Running head: VOICE RECOGNITION TECHNOLOGY: IMPROVING MEDICAL RECORDS

Leveraging Technology:

Using Voice Recognition to Improve Medical Records Production

at Walter Reed Army Medical Center

William L. Novakoski

Baylor University

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
AUG 1999	**Final**	**Jul 1998 - Jul 1999**

4. TITLE AND SUBTITLE
Leveraging Technology: Using Voice Recognition to Improve Medical Records Production at Walter Reed Army Medical Center

5a. CONTRACT NUMBER
5b. GRANT NUMBER
5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
LTC William L. Novakoski

5d. PROJECT NUMBER
5e. TASK NUMBER
5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Walter Reed Army Medical Center 6825 16th St. Washington, DC 20307

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
US Army Medical Department Center and School Bldg 2841 MCCS-HRA (US Army-Baylor Program in HCA) 3151 Scott Road, Suite 1412 Fort Sam Houston, TX 78234-6135

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
9-99

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT
Medical records documentation is burdensome for health care providers in terms of both the time and costs involved in their production. Recent advances in voice recognition technology have made it an alternative to transcription services. This pre-implementation study, conducted within the Department of Pathology and Area Laboratory Services at Walter Reed Army Medical Center, sought to determine if voice recognition technology could be leveraged to improve the production of its anatomic pathology reports. Work process analyses, a transcription services satisfaction survey, and a financial analysis were performed to determine if the time and costs to produce these reports could be reduced. The work process analyses showed the voice recognition system could substantially simplify the process of producing these reportsfrom twelve to four steps. The satisfaction survey showed the pathologists were dissatisfied with the current transcription services. The foreign-born pathologists with accents were particularly dissatisfied with the accuracy, and younger pathologists were dissatisfied with the timeliness of producing the transcribed reports. Use of the voice recognition system could result in a cost savings of $520,000 over five years by eliminating the need for six medical records transcriptionist positions. These results indicate voice recognition could be used to reduce the time and costs involved in the production of the pathology reports, however, the results need to be confirmed following the implementation of the voice recognition system.

15. SUBJECT TERMS
Voice Recognition Technology, Medical Records, Pathology, Transcrition Services

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**UU**	**52**	

Acknowledgments

The author would like to acknowledge and thank COL Renata Greenspan, Chief, and COL James Hawkins, Laboratory Manager, Department of Pathology and Area Laboratory Services, Walter Reed Army Medical Center, and their staff, for completing the satisfaction surveys and providing information on the workflow process. The author would also like to thank LTC Barclay Butler, Department of Information Management, and LTC Jay Carlson, Department of Obstetrics and Gynecology, for their instructions in the capabilities of this information technology. The author thanks MAJ David Brant and Ms. Alice Echard, Headquarters, for their continued encouragement throughout the year, especially when, after many months of work, it looked as if the project would not be completed. Lastly, the author thanks Nancy, Joseph, and Katie, his family, who are his greatest source of happiness.

Abstract

Medical records documentation is burdensome for health care providers in terms of both the time and costs involved in their production. Recent advances in voice recognition technology have made it an alternative to transcription services. This pre-implementation study, conducted within the Department of Pathology and Area Laboratory Services at Walter Reed Army Medical Center, sought to determine if voice recognition technology could be leveraged to improve the production of its anatomic pathology reports. Work process analyses, a transcription services satisfaction survey, and a financial analysis were performed to determine if the time and costs to produce these reports could be reduced. The work process analyses showed the voice recognition system could substantially simplify the process of producing these reports from twelve to four steps. The satisfaction survey showed the pathologists were dissatisfied with the current transcription services. The foreign-born pathologists with accents were particularly dissatisfied with the accuracy, and younger pathologists were dissatisfied with the timeliness of producing the transcribed reports. Use of the voice recognition system could result in a cost savings of $520,000 over five years by eliminating the need for six medical records transcriptionist positions. These results indicate voice recognition could be used to reduce the time and costs involved in the production of the pathology reports, however, the results need to be confirmed following the implementation of the voice recognition system in the pathology department.

Table of Contents

List of Tables

List of Figures

Introduction

Medical records documentation has long been a source of frustration for health care providers. Physicians, in particular, often resent the time consumed in producing medical records, believing this burdensome task does little to improve patient care (Hershey, McAloon & Bertram, 1989). Beyond the clinical need of medical documentation for good follow-up care of the patient, the providers must be increasingly comprehensive and specific in their documentation. Today, medical documentation must be sufficient for medico-legal reasons, to meet Health Care Financing Administration (HCFA) coding mandates, to comply with the Joint Commission on Accreditation of Healthcare Organizations (JCAHO) requirements, and to satisfy managed care organizations' utilization management criteria. For many military providers, the lack of adequate administrative support staff to assist in the completion of medical records exacerbates their frustration.

Voice recognition technology has been under development for the past three decades. Physicians saw the advent of voice recognition technology as a means to reduce the burden of producing the required medical documentation. Physicians and health care administrators alike recognized the potential to reduce both the time and costs involved in producing clinical records. Early voice recognition system testing disappointed clinicians because of problems with accuracy, limited vocabularies, and slowness (Bergeron, 1996). Thus, the benefits of reducing medical records production times and costs were unachievable. During the past two years, performance breakthroughs in affordable, large vocabulary, speaker independent voice recognition systems have re-ignited the medical community's interest (Bergeron, 1997; Wolinsky, 1998).

