Effect of Human Versus Computerized Voice Recognition Transcription on Billing Level in an Orthopaedic Surgery Practice

B Dezfuli, M Chilvers

Citation

Abstract

Computer based speech recognition transcription software (CBT) use has increased in prevalence in the last decade. However, its effects on level of billing in an orthopaedic practice have not been published. During a one-year period, patients were seen by the author at one of two Orthopaedic Surgery Clinics at the University of Arizona. One clinic utilized human transcriptionists (HT) and another CBT. A total of 1,758 notes were generated, 900 completed by HT and 858 by CBT. Of all HT notes, 70 (8%) were billed as level 4, whereas only 27 (3%) notes generated by CBT billed as level 4 (p< 0.01). When analyzing only new patient reports, 29 (9%) HT reports billed as level 4, whereas only 11 (3%) notes generated by CBT billed as level 4. CBT does not result in a higher level of billing. Caution must be used with implementation of this new technology on the basis of cost savings.

INTRODUCTION

Computerized based speech recognition transcription software (CBT) has been increasingly implemented in medical practices over the last decade.[1] Benefits of the shift from human transcriptionists (HT) may include decreased costs associated with creating transcribed documents and less transcription delays.[2] Drawbacks of the new technology include excessive training time, increased time to dictate individual reports, and increased errors.[3]

Specific medical specialties that have implemented CBT within the literature with reasonable success include pathology and radiology.[1,2] Little research has been done on the implementation of CBT on specialties with greater patient interactions. The concern with CBT is that the increased requirements to review and edit reports negatively impact interaction times with patients. Pathologists and radiologists are somewhat removed from the live patient interaction.

The purpose of this study was to evaluate the impact of CBT on the level of billing in a standard orthopaedic practice. We hypothesize the use of CBT will result in a lower average level of billing than reports generated by HT, which will serve as control.

METHODS

During a one-year period, patients were seen by one of the authors (M.C.) at one of two Orthopaedic Surgery Clinics at the authors' institution. One clinic utilized HT while the other clinic utilized CBT (Dragon NaturallySpeaking, Nuance Communications, Burlington, MA) to generate reports. Reports were then analyzed for level of billing using evaluation and management codes 99201 – 99205 and 99211 – 99215. The 9920- prefix denotes new patient visits. The 9921- prefix denotes return patient visits. The final digit 1-5 denotes level of billing of the reports. Comparisons between groups were performed with Fisher's Exact Test. Differences were considered significant if the P value was less than 0.05. Finally, a multinomial logistic regression was used to explore the effects of whether the patient was new with transcription type on level of billing.

RESULTS

Table 1 and Figure 1 report the results of the study. There were a total of 1,759 outpatient visits during the study period. Of these, 659 (37%) were new patient visits and 1099 (63%) were return visits. At the clinic where HT generated reports, 900 (51%) visits occurred. At the clinic where reports were generated through CBT, 858 (49%) visits occurred. The distribution of visit types (new versus return) between the two clinics was not significantly different (P >

0.05). The distribution of number of visits between the two clinics was not significantly different (P > 0.05).

In either transcription method, no reports generated a level of billing 1 or 5. Of 1,759 reports, 155 (9%) were level 2, 1506 (86%) were level 3, and 97 (6%) were level 4. Of 900 HT reports, 103 (11%) were level 2, 727 (81%) were level 3, and 70 (8%) were level 4. Of 858 CBT reports, 52 (6%) were level 2, 779 (91%) were level 3, and 27 (3%) were level 4.

When analyzing the 659 new patient reports, 9 (1%) were level 2, 610 (93%) were level 3, and 40 (6%) were level 4. Of 325 new patient reports generated through HT, 5 (1%) were level 2, 291 (90%) were level 3, and 29 (9%) were level 4. Of 334 new patient reports generated through CBT, 4 (1%) were level 2, 319 (96%) were level 3, and 11 (3%) were level 4. The distribution of level of billing between the two transcription methods was significantly different (P < 0.01). A significantly higher percentage of new patient reports were billed as level 4 using HT (9% vs 3%, P < 0.05).

When analyzing the 1099 return patient reports, 146 (13%) were level 2, 896 (82%) were level 3, and 57 (5%) were level 4. Of 575 return patient reports generated through HT, 98 (17%) were level 2, 436 (76%) were level 3, and 41 (7%) were level 4. Of 524 return patient reports generated through CBT, 48 (9%) were level 2, 460 (88%) were level 3, and 16 (3%) were level 4. The distribution of level of billing between the two transcription methods was significantly different (P < 0.001). A significantly higher percentage of return patient reports were billed as level 3 using CBT (88% vs 77%, P < 0.05).

A multinomial logistic regression was used to explore the effects of whether the patient was new with transcription type on level of billing. In these analyses, level 2 is considered the reference level. New patients were 11.0 times more likely to be billed at level 3 (p < 0.001) and 11.4 times more likely to be billed at level 4 (p < 0.001) compared with returning patients. After adjusting for whether the patient is new or returning, CBT reports were 2.1 times more likely to be billed at level 3 (P < 0.001) but are 0.8 times as likely to be billed at level 4 (P = 0.318) compared to HT.

