Use of Voice Recognition Software in an Outpatient Pediatric Specialty Practice

Robert M. Issenman, MD, FRCPC, and Iqbal H. Jaffer

ABSTRACT. Background. Voice recognition software (VRS), with specialized medical vocabulary, is being promoted to enhance physician efficiency, decrease costs, and improve patient safety. This study reports the experience of a pediatric subspecialist (pediatric gastroenterology) physician with the use of Dragon Naturally Speaking (version 6; ScanSoft Inc, Peabody, MA), incorporated for use with a proprietary electronic medical record, in a large university medical center ambulatory care service.

Methods. After 2 hours of group orientation and 2 hours of individual VRS instruction, the physician trained the software for 1 month (30 letters) during a hospital slowdown. Set-up, dictation, and correction times for the physician and medical transcriptionist were recorded for these training sessions, as well as for 42 subsequently dictated letters. Figures were extrapolated to the yearly clinic volume for the physician, to estimate costs (physician: $110 per hour; transcriptionist: $11 per hour, US dollars).

Results. The use of VRS required an additional 200% of physician dictation and correction time (9 minutes vs 3 minutes), compared with the use of electronic signatures for letters typed by an experienced transcriptionist and imported into the electronic medical record. When the cost of the license agreement and the costs of physician and transcriptionist time were included, the use of the software cost 100% more, for the amount of dictation performed annually by the physician.

Conclusions. VRS is an intriguing technology. It holds the possibility of streamlining medical practice. However, the learning curve and accuracy of the tested version of the software limit broad physician acceptance at this time. Pediatrics 2004;114:e290–e293. URL: http://www.pediatrics.org/cgi/content/full/114/3/e290; voice recognition software, transcription, pediatric.

ABBREVIATION. VRS, voice recognition software.

Technologic innovation often brings the promise of increased ease of use and efficiency. This is followed by a phase in which the innovation must be incorporated into existing processes, followed by normalization of use, which allows a true assessment of the cost-benefit relationship.1–3

An epidemic of sudden acute respiratory syndrome forced a partial hospital closure in April 2003. The sudden decrease in the number of clinic patients treated afforded the university hospital-based physicians the time required to conduct a pilot trial of Dragon Naturally Speaking, a proprietary software system using voice recognition software (VRS). VRS purports to transcribe speech to written text with high accuracy and efficiency. Hospital physicians agreed to test the incorporation of Dragon Naturally Speaking (version 6) dictation into the conventional health record-updating process used in the clinic.

VRS is marketed as a time-saving approach that can increase performance for weaker typists and decrease the time spent in front of a computer. The software must be trained initially to the user’s voice. Then the user can speak in a normal voice and tone; the computer recognizes the speech and displays it as text. In the past few years, VRS has continually been upgraded to provide greater speed and accuracy. A review of the literature indicated that VRS systems have been successfully used to transcribe reports by radiologists. Many studies reported a high level of satisfaction with the systems, with the stipulation that the software was sufficiently enriched with specialized medical vocabulary.

METHODS

Two pediatric staff physicians and 2 radiologists at McMaster Children’s Hospital elected to be included in the trial of Dragon Naturally Speaking (ScanSoft Inc, Peabody, MA), which was concurrently introduced into the Departments of Radiology and Internal Medicine at Hamilton Health Sciences (Hamilton, Ontario, Canada). Each physician was given a 2-hour didactic lecture on the features of the program, after an opportunity to review the instruction manual provided by the software developer. Subsequently, the physicians were given a 2-hour personalized training session on the use of the headset microphone and Dragon Naturally Speaking software and on software incorporation into the Meditech hospital information system (Meditech, Waltham, MA). This provided an opportunity to train the physician’s office personal computer (Pentium 4, 256-MB, personal computer; Dell Corp, North York, Ontario, Canada) in the phraseology and information of the particular physician. Dragon Naturally Speaking incorporated physician dictation into Microsoft Word in Office 2000 on a Windows 98 (2nd ed) platform (Microsoft Corp, Redmond, WA).