Conditions that prompted the study

Clinicians and administrators in Walter Reed Army Medical Center's (WRAMC) Department of Pathology and Area Laboratory Services (DPALS) have been interested in voice recognition technology as an alternative to inefficient, costly medical records transcription services for a number of years. In fact, in 1995, the department purchased a voice recognition system from Kurzweil Applied Intelligence. Following installation, however, the system was soon abandoned by the pathologists, who found it user-unfriendly for a number of reasons. The system was too slow-it could process the pathologist's dictation only at the rate of 90 words per minute instead of a normal speaking rate of approximately 120 to 150 words per minute. Discrete speech patterns (i.e., the requirement that speakers pause before each word) made its use unnatural and difficult. Lastly, the accuracy of the system in transcribing the correct word was not satisfactory to the pathologists. Because of their dissatisfaction, no cost savings were realized in pathology from this early purchase of the voice recognition system (personal communication with COL Hawkins, October 1, 1998).

A surgeon, newly reassigned from William Beaumont Army Medical Center (WBAMC), began to champion the adoption of voice recognition technology at WRAMC during the summer of 1998. While at WBAMC, this surgeon had investigated the use of voice recognition technology for use in dictating data into the ambulatory data system, thus alleviating physicians of time-burdening, handwriting efforts. The surgeon briefed WRAMC's Deputy Commander for Clinical Services (DCCS) on the capabilities of this technology, and arranged for demonstrations of the Kurzweil/Lernout & Hauspie product to the Deputy Commander for Administration (DCA), the Chief, Information Management, the Chief of the DPALS, the administrative resident, and other key clinical and administrative staff. Following what proved to be an

impressive demonstration of the voice recognition system by the vendor, the executive leadership of WRAMC including the Commander, the DCCS and the DCA, along with the leadership in DPALS decided to purchase the updated Kurzweil product for use within its Anatomic Pathology Service. In addition, the WRAMC leadership agreed that the administrative resident would conduct a study to determine the pathologists' acceptance of the system and any cost reductions achieved using the voice recognition technology. They believed that if voice recognition technology improved medical records production within the pathology department, it might also be used to improve medical records production in other inpatient and outpatient services within WRAMC. In addition, the Commanding General of the North Atlantic Regional Medical Command (NARMC), who had also been briefed on this technology, envisioned potential savings at other military treatment facilities (MTF) throughout his command.

Statement of the Problem

DPALS is looking for ways to gain efficiencies in producing its anatomical pathology reports. Currently, the reports take 2 to 5 days to produce-from receipt of the tissue specimen to completion of the final report. Each month, approximately 25% of the final reports fail to be completed within the 2-day College of American Pathologists' (CAP) standard (personal communication with Dr. Horton, May 28, 1999). The department's annual transcription costs for producing these reports from the pathologists' dictations exceed $210,000. Both the clinical staff and the department administrators desire to improve upon the production of these reports. Clinical pathologists are specifically interested in producing accurate reports in a more timely manner to improve services for the attending physicians taking care of the patients and to comply with CAP standards. The department administrators would like to improve the work processes

and reduce the cost of producing these reports (personal communication with COL Hawkins, October 1, 1998).

Voice recognition technology had been proposed by the newly-assigned surgeon as a leveraging technology solution which could satisfy the interests of both the clinical pathologists and the department administrators. A voice recognition system could reduce the time and cost of generating these reports by automating the transcription steps in the process. For these timeliness and cost improvements to be realized, the voice recognition technology system must gain the acceptance of the individual pathologists, the end-users of the technology. If this technology does not meet the needs of the pathologists, they may once again abandon its use as did the WRAMC pathologists who attempted to use this technology in the past.

Literature Review

The production and management of clinical reports represents a considerable expense of both time and dollars in today's health care setting (Rosenthal, Bos, Sokolowski, Mayo, Quigley, Powell & Teel, 1997). For these reasons, during the 1980's, clinicians began to test various applications of the newly emerging voice recognition technology in clinical settings. Voice recognition systems were tested for anesthesia record keeping, radiology reporting, dental charting and gastrointestinal endoscopy reporting (Leeming, Proter, Jackson, Bleich & Simon, 1981; Sarnat, 1983; Matumoto, Iinuma, Tateno, Ikehira, Yamasaki, Fukuhisa, Tsunemoto, Shishido, Kubo, & Inamura, 1987; Feldman & Stevens, 1990; Cass, 1992). These early voice recognition system tests struggled to overcome problems with accuracy, limited vocabularies, and slowness. Error rates of one or two words for every ten words dictated, vocabularies of only a few thousand words, and dictation rates of only 30 to 40 words per minute were hindrances to the adoption of this technology. The biggest hindrance of all, however, has been that the voice

recognition systems required the user to speak in the unnatural discrete speech manner (Bergeron, 1997).

Scientists and academicians debated the state of voice recognition technology at a 1993 National Academy of Science symposium entitled, "Human-Machine Communication by Voice." At the symposium, Makhoul and Schwartz (1995) proclaimed a paradigm shift in voice recognition technology due to the availability of "high-accuracy, speaker-independent, continuous speech recognition for large vocabularies possible in real time, on off-the-shelf workstations, without the aid of special hardware" (p. 9956). Levinson, the moderator of the symposium, argued a dissenting opinion on the state of voice recognition technology. He asserted that the real paradigm shift would occur when computers understood speech, not these "incremental advancements in speech transcription" (Levinson, 1995). The presentations at this meeting tended to be either very scientific or philosophic and contributed little to the immediate application of this technology in business or clinical settings.

