

- (b) An enhancing 3.5 cm liver lesion is present, *consistent with* a cavernous hemangioma, *which can be confirmed with ultrasound or MRI*.

While both sentences contain the same finding and degree of confidence (i.e., consistent with); the addition of a follow-up recommendation adds a measure of uncertainty, without explicitly stating it through terms of uncertainty. This has the potential to adversely influence outcomes without directly affecting diagnostic accuracy; through delayed diagnosis, additional testing, and increased cost. This illustrates the fact that report uncertainty can introduce “hidden” costs to medical care, without directly introducing diagnostic error.

Another form of communication used to express uncertainty is a disclaimer, which is defined as an introductory expression used to explain or attempt to rationalize uncertainty. An example of a commonly used disclaimer in radiology reporting relates to quality deficiencies; which can take the forms of poor technique, patient noncompliance, or limitations in clinical data. These disclaimers may serve a vital function in communicating external causes of uncertainty, which should affect the accuracy, reliability, and confidence in which the report findings are interpreted and acted upon.

To date, the assessment of uncertainty and diagnostic confidence in radiology reporting has been largely subjective in nature and left to each individual reader’s perceptions. As data mining and knowledge discovery applications begin to play greater roles in medical reporting and analysis, technology such as natural language processing (NLP) can be applied to the task of uncertainty detection and analysis with the ultimate goal of correlating report uncertainty, diagnostic accuracy, and clinical outcomes [9]. Analysis of report uncertainty could potentially provide important insights as to the contextual and user-specific factors associated with uncertainty, and subsequent impact uncertainty plays on clinical outcomes.

Analysis of Speech to Assess Stress and Uncertainty

Once the clinical imperative of report uncertainty has been established, the next step is creating an objective (i.e., non-operator dependent) method for report uncertainty analysis. One proposed method is the application of NLP to identifying and quantifying report uncertainty through linguistic (i.e., text based) analysis. From a practical perspective, the NLP application could run in parallel to the method of report input (e.g., speech recognition), and visually highlight report language associated with uncertainty for real-time review (and possible intervention) by the authoring radiologist. Once the report has been finalized, the NLP application could record the derived uncertainty data into an uncertainty database along with a

number of associated clinical, technical, and patient-specific data elements for longitudinal analysis (Table 1). The correlation of these multiple data elements becomes critical in understanding the causative factors, confounding variables, and clinical ramifications of report uncertainty in medical imaging.

In addition to identifying report uncertainty through textual analysis using NLP, a potentially complementary analysis of uncertainty can be performing through speech analysis. The premise for this application lies in the longstanding use of voice stress analysis (VSA) for lie detection [16–18]. The theory behind this application is that deception will elicit physiologic changes in speech, which can be objectively identified through computerized speech analysis [19–22]. While deception in law enforcement applications is generally thought of as a deliberate act to mislead, the proposed application of uncertainty detection in medical reporting is instead one of non-deliberate (i.e. unintentional) intent. In this case, as the radiologist is creating the report, he/she may subconsciously introduce uncertainty through equivocation and hesitancy. Regardless of the deliberateness of the act, the same theoretical tenets hold true for both applications. Deception, in the form of deliberate lies or non-deliberate equivocation will in theory provide physiologic change in speech, which differs from baseline (i.e., control) speech patterns [23]. At the same time, equivocation employs the use of adjectives and adverbs in order to qualify the meaning of statements and introduce ambiguity and uncertainty [24].

In order to detect subtle variations in voice stress and improve the accuracy of voice stress analysis, a personalized voice profile can be created for each individual end-user, which can be readily accomplished through the routine use of speech recognition technologies for medical reporting. While conventional speech recognition applications create a personalized voice profile related to the unique elements of

each end-user's voice and vocabulary, the proposed technology would significantly expand the purview of these voice profiles by correlating intrinsic voice characteristics with real-time stress measures using voice stress analysis. The resulting personalized voice stress profile would create a personalized record depicting how an individual end-user's intrinsic speech characteristics change over time in accordance with changing measures of stress, workload, and task complexity. By correlating these data with performance outcome measures (e.g., diagnostic accuracy), one could create a method for predicting the relationship between stress-induced changes in voice, subconscious deception (i.e., uncertainty, lack of confidence), and task performance. Task performance can be analyzed by assessing the specific task being performed (e.g., exam type), complexity (e.g., size of dataset), clinical context (e.g., clinical indication), and quality measures (e.g., diagnostic accuracy). In the example of a radiologist tasked with interpretation of a brain MRI, the combined data could provide insight as to how the individual radiologist's speech characteristics and stress changed relative to baseline during the course of interpretation and reporting, how this compares with measures of different exam types and complexities, and how this correlates with outcome analysis (e.g., diagnostic accuracy based upon peer review or additional test results). This analysis could even be extended on a more granular level to assess uncertainty related to individual findings within a single report.