Dictation appears as a Word document, and the physician corrects any errors and trains the computer to recognize characteristic and recurring phrases. The physician e-mails the corrected transcript, with Microsoft Outlook, to the office administrative assistant, who pastes the transcribed letter into the Meditech electronic medical record. The letter is immediately accessible on personal computers throughout the hospital, and the paper copy is posted to the referring physician.

In this trial, the process of using Dragon Naturally Speaking was compared with the conventional dictation process. In the
TABLE 1. Dictation, Correction, and Total Time per Letter With Dragon Naturally Speaking Versus the Control Process During the First Month of Training

<table>
<thead>
<tr>
<th></th>
<th>VRS</th>
<th>Correction</th>
<th>Total/Letter</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set-Up</td>
<td>1 min</td>
<td>5 min</td>
<td>8 min</td>
<td>14 min</td>
</tr>
<tr>
<td>Control</td>
<td>3 min</td>
<td>17 s</td>
<td>3.3 min</td>
<td></td>
</tr>
</tbody>
</table>

TABLE 2. Dictation, Correction, and Total Time per Letter With Dragon Naturally Speaking Versus the Control Process After Training

<table>
<thead>
<tr>
<th></th>
<th>VRS</th>
<th>Correction</th>
<th>Total/Letter</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set-Up</td>
<td>1 min</td>
<td>3 min</td>
<td>5.6 min</td>
<td>9.6 min</td>
</tr>
<tr>
<td>Control</td>
<td>3 min</td>
<td>17 s</td>
<td>3.3 min</td>
<td></td>
</tr>
</tbody>
</table>

control process, the physician dictates a letter with a handheld, portable, dictating unit, using Micro-format tapes. An experienced medical transcriptionist transcribes the dictation as a Microsoft Word document, which the office assistant pastes into the electronic medical record. The physician then accesses the letter on an office computer, corrects any errors, and authorizes the letter with an “electronic signature.” The paper copy is sent to the referring physician, and the letter is available in the electronic medical record.

In this study, we compared the time required to dictate and correct letters typed by a transcriptionist with the physician’s time using Dragon Naturally Speaking with a medical vocabulary (an option not available on standard versions). Only 1 of the 4 physicians elected to continue the trial, because of frustration regarding adequately training the software. The trial involved a single physician (R.M.I.), who timed the various steps in both processes with the built-in computer clock, rounding to the nearest minute. Corrections were timed by a research assistant with a stopwatch and were recorded in seconds. The hospital slowdown in April 2003 allowed sufficient time for a VRS trial. During that period, 30 letters were dictated, to familiarize the software with the particular vocabulary of the pediatric specialty (gastroenterology). The trial began at the end of 1 month of orientation and training. Results were entered into a Microsoft Excel spreadsheet. Means and SDs were calculated, and Student’s t test and χ² test were used for comparative statistical analysis.

RESULTS

On average, the physician performs 600 new consultations and 1200 repeat visits per year in the outpatient clinic. Letters for approximately one-half of the consultations and follow-up visits (total: 900) are dictated by fellows, residents, and nurse practitioners, using the central hospital transcription pool. All clinicians use a generic template adapted to the requirements of a dictation letter describing a follow-up visit. The average length of each letter was 225 words, as determined with the word count feature included in Microsoft Word. Mistakes were counted if the transcriptionist or software mistyped a word. Errors made by the dictating physician and other corrections were not counted. When performing corrections, the physician types at 27 words per minute, with a 90% accuracy rate. Results are indicated in Tables 1 through 5 and Figs 1 and 2. Student’s t tests determined that the results were statistically significant with respect to time and financial costs incurred, to a level of P ≤ .001.

TABLE 3. Accuracy Rate per Dictated Letter

<table>
<thead>
<tr>
<th>Word Count</th>
<th>Errors</th>
<th>Accuracy, %</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>VRS</td>
<td>225</td>
<td>19.6</td>
<td>90.8</td>
</tr>
<tr>
<td>Control</td>
<td>259</td>
<td>0.07</td>
<td>99.97</td>
</tr>
</tbody>
</table>

DISCUSSION

Physicians have been notoriously resistant to the adoption of new technology. This has been attributed to innate conservatism of the medical profession. Alternately, physicians have been said to mirror normal human populations in adapting to change. The “visionaries” search out new technologies. The “early adopters,” who represent ~16%, closely watch these individuals. Acceptance by these local opinion leaders leads to adoption by the “early majority.” The “late majority” gradually accepts the technology, leaving only the “resistors,” who reject novelty for its own sake.