Lernout & Hauspie (L&H) (which purchased Kurzweil Applied Intelligence in 1996), Dragon Systems, Inc., and International Business Machines (IBM) are the three leading manufacturers of voice recognition technology products (21st Century Eloquence, 1998). In the late 1980's these manufacturers began to develop healthcare specific voice recognition products. Kurweil commissioned two studies to assess the financial and strategic benefits of using their voice recognition system in health care settings during the early 1990's. One study, conducted in an emergency department setting providing care for 51,000 patients annually, found combined financial benefits for the hospital and physician group to be over $645,000 during a five year period when compared with their handwritten records (DMR Group, Inc., 1992). The second study compared the use of Kurweil's product against traditional dictation services in the

emergency departments of two hospitals. This study found that while the process of producing the emergency room record was simplified, the physicians' dictation times were unchanged. In addition, the time to achieve break-even on the full investment was six to seven months in a 20,000-visit emergency department. Lastly, the study showed a five year cost savings of $472,000 compared to the dictation services (Little, 1994). Of course, these manufacturer commissioned studies were performed for marketing purposes and the results must be skeptically reviewed.

Threet & Farques (1999) published their pilot project results of an economic evaluation of voice recognition technology use in the Family Practice Clinic of the Naval Hospital Roosevelt Roads, Puerto Rico. This study evaluated DragonDictate, a commercial, off-the-shelf discrete speech product manufactured by Dragon Systems, Inc., for an 9-month period during 1996 and 1997. Their provider perception surveys indicated that the use of voice recognition technology would result in more thorough medical records documentation, eliminate the costs of hiring medical transcriptionists, and reduce the overload that providers feel with their health information management system. They also queried patients with a patient satisfaction survey to assess the effect of voice recognition technology on the patient-provider encounter. From these surveys, they concluded dictating during the encounter did not adversely effect the patient-provider relationship when compared with handwritten notetaking. Accuracy, words per minute, and input time were generally found to improve with increased provider usage of the system. They also noted four disadvantages; potential noise interference, user speech pattern variability, concerns for patient's privacy violations, and words per minute limitations of this discrete speech product (Threet & Farques, 1999; Threet, 1997).

Today, past accuracy problems can largely be overcome by training the voice recognition system. This training entails reading a few paragraphs of specified text into the microphone followed by a period of processing time in which the computer creates a profile specific to the user's voice. Following this training, users can expect accuracy rates to 95% or higher. The discrete speech limitation was also remedied by the industry during 1997. Today, continuous speech allows the user to dictate text to their computers by speaking naturally at a rate between 100 to 140 words per minute (McCune, 1998).

Massachusetts General Hospital's Department of Radiology began to alpha test a continuous speech voice recognition system in November, 1995. Three years later, 55% of the department was using voice recognition to produce 700 to 800 reports per workday. In their experience, the predominant benefit was the decreased report turnaround time of the voice recognition produced reports versus the transcribed reports; 4 versus 2.4 days respectively. The department achieved a $350,000 cost savings within the first two years of voice recognition use largely by reducing the number of full-time equivalent (FTE) transcriptionist positions from 22 to 7. In addition, Massachusetts General Hospital found that the accuracy of the radiology reports was improved with the voice recognition system. Their greatest challenge was to convince the radiologists to use the technology when its utilization did not immediately benefit the end user. They overcame this challenge by providing technical support and user-friendly, convenient training (Mehta, Dreyer & Thrall, 1999).

L&H demonstrated its pathology-specific product, Kurzweil Clinical Reporter Version 2.0 for Pathology, to the American Society of Clinical Pathologists meeting in September, 1997. According to its news release ("Lernout and Hauspie demonstrates" 1997), this pathology product "integrates L&H's large vocabulary continuous speech recognition technology, a

pathology knowledge base, developed in conjunction with practicing pathologists, an automatic report writer, and complete integration services." The news release went on to claim that health care institutions could expect the following benefits from this product; savings of 70 - 100% in transcription costs, immediate availability of final reports, common protocols to ensure the consistency of information at the grossing station, significant data input efficiency in the microscopic description and final diagnosis, and point-of-care, structured recording of pertinent patient data for outcome studies (p. 1). Pathology departments now using this L&H pathology product include the Hospital of the University of Pennsylvania, William Beaumont Hospital in Royal Oak, Michigan, and Saint Luke's Health Network in Bethlehem, Pennsylvania ("Hospital of the University" 1997; Lernout & Hauspie helps drive" 1999).

Purpose

The purpose of this study was to determine if voice recognition technology could be leveraged to improve the process and thereby reduce the time and costs of producing anatomic pathology reports within DPALS. Although cost-benefit analyses are usually performed as a pre-adoptive evaluation of a new technology, this study was accomplished concurrently with the implementation of the voice recognition system.

Methods and Procedures

Setting

WRAMC is the largest of twelve MTFs located within the 21-state NARMC (NARMC, 1998). Today WRAMC is a large tertiary care teaching hospital in northwest Washington, DC, serving a beneficiary population of more than 440,000 in the National Capitol Area. In addition to its local beneficiaries, WRAMC serves as a referral center for military hospitals and deployed

U.S. military forces throughout the world. Its current mission statement is to:

> provide quality, comprehensive health care that is cost competitive and accessible; serve as a national resource for specialty care and medical issues unique in DoD and other federal agencies; maintain individual and collective readiness in support of the DoD Health Care System; and provide research, education and training in support of the DoD Health Care System (Walter Reed Army Medical Center, 1998).