Figure 1

Table 1. Billing Level as a Function of Clinic Type

	Both Clinics		Human Transcriptionist Clinic			Computer Based Voice Recognition Software Clinic		
Total Patient Visits (New & Return)	#	%	#	%	P	#	%	P
Billed Level 2	155	8.82%	103	11.44%	< 0.05	52	6.06%	< 0.05
Billed Level 3	1506	85.67%	727	80.78%	< 0.05	779	90.79%	< 0.05
Billed Level 4	97	5.52%	70	7.78%	< 0.05	27	3.15%	< 0.05
Total	1758		900			858		
New Patient Visits	#	%	#	%	P	#	%	P
Billed Level 2	9	1.37%	5	1.54%	NS	4	1.20%	NS
Billed Level 3	610	92.56%	291	89.54%	< 0.05	319	95.51%	< 0.05
Billed Level 4	40	6.07%	29	8.92%	< 0.05	11	3.29%	< 0.05
Total	659		325			334		
Return Patient Visits	#	%	#	%	P	#	%	P
Billed Level 2	146	13.28%	98	17.04%	< 0.05	48	9.16%	< 0.05
Billed Level 3	896	81.53%	436	75.83%	< 0.05	460	87.79%	< 0.05
Billed Level 4	57	5.19%	41	7.13%	< 0.05	16	3.05%	< 0.05
Total	1099		575			524		

Figure 2

Figure 1. Level of Billing as a Function of Transcription Method Used.
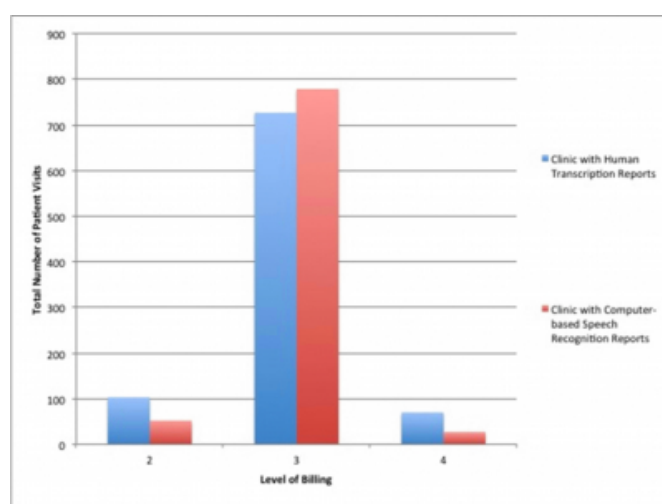
DISCUSSION

The study above indicated that CBT might not be an acceptable alternative to HT for producing outpatient notes with respect to billing. The majority of reports were billed as level 3. Compared with CBT, HT were associated with a lower percentage of level 3 and higher percentage of level 2 and level 4 reports produced. With subgroup analysis, the higher percentage of level 2 reports was attributable to return patient visits. New patient reports, which bill higher than return patient reports, were billed higher with HT than CBT.

Prior studies have shown conflicting results on CBT efficiency. Some studies show CBT prolonging dictation and

proofreading times,[4] while others state improvement[5] Nevertheless, physician surveys used to evaluate experience with CBT reveal high dissatisfaction rates.[6]

Limitations of this study were based on study design. First, this was a retrospective study and although the patient populations between the clinics were expected to be similar, it was not evaluated. Therefore, it is conceivable that the differences between groups were a result of different patient populations.

This study does not show a superiority of such new technology with respect to level of billing. Thus the clinician must balance to cost and benefit of this new technology. This study showed that using HT is associated with billing more level 4 reports. Coupled with prior research showing dictations times with CBT prolonged with higher error rate than HT, it may not be of greater financial benefit or efficiency for orthopaedic practices to switch to CBT.

ACKNOWLEDGEMENTS

References

1. White KS. Speech recognition implementation in radiology. Pediatr Radiol. 2005;35:841-846.
2. Kang HP, Sirintrapun SJ, Nestler RJ, Parwani AV. Experience with voice recognition in surgical pathology at a large academic multi institutional center. Am J Clin Pathol. 2010 Jan;133(1):156-9.
3. McGurk S, Brauer K, Macfarlane TV, Duncan KA The effect of voice recognition software on comparative error rates in radiology reports. Br J Radiol. 2008 Oct;81(970):767-70.
4. Al-Aynati, M.M., Chorneyko, K.A., 2003. Comparison of voice-automated transcription and human transcription in generating pathology reports. Archives of Pathology & Laboratory Medicine, 127(6), 721-725.
5. Borowitz, S.M., 2001. Computer-based speech recognition as an alternative to medical transcription. Journal of the American Medical Informatics Association, 8(1), 101-102.
6. Alapetite, A., Andersen, H.B., and Hertzum, M. 2009. Acceptance of speech recognition by physicians: A survey of expectations, experiences, and social influence. International Journal of Human-Computer Studies, 67, 1, 36-49.

Author Information

Bobby Dezfuli, MD

Department of Orthopaedic Surgery, University of Arizona Medical Center

Margaret Chilvers, MD

Department of Orthopaedic Surgery, University of Arizona Medical Center