The use of computers has gradually been accepted in medical practice. However, most physicians have embraced only those technologies that facilitate their work, such as “read-only” laboratory results, in contrast to the resistance to “physician order entry,” which, although safer and more economical for the patient and the institution, imposes an efficiency cost on the physician.

There is little literature on the use of VRS in medical care. Most successful implementations have been performed in radiology or emergency services. The primary reason for this observation seems to be the fact that, within these 2 medical specialties, there are several highly repetitive phrases that can be well understood by the computer. This is not the case in pediatric subspecialty practices. Previous studies espoused VRS because of the obvious attraction of using VRS to replace illegible and potentially dangerous handwritten notes in medical records. Systems can be made to work with clinical care templates, which order data collection and ensure consistency, accuracy, and completeness, despite a variety of caregivers. These templates also provide clinical data, which can be analyzed with a “natural language processor,” to extract useable data from dictated text.

The use of the VRS brought 1 advantage. Transcription by the physician was immediately correct.
and available once completed. This decreased the turnaround time from ~1 week to <1 day, an advantage that would not be evident in a system with already established, efficient turnaround times.

The use of VRS was 66% less efficient in total time. With inclusion of the capital cost of Dragon Naturally Speaking licensing, VRS cost twice as much as conventional transcription, on an annual basis, when the cost of physician time was included (calculated as $110 per hour). The time needed for correction might gradually decrease with time, but it is unlikely that average physicians (early majority and late majority) would be willing to invest the training time required. Commercial software producers promote the use of a blended system, in which a skilled transcriptionist corrects the dictation initially transcribed by the VRS, as a practical alternative. This could reduce physician correction time, provided the transcriptionist could decipher the meanings of some of the irrational expressions concocted by the software.

One of the major limitations of this study was the fact that the data analysis was based on the experience of a single physician. Although 4 physicians were recruited to participate, they became increasingly frustrated with the inability of the software to produce the desired results. Although all of the physicians were open to the use of VRS in their practice, 3 of the 4 physicians subsequently left the study because of inherent time constraints and difficulties in improving the accuracy of the software.

CONCLUSIONS

VRS has been continuously improved in the past few years. It currently remains impractical, depending on the relative value placed on physician time and transcriptionist time. In certain circumstances, VRS offers rapid turnaround, with instant availability of typed legible transcription. It is best suited to practices (procedures) in which highly repetitive phrases reoccur. In some instances, routing the VRS transcription to the transcriptionist for correction may reduce the inefficiency. There is a high standard for accuracy in medical environments, which may

TABLE 5. Cost Comparison

<table>
<thead>
<tr>
<th></th>
<th>Purchase</th>
<th>Transcriptionist Cost/y</th>
<th>Physician Cost/y</th>
<th>Secretary Cost/y</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>VRS</td>
<td>$400</td>
<td>$14 190</td>
<td>$700</td>
<td>$15 290</td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>$1320</td>
<td>$4950</td>
<td>$700</td>
<td>$6970</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 1. Time involved in transcription of individual letters with VRS, compared with the time needed by the transcriptionist.

Fig. 2. Accuracy of VRS, compared with the accuracy of the transcriptionist.
not be suited to some of the VRS-created irrational syntax. At this time, the software does not seem to have reached the threshold for general adoption.

ACKNOWLEDGMENTS

We acknowledge the continuing help and support of Dale Anderson, Angelo Zingaro, and Mark Farrow of the Hamilton Health Sciences Information Technology Department.

REFERENCES

Use of Voice Recognition Software in an Outpatient Pediatric Specialty Practice
Robert M. Issenman and Iqbal H. Jaffer
Pediatrics 2004;114:e290
DOI: 10.1542/peds.2003-0724-L

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://pediatrics.aappublications.org/content/114/3/e290