The Department of Pathology and Area Laboratory Services offers a wide range of clinical and anatomic pathology services within WRAMC, and serves as a reference and consulting laboratory for military medical treatments facilities throughout the northeastern United States. Another important part of the DPALS mission is to train graduate physicians in the specialty of pathology. To accomplish this, DPALS operates a four-year pathology residency in cooperation with the National Naval Medical Center in Bethesda, Maryland. Currently, DPALS employs nine full-time staff pathologists and twenty resident pathologists. Together, this staff interprets approximately 20,000 surgical specimens and performs 50 to 100 autopsies per year (Walter Reed Army Medical Center, 1999).

Product Description

Kurweil Clinical Reporter™ for Pathology is a voice recognition software system specifically developed to produce quality pathology reports. It features an active vocabulary, to include medical and pathology specific words, and is expandable to 64,000 words. The system contains a pathology knowledge base which prompts the user to follow established practice guidelines, minimizing the risks of inaccurate or incomplete reporting. Its ability to recognize continuous and natural speech is designed to allow the pathologist to dictate structured or detailed free text notes quickly and easily. The system adjusts to the individual user's voice during an initial system training period called "enrollment." During this enrollment, the user reads prescribed text to the system for 30 to 60 minutes. United States-born users can expect

90 - 95% accuracy following this initial enrollment process. Foreign-born users with accents should expect less accuracy even after the enrollment process. All users will experience better accuracy with repeated use as the system continuously adjusts to their voices (Lernout & Hauspie, 1998).

Kurweil Clinical Reporter™ for Pathology can run on either the Windows® 95 or Windows NT® 4.0 operating systems. The minimum hardware requirements include Pentium® 200 MHz processor with MMX™, 128 MB RAM, CD-ROM drive, 4 GB hard drive, AWE 64 sound board, SVGA video adapter, and a headset or handset microphone. The system components, quantities, unit prices, and total price actually purchased by DPALS are shown in Table 1.

Analysis

Work process analyses of the transcription services and the voice recognition processes were documented through interviews with two staff pathologists and the department administrator and by the administrative resident's direct observation of work practices. The work processes for both the transcription services and the voice recognition process were considered to have started with the accession of the labeled specimen into the pathology department. The final step in the work processes with both methods of producing the reports was when the final report was signed and ready for return to the requesting physician. For the work process documentation to be considered valid, the two pathologists and the department administrator had to agree on each step of the transcription services and the voice recognition processes. The two work processes were then compared to determine the steps that had either been eliminated or simplified after the voice recognition system had gone into use.

The pathologists' level of satisfaction with the transcription services was assessed with a survey (Appendix A). The survey was distributed to the pathologists by a staff pathologist who had been selected as the clinical champion to oversee the implementation of the voice recognition system. The pathologists completed this questionnaire prior to their training on the voice recognition technology. This survey will serve as a baseline for comparison of the pathologists' satisfaction with the voice recognition system. Following full implementation of the voice recognition system, the pathologists will be asked to complete a second satisfaction survey (Appendix B). Both surveys asked the pathologists to rate their level of satisfaction for five items on a modified 5-point Likert scale. The first three items asked about accuracy, timeliness and ease of use for each of the two methods. The content validity of these items was verified from a literature review which revealed these three items to be relevant to physicians' satisfaction with this technology in numerous other studies. The fourth item asked about the effect on Graduate Medical Education (GME), Quality Improvement (QI) and Risk Management (RM). This item was added at the request of the clinical champion. The final item asked the pathologists to give an overall rating of satisfaction with each method. The construct validity of the survey was verified by prescreening with the administrator and clinical champion (Cooper & Emory, 1995). Both surveys also asked the pathologists for additional comments and a number of demographic questions. The second survey asked about the length of the enrollment process and the number of times the voice recognition system had been used subsequent to enrollment. Descriptive statistics were calculated using SPSS® statistical software.

Confidentiality was provided for the survey respondents in accordance with good research practices. The purpose of having the identifying information on the surveys was to match the individual pathologist's level of satisfaction with the dictated reports and reports

generated with the voice recognition system. A statement of confidentiality was included on each survey.

The financial analysis was accomplished by comparing data on the costs of producing the anatomic pathology reports for fiscal year 1999 (FY99) using its general schedule salaried medical transcriptionists (MRT) and the expected costs of producing the reports using the voice recognition technology. Table 2 shows the FY99 salaries of the eight currently employed MRTs. The cost of the dictation equipment was considered a "sunk cost" for this analysis and was excluded in the cost computation of the transcriptionist's produced reports. The costs of using the voice recognition technology included the salaries and benefits of the two transcriptionists to be retained for other duties, as well as the expenditures for the hardware, software, and service contract. In addition, the financial analysis was accomplished by forecasting the costs for each system five years into the future. The difference between the costs of producing the reports was the cost savings or revenue gain with using the voice recognition system. An average annual return-on-investment (ROI) was calculated for a five year period. In addition, profitability analysis was performed by calculating both the net present value (NPV) and the internal rate of return (IRR) of the revenue gains for the five year period. The last measure of the financial analysis was time break-even or the payback time expected to recover the investment in the voice recognition system.

Results

The work process documentation for the transcription services and the voice recognition technology is shown in Figures 1 and 2, respectively. The work processes both began with the accession of the anatomic specimen into the laboratory. On the first day of the process the

pathologist performed a gross examination which consisted of visually inspecting, weighing and measuring the tissue. The gross examination was completed only after the typed report had been reviewed, corrected, and signed-out by the pathologist. This six step gross examination process was expected to take one day. If, however, the report required corrections, the transcription and review steps were repeated adding a second or third day to the process. Using the voice recognition system the dictation, review, corrections, and sign-out were combined into a single step. Thus, the gross examination was reduced from six or more steps to a two step process and was completed in a single day.