In addition to measuring and analyzing periodic fluctuations in stress specific to context and individual user baseline, speech analysis could also be used to identify specific speech patterns associated with increased uncertainty such as such as hesitations in speech, increased frequency of specific words (e.g., ums and uhs), and longer latency periods [25]. The end result would be a comprehensive speech analysis system which identifies uncertainty through language (i.e., report content), speech patterns, and end-user-specific voice stress variations.

A number of limitations exist in the use of VSA for traditional lie detection, which in theory could be obviated by the proposed technology. These include the degree of inter-operator variability and subjectivity, ability for subjects to “game” the system, and lack of extensive baseline data relative to the subjects being evaluated. The “gaming” of the system has been well described by the use of physical (e.g., tongue biting) and mental (e.g., counting backwards) countermeasures subjects use when answering control questions in order to alter stress responses and confuse analysis of test questions [26]. The proposed technology would counter these limitations through the use of objective computerized analysis of voice characteristics and stress, along with the creation of an in-depth personalized voice profile based upon longitudinal data collection over extended use of speech recognition.

Table 1 Associated data for report uncertainty analysis

1	Clinical data (clinical indication, medical/surgical history, laboratory/pathology data)
2	Exam and protocol selection (exam appropriateness, protocol optimization)
3	Image quality (contrast/spatial resolution, positioning, artifacts, exposure index, image processing)
4	Technology used (imaging modality, information system technologies, decision support)
5	Patient profile characteristics (age, body habitus, compliance, morbidity, ambulatory status)
6	Exam complexity (modality, anatomic region, size and complexity of imaging dataset)
7	Historical imaging data (availability of correlating imaging exams and reports)
8	Report findings (characteristics of abnormalities reported [size, conspicuity, chronicity, and morphology])

Implementation and Intervention Strategies

The ability to combine language and speech databases for uncertainty detection and characterization could provide a synergistic mechanism for understanding how uncertainty is manifested in both context- and user-specific fashions. As the individual language and voice stress uncertainty databases identify elevated uncertainty measures, automated prompts can be sent to notify the end-user of the observation, along with statistical data relating the magnitude and of the measurement relative to baseline uncertainty measures. The corresponding uncertainty data can in turn be correlated with outcomes analysis data (e.g., diagnostic accuracy) to establish the relationship between uncertainty and quality specific to the context of the task being performed (e.g., clinical indication, exam type) and individual user. The methodology for longitudinal outcomes analysis has been previously described and would consist of user and context-specific databases which correlate individual report findings with subsequent report data, peer review, clinical testing, laboratory data, and clinical feedback [27, 28].

Realizing that each individual end-user would have their own preferences as to how the uncertainty data would be communicated and acted upon, customized implementation pathways could be created to maximize technology acceptance and integration into routine workflow. For example, radiologists may differ in accordance with how the uncertainty prompts would be communicated (e.g., visual versus auditory), when the action would take place (e.g., immediate versus end of task completion), the specific threshold which mandates an automated prompt, and intervention options (e.g., computerized decision support and data mining tools).

The ability to correlate uncertainty and outcomes data could potentially provide insights into the clinical significance of uncertainty and clinical value of real-time intervention. As an example, if quality assurance (QA) analysis of an individual radiologist's database shows a specific and reproducible uncertainty pattern associated with higher than expected QA discrepancies, then this user-specific voice stress pattern could be prospectively programmed to identify "high risk" reports (or individual report findings) associated with inaccurate diagnosis. The ultimate goal would be the creation of a context and user-specific database which could prospectively identify specific voice stress and language patterns associated with poor outcomes data. By identifying these patterns in real time, automated alerts and prompts could be sent to notify the radiologist of the concern, present supporting data, and present customizable intervention options (e.g., computerized decision support, subspecialist consultations, context-specific educational aides, and supplemental clinical data).

Conclusion

For better or worse, uncertainty is an inevitable component of medical practice, which has the potential to affect the perceived quality of service delivery. In order to accurately address the clinical impact uncertainty has on clinical care, it is essential that uncertainty be characterized into external and internal forms, tracked and analyzed in accordance with its context- and user-specific attributes, and correlated with clinical outcomes. Speech analysis, both in written and verbal forms provides an accessible means to accomplish these tasks through the routine use of speech recognition technologies, which have become ubiquitous in everyday medical practice. The creation of personalized speech profiles creates a portable means for tracking and analyzing uncertainty for each individual end-user and creating customizable intervention strategies to maximize quality performance. Rather than reject this technology out of concerns for intrusiveness and excessive oversight, the radiologist community should embrace the opportunity to better understand uncertainty and intervene at the point of care. This could in theory offer improvement to both subjective and objective measures of radiology report quality, which remains the single most important goal of radiology service providers.