On the first day, following the gross examination, the specimen was frozen or chemically fixed and cut into sections for viewing on a microscopic slide. On the second day the pathologist performed the microscopic examination of the tissue. Using the transcription services the microscopic examination of the tissue was expected to be another six step, one day process. As with the gross examination, corrections required a recycling of the transcribed reports adding steps and days to the process. With the voice recognition system the microscopic examination was also reduced to a two step process by combining the dictation, review, corrections, and sign-out steps.

Both the transcription services and voice recognition work processes were designed as two day processes to allow the time required for the slide preparation. The voice recognition process, however, eliminated the recycling of reports to the transcriptionists, thereby ensuring additional days are not required to produce the final report.

Thirteen (45%) of 29 staff and resident pathologists assigned to DPALS completed the transcription services survey. Several of the residents were on training rotations external to WRAMC and were unavailable. All thirteen pathologists surveyed were Army officers in the

ranks of captain to colonel (Table 3). The pathologists ranged in age from 28 to 59 years with 62% (8) being less than forty years of age. Sixty-two percent (8) of these pathologists were male and 38% (5) were residents in their pathology specialty training. Nine pathologists (69%) had worked at WRAMC less than five years. Three (23%) of the pathologists were foreign-born and spoke with heavy accents.

Transcription services satisfaction scores are shown on Table 4. The values shown are the mean scores of the pathologists' Likert Scale scores. A Likert Scale score of one indicated the pathologist was very dissatisfied with the transcription services and a score of five indicated the pathologist was very satisfied with that item. Overall, the pathologists were slightly dissatisfied with the transcription services ($\underline{M}$ = 2.8) and slightly dissatisfied with the accuracy of the transcribed reports ($\underline{M}$ = 2.7). The scores showed the pathologists to be somewhat satisfied with the timeliness of the reports ($\underline{M}$ = 3.5) and the ease of using the transcription services ($\underline{M}$ = 3.5). The pathologists felt neither satisfied nor dissatisfied with the effect of the transcription services on GME, QI, and RM ($\underline{M}$= 3.0).

Univariate analysis was performed to determine if the pathologists' satisfaction with the transcription services varied by key demographic variables (Table 5). The overall satisfaction of the pathologists who had accents was significantly lower than that of the pathologists without accents ($\underline{p}$ = 0.041). The same accented pathologists' satisfaction with the accuracy of the transcribed reports was lower than the non-accented pathologists' satisfaction ($\underline{M}$ = 2.0 versus M = 2.9 respectively) but did not achieve statistical significance ($\underline{p}$ = 0.095). The younger pathologists were significantly less satisfied with the timeliness of the transcribed reports ($\underline{M}$ = 2.9 versus $\underline{M}$ = 4.6, $\underline{p}$ = 0.006).

The transcription services satisfaction survey included blank spaces for the pathologists to add written comments. Three favorable comments and three unfavorable comments were noted on the survey forms (Table 6). One comment contained both a favorable and an unfavorable element. Three pathologists commented unfavorably on the accuracy while two pathologists praised the praised the transcriptionists' accuracy. One pathologist wrote favorably about the turn-around time of the reports.

The financial analysis was performed to compare the costs of the transcription-produced against the voice recognition-produced reports (Table 7). This analysis showed that the NPV of the cost reduction would exceed $520,000 in five years. This cost savings would be achieved by eliminating the six lowest MRT positions. The total salary savings would be $147,200 in FY99 and rise at 4% each subsequent year. The ROI and IRR were calculated, as 119% and 146%, respectively. This analysis also showed the expected time to pay back the cost of the voice recognition system would be ten months.

Discussion

WRAMC has struggled over the past several years to continue to perform its mission in the face of declining budgets allocated from the U. S. Army Medical Command (MEDCOM). In fact, the cumulative budget shortfall for the past four years was $16 million. In response to this shortfall, WRAMC has made numerous efforts to improve its operating efficiency. Its efforts have included the development of an integrated health care system with the MTFs at Forts Belvoir and Meade, the consolidation of inpatient wards, the delivery of more services in lower cost outpatient settings such as ambulatory procedure and short stay units, and the reduction of

its military and civilian workforce. The budget and personnel cuts have been felt down to the department and service levels (personal communication with COL Heckert, September 8, 1998).

WRAMC has sought opportunities to use advanced technology to improve the efficiency of delivering health care services at a lower cost. WRAMC's technology improvement efforts are in keeping with a goal of the Army Medicine Strategic Vision of leveraging technology to "capitalize on information technology, exploit emerging technology and develop business-driven technology solutions" (Army Medicine, 1998). Thus, it is with great interest that WRAMC's leadership has supported and followed the voice recognition system implementation within DPALS.

Based on this study, the work process of producing the anatomic pathology reports is expected to be substantially simplified using the voice recognition system. What has been a twelve step or more process will be reduced to a four step process. The time savings in producing the reports will occur because the dictations will not be sent to the MRTs for typing and resent to the MRTs for correcting errors.

The survey indicates that overall the pathologists are currently dissatisfied with the transcription services ($\underline{M}$ = 2.8). This low satisfaction score indicates that there is considerable room for improvement in the production of the reports. The pathologists were most dissatisfied with the accuracy of the reports ($\underline{M}$ = 2.7) and most satisfied with the timeliness ($\underline{M}$ = 3.5) and the ease ($\underline{M}$ = 3.5) of using the transcription services. The accuracy was of particular concern to the pathologists with accents compared to the pathologists without accents ($\underline{M}$ = 2.0 versus $\underline{M}$ = 2.9, respectively). Correctly capturing the intended words will be an important test for the voice recognition system to pass. Because the voice recognition system will learn each pathologist's

voice during the enrollment process and continue to learn their voice with each subsequent use, the system is expected to perform with a high degree of accuracy.

The younger pathologists were significantly less satisfied with the timeliness of the transcribed reports than the older pathologists ($\underline{M}$ = 2.9 versus $\underline{M}$ = 4.6 respectively). In fact, the DPALS April QI meeting minutes noted that only 74% of the reports were completed within two days and 95% were completed within three days. The simplified voice recognition work process is expected to improve this completion time by achieving near 100% completion within the two-day CAP standard.

The financial analysis for using the voice recognition system appears very favorable-a cost savings of $520,000 over a five year period. The cost savings will be achieved by eliminating six MRT positions. While this represents a cost savings in producing these reports, it does not necessarily equate to a surplus within the DPALS budget. Any cost savings on the part of DPALS by using the voice recognition is likely to result in lower departmental budgets in future years. Because of civilian personnel management regulations, these MRTs must be allowed to compete for other vacant positions within WRAMC for which they might qualify. Therefore, the overall WRAMC personnel numbers may or may not change with the elimination of these six MRT positions.

It is important to note that the incentives for using voice recognition to improve the production of these pathology reports are different at the organizational and departmental levels. Any cost savings will benefit the organization, WRAMC. The department's incentive is to simplify its work process, reduce the report throughput time to meet QI standards, and improve the satisfaction of the pathologists.

An important limitation of this study was that it was conducted prior to the full implementation of the voice recognition system. The work process simplification and cost savings must be verified after the system is fully implemented. In addition, because of the importance of the pathologists' satisfaction, a follow-up voice recognition satisfaction survey should be performed. To this end, a voice recognition system satisfaction survey is included in this study. Despite this major limitation, the results of this study are consistent with other recently published studies which document that voice recognition can improve the medical records documentation (Bergeron, 1996; Mehta, Dreyer, and Thrall, 1999; Threet & Farques, 1999).

Conclusions and Recommendations

This study concludes that voice recognition technology can be leveraged to reduce the time and costs of producing pathology reports. These improvements will result from eliminating the errors and costs associated with transcription services. Because this study was performed before the implementation of the voice recognition system, a follow-up study is recommended. The follow-up study should verify the work process and the financial analysis and survey the pathologists for their satisfaction with the voice recognition system. It is further recommended that the results be briefed to the WRAMC and NARMC leadership. WRAMC and NARMC should consider implementation of voice recognition technology within their other departments and MTFs.

References

Army Medicine. (1998). *Strategic Vision 1998.* Available online: www.armymedicine.army. mil/armymed/default.htm.

Bergeron, B. P. (1996). Voice recognition: An enabling technology for modern health care? Proceedings of the American Medical Informatics Association-Annual Fall Symposium, 802-806.

Bergeron, B. P. (1997). Usable voice-recognition technology; It's finally arrived. Postgraduate Medicine, 102(5), 38-44.

Cass, O. W. (1992). Automated speech technology for gastrointestinal endoscopy reporting and image recording. Proceeding of the American Medical Informatics Association-Annual Symposium on Computer Applications in Medical Care, 1991, 968-969.

Cooper, D. R., & Emory, C. W. (1995). Business Research Methods. Chicago: Irwin.

DMR Group, Inc./Strategic Technologies. (1992). A professional benefit study: VoiceEM® within a hospital emergency department. Wellesley, MA: Author.

Feldman, C. A., & Stevens, D. (1990). Pilot study on the feasibility of a computerized speech recognition charting system. Community Dentistry and Oral Epidemiology, 18(4), 213-215.

Hershey, C., McAloon, M., & Bertram, D. (1989). The new medical practice environment: Internists view of the future. Archives of Internal Medicine, 149, 1745-1749.

Hospital of the University of Pennsylvania (HUP) selects L&H's Kurzweil clinical reporter software for creating pathology reports by voice. (1997, July 14). Lernout & Hauspie News Release, p. 1. Available online: www.lhsl.com/news/releases/19970714-NISTGrant.asp.

Leeming, B. W., Proter, D., Jackson, J. D., Bleich, H. L., & Simon, M. (1981). Computerized radiologic reporting with voice data-entry. Radiology, 138, 585-588.

Lernout & Hauspie demonstrates first in a series of continuous speech voice-enabled reporting solutions for the medical market. (1997, September 22). Lernout & Hauspie News Release. Available online: www.lhsl.com/news/releases/19970922-ContinuousPath.asp.

Lernout & Haupie (1998). L&H's Kurweil Clinical Reporter for Pathology (Marketing brochure). Burlington, MA: Author.

Lernout & Hauspie helps drive use of speech technology in healthcare; announces new contracts. (1999, April 13). Lernout & Hauspie News Release. Available online:www.lhsl.com/ news/ releases/19970413-medmomentum.asp.

Levinson, S. E., (1995). Speech recognition technology: A critique. Proceedings of the National Academy of Sciences, 92, 9953-9955.

Little, A. D., Inc. (1994). VoiceEM® and traditional transcription: A study of the costs and benefits. Cambridge, MA: Author.

Makhoul, J., & Schwartz, R. (1995). State of the art in continuous speech recognition. Proceedings of the National Academy of Sciences, 92, 9956-9963.

Matumoto, T., Iinuma, T. A., Tateno, Y., Ikehira, H., Yamasaki, T., Fukuhisa, K., Tsunemoto, H., Shishido, F., Kubo, Y., & Inamura, K. (1987). Automatic radiologic reporting system using speech recognition. Medicine and Progressive Technology, 12, 243-257.