References

1. Blanch DC, Hall JA, Roker DL, et al: Is it good to express uncertainty to a patient? Correlates and consequences for medical students in a standardized patient visit. *Patient Educ Consult* 76:300–306, 2009
2. Luther VP, Crandall SJ: Ambiguity and uncertainty: neglected elements of medical education curricula? *Acad Med* 86:799–800, 2011
3. <http://en.wikipedia.org/wiki/uncertainty>
4. Haynes B, Haines A: Getting research findings into practice: barriers and bridges to evidence based clinical practice. *Br Med J* 317:273–276, 1998
5. McAlister FA, Straus SE, Guyatt GH, et al: Users' guides to the medical literature: integrating research evidence with the care of the individual patient. *JAMA* 283:2829–2836, 2000
6. Gerrity MS, Earp JAL, DeVellis RF, et al: Uncertainty and professional work: perceptions of physicians in clinical practice. *Am J Sociol* 97:1022–1051, 1992
7. Reiner BI: Medical imaging data reconciliation. Part 2: clinical order entry/imaging report data reconciliation. *J Am Coll Radiol* 10:720–724, 2011
8. Reiner B, Kruspinski EA: The insidious problem of fatigue in medical imaging practice. *J Digit Imaging* 1:3–6, 2012
9. Reiner B: Uncovering and improving upon the inherent deficiencies of radiology reporting through data mining. *J Digit Imaging* 23(2):109–118, 2010
10. Reiner BI, Siegel EL, Knight N: Radiology reporting: past, present, and future: the radiologist perspective. *J Am Coll Radiol* 5:313–319, 2007
11. Reiner B, Siegel E, Protopapas Z, et al: Impact of filmless radiology on the frequency of clinician consultations with radiologists. *AJR* 173:1169–1172, 1999
12. Reiner BI, Siegel EL, Siddiqui K: Evolution of the digital revolution: a radiologist perspective. *J Digit Imaging* 16:324–330, 2003

13. Sniezek JA: An examination of group process in judgmental forecasting. *Int J Forecast* 5:171–178, 1992
14. Sniezek JA, Henry RA: Accuracy and confidence in group judgment. *Organ Behav Hum Decis Process* 43:1–28, 1989
15. Renooij S, Witteman C: Talking probabilities: communicating probabilistic information with words and numbers. *Int J Approx Reason* 22:169–194, 1999
16. Brenner M, Branscomb HH: The psychological stress evaluation, technical limitations affecting lie detection. *Polygraph* 8:127–132, 1979
17. Brockway BF, Plummer DB, Lowe BM: The effects of two types of nursing reassurance upon patient vocal stress levels as measured by a new tool, the PSE. *Nurs Res* 25:440–446, 1976
18. Van der Car DH, Greaner J, Hibler N, et al: A description and analysis of the question and validity of the psychological stress evaluator. *J Forensic Sci* 25:174–188, 1980
19. Williams CE, Stevens KN: Emotions and speech: some acoustic correlates. *J Acoustic Soc Am* 52:1238–1250, 1972
20. Streeter LA, Macdonald NH, Apple W, et al: Acoustic and perceptual indicators of emotional stress. *J Acoust Soc Am* 73:1354–1360, 1983
21. Rostolland D: Acoustic features of shouted voice, part 1. *Acustica* 50:118–125, 1982
22. Lieberman P, Michaels S: Some aspects of fundamental frequency and envelope amplitude as related to the emotional content of speech. *J Acoust Soc Am* 7:922–927, 1962
23. Ben-Shakhar G, Elaad E: The validity of psychophysiological detection of information with the guilty knowledge test: a meta-analytic review. *J Exp Psychol* 9:261–269, 2003
24. Buller DB, Burgoon JK: Interpersonal deception theory. *Commun Theory* 6:203–242, 1996
25. DePaulo BM: Nonverbal behavior and self-presentation. *Psychol Bull* 111:203–243, 1992
26. Honts CR, Raskin DC, Kircher JC: Mental and physical countermeasures reduce the accuracy of polygraph tests. *J Appl Psychol* 79:252–25, 1994
27. Reiner BI: Medical imaging data reconciliation. Part 3: reconciliation of historical and current radiology report data. *J Am Coll Radiol* 11:768–771, 2011
28. Reiner BI: Medical imaging data reconciliation. Part 4: reconciliation of radiology report and clinical outcome data. *J Am Coll Radiol* 8(12):858–862, 2011