McCune, J. C. (1998). Telling is believing: Let your speech replace your keyboard with today's voice-recognition technology. Management Review, 87(5), 19-21.

Mehta, A., Dreyer, K. J., & Thrall, J. H. (1999). Voice recognition technology. Imaging Economics, 12(1), 56-58.

North Atlantic Regional Medical Command. (1998). *North Atlantic Regional Medical Command.* Available online: www.narmc.amedd.army.mil/.

Rosenthal, D. F., Bos, J. M., Sokolowski, R. A., Mayo, J. B., Quigley, K. A., Powell, R. A., & Teel, M. M. (1997). A voice-enabled, structured medical reporting system. Journal of the American Informatics Association, 6(4), 436-441.

Sarnat, A. J. (1983). Computerized speech recognition for anesthesia recordingkeeping. Medical Instrumentation, 17, 25-27.

Threet, E. (1997). Economic analysis of voice recognition for the clinicians' desktop at the Naval Hospital Roosevelt Roads. Unpublished master's thesis, Naval Postgraduate School., Monterey, CA.

Threet, E., & Farques, M. P. (1999). Economic analysis of voice recognition for the clinicians' desktop at the Naval Hospital Roosevelt Roads. Military Medicine, 164, 119-126.

Walter Reed Army Medical Center. (1998). *Mission, vision and values.* Available online: www.wramc.amedd.army.mil/welcome/misvisval.htm.

Walter Reed Army Medical Center. (1999). *Welcome to the department of pathology.* Available online: www.wramc.amedd.army.mil/departments/pathology/welc.htm.

Wolinsky, H. (1998). Is the time right for speech recognition software? American College of Physicians-ASIM Observer, 8-9.

21st Century Eloquence. (1998). *Products: Voice recognition engines.* Available online: http://voicerecognition.com/1998/products/.

Table 1

System Purchased by Department of Pathology and Area Laboratory Services

Model number	Description	Qty	Unit price	Price
CR-SWUG-01	Clinical Reporter for Pathology update, ten user software license includes: large vocabulary Surgical Pathology reporting system with a knowledge base for gross descriptions and diagnoses.	10	$1,500	$15,000
CR-PATH-R-12	Clinical Reporter for Pathology upgrade, 12 User Resident software license includes: large vocabulary Surgical Pathology reporting system with a knowledge base for gross descriptions and diagnoses.	1	N/C	N/C
WS-1210	Pentium II 400 MMX, 128 Megs of 10ns SDRAM 17” SVGA Color Monitor, Dual IDE Ports/FFD & I/O Controller Serial Ports & One Parallel Port, I/O Controller Card, 5.1 GB Ultra ATA Hard Drive, (3.5”0 1.44M High Density Floppy Drive, 24X Speed Internal CD-ROM Drive,	16	$3,640	$58,240

Table 1 Continued

Model number	Description	Qty	Unit price	Price
	Creative Labs Sound Blaster 64 Sound Card, 3Com 10/100 Ehternet Card, PS/2 Keyboard, PS/2 Mouse, McAfee VirusScan, PC Anywhere, Microsoft Windows.			
WS-1005	Windows NT Server: Pentium Pro 200, dual capability motherboard, 1 processor, 64 MB RAM, 2Gb mirrored external or internal sled SCSI drives, 500Mb ide hard drive, 12x ide CD-ROM, 33.6k full duplex fax/modem, 3Com 10/100 Ehternet Card, 15” color monitor, SVGA graphics, UPS, Seagate tape backup system, Windows NT 4.0 software and 5 workstation client license. 90 day warranty.	1	$5,415	$5,415
MI-1006	Shure Headset Microphone and On-off foot pedal with keyboard interface. 90 day warranty.	25	$220	$5,500

Table 1 Continued

Model number	Description	Qty	Unit price	Price
MAINT	Annual Support Contract on 6 additional Pcs and 15 Headsets.	1	$3,017	$3,017
		TOTAL		$87,172

Note. From Quotation #98-09-737J dated 2 September 1998 by E. Bruce Sopko, Regional Sales Manager, Lernout & Hauspie.

Table 2.

General Schedule Medical Records Transcriptionists' Annual Salaries

Grade & Step	Salary[a]
GS 7-10	$35,760
6-8	30,530
5-6	25,908
5-5	25,168
5-4	24,428
5-3	23,688
5-1	22,208
4-10	25,800
Total	$213,490

Note. From Salary Table 1999-DCB. Available online: www.opm.gov.

[a]Salary includes fringe benefits and 7.87% locality pay for General Schedule employees in the Washington-Baltimore area.

Table 3

Demographics of Pathologists Who Completed the Transcription Services Satisfaction Survey

Variable	No.	(%)	Range	Mean
Active Duty	13	(100)	--	--
Rank[a]				
O3	4	(31)	--	--
O4	5	(38)	--	--
O5	1	(8)	--	--
O6	3	(23)	--	--
Position				
Staff	8	(62)	--	--
Resident	5	(38)	--	--
WRAMC Years				
<5 years	9	(69)	1 - 4	4.7
>5 years	4	(31)	7 - 14	9.3
Age				
<40 years	8	(62)	28 - 36	32.3
≥40 years	5	(38)	40 - 59	48.6

Table 3 Continued

Variable	No.	(%)	Range	Mean
Gender				
Male	8	(62)	--	--
Female	5	(38)	--	--
Accent	3	(23)	--	--
	10	(77)	--	--

Note. [a]O3 = Captain, O4 = Major, O5 = Lieutenant Colonel, O6 = Colonel.

Table 4

Pathologists' Satisfaction Scores With Transcription-produced Reports

Question	Mean	Range
How satisfied were you with the accuracy of the transcriptions?	2.7	1- 5
How satisfied were you with the timeliness of completing the report (from time of dictation to time when the report was ready for signature)?	3.5	2 - 5
How satisfied were you with the ease of using the dictation services?	3.5	1 - 5
How satisfied were you with the effect of the dictation services on GME, QI, and RM?	3.0	1 - 5
Overall, how satisfied were you with using the dictation service?	2.8	1 - 5

Table 5

Univariate Analysis: Pathologists' Satisfaction With Transcription-produced Reports

Variable	Mean	Range	p value
Satisfaction Question			
Accuracy?			
Accent	2.0	2	0.095
No accent	2.9	1 - 5	
Age-<40 years	2.6	1 - 5	0.829
Age-≥40 years	2.8	2 - 5	
Male	2.8	1 - 5	0.855
Female	2.6	2 - 5	
Staff	2.8	2 - 5	0.875
Resident	2.6	1 - 5	
WRAMC-<5 years	2.4	1 - 5	0.396
WRAMC-≥5 years	3.3	2 - 5	

Table 5 Continued

Satisfaction Question Variable	Mean	Range	p value
Timeliness?			
Accent	3.3	2 - 4	0.757
No accent	3.6	2 - 5	
Age-<40 years	2.9	2 - 5	0.006
Age-≥40 years	4.6	4 - 5	
Male	3.3	2 - 5	0.333
Female	4.0	2 - 5	
Staff	3.9	2 - 5	0.290
Resident	3.0	2 - 5	
WRAMC-<5 years	3.4	2 - 5	0.834
WRAMC-≥5 years	3.8	2 - 5	

Table 5 Continued

Variable	Mean	Range	p value
Satisfaction Question			
Ease?			
Accent	3.0	2 - 4	0.439
No Accent	3.6	1 - 5	
Age-<40 years	3.5	1 - 5	0.888
Age-≥40 years	3.4	2 - 5	
Male	3.8	1 - 5	0.173
Female	3.0	2 - 4	
Staff	3.6	2 - 5	0.298
Resident	3.2	1 - 5	
WRAMC-<5 years	3.3	1 - 5	0.540
WRAMC-≥5 years	3.8	3 - 5	

Table 5 Continued

Variable	Mean	Range	p value
Satisfaction Question			
Effect on GME, QI, & RM?			
Accent	2.7	2 - 4	0.603
No accent	3.1	1 - 5	
Age-<40 years	2.9	1 - 4	0.648
Age-≥40 years	3.2	2 - 5	
Male	3.1	1 - 4	0.648
Female	2.8	2 - 5	
Staff	3.4	2 - 5	0.106
Resident	2.4	1 - 3	
WRAMC-<5 years	2.7	1 - 4	0.449
WRAMC-≥5 years	3.8	2 - 5	

Table 5 Continued

Satisfaction Question Variable	Mean	Range	p value
Overall?			
Accent	2.0	2	0.042
No accent	3.0	1 - 5	
Age-<40 years	2.6	1 - 5	0.613
Age-≥40 years	3.0	2 - 5	
Male	2.8	1 - 5	0.648
Female	2.8	2 - 5	
Staff	2.9	2 - 5	0.737
Resident	2.6	1 - 5	
WRAMC-<5 years	2.6	1 - 5	0.449
WRAMC->5 years	3.3	2 - 5	

Table 6

Pathologists' Written Comments on Transcription Services Satisfaction Survey

Favorable Comments (N = 3)
"For surgical pathology-these guys do a great job. Sure, there are typos, but it's our job to make sure the report is correct. I was surprised to learn that they don't type autopsies... I'm sure there's a good reason, but it's sure a pain to type them up myself... the templates are wrong."
"I am used to typing my own reports at AFIP, so I consider our typing services a luxury."
"Overall, the accuracy of the dictations has improved dramatically during the past two years. I am somewhat dissatisfied with the length of time between dropping the cases off and when the cases would be on the computer for sign out. To be fair, the amount/volume of dictations per day is at times very high. Autopsy dictation-gave up. Typed them in myself."
Unfavorable Comments (N = 3)
"Let's be honest here-the transcription service at WRAMC is notorious and known Army-wide, across the world since the new building was built, speaking with people who trained here 20 years ago; improving little over the past 6 months, still incompetent."

Table 6 Continued

"Mrs. X is terrible-I have 3-5 typos per report-often glaring-I know I'm supposed to proof-but with such a high error rate some slip through."
"Obvious lack of proofreading on the part of the transcriptionists. I spent far too much time correcting their errors."
Mixed Comment (N = 1)
"Cases turned around rapidly but sometimes important errors in them that take some time to fix."

Table 7

Financial Analysis: Transcription-produced v. Voice Recognition-produced Pathology Reports

Assumptions: Costs & Savings

1. Equipment, software & on-site training	$87,172	Sixteen Kurzweil systems, hardware & software
2. Annual service contract	$3,017	Begins Year 2
3. Incremental revenue recapture	$147,200	Six MRT FTE salaries @ 4% (annual increase)

Year	Cost	Cost Savings	$Net	PV Factor @10%	$PV
	$87,172		($87,172)	1.00	($87,172)
1		$147,200	$147,200	0.9091	$133,820
2	$3,017	$153,088	$150,071	0.8264	$124,019
3	$3,017	$159,212	$156,195	0.7513	$117,349
4	$3,017	$165,580	$162,563	0.7830	$127,289
5	$3,017	$172,203	$169,186	0.6209	$105,048
	$99,240	$797,283	$698,043		$520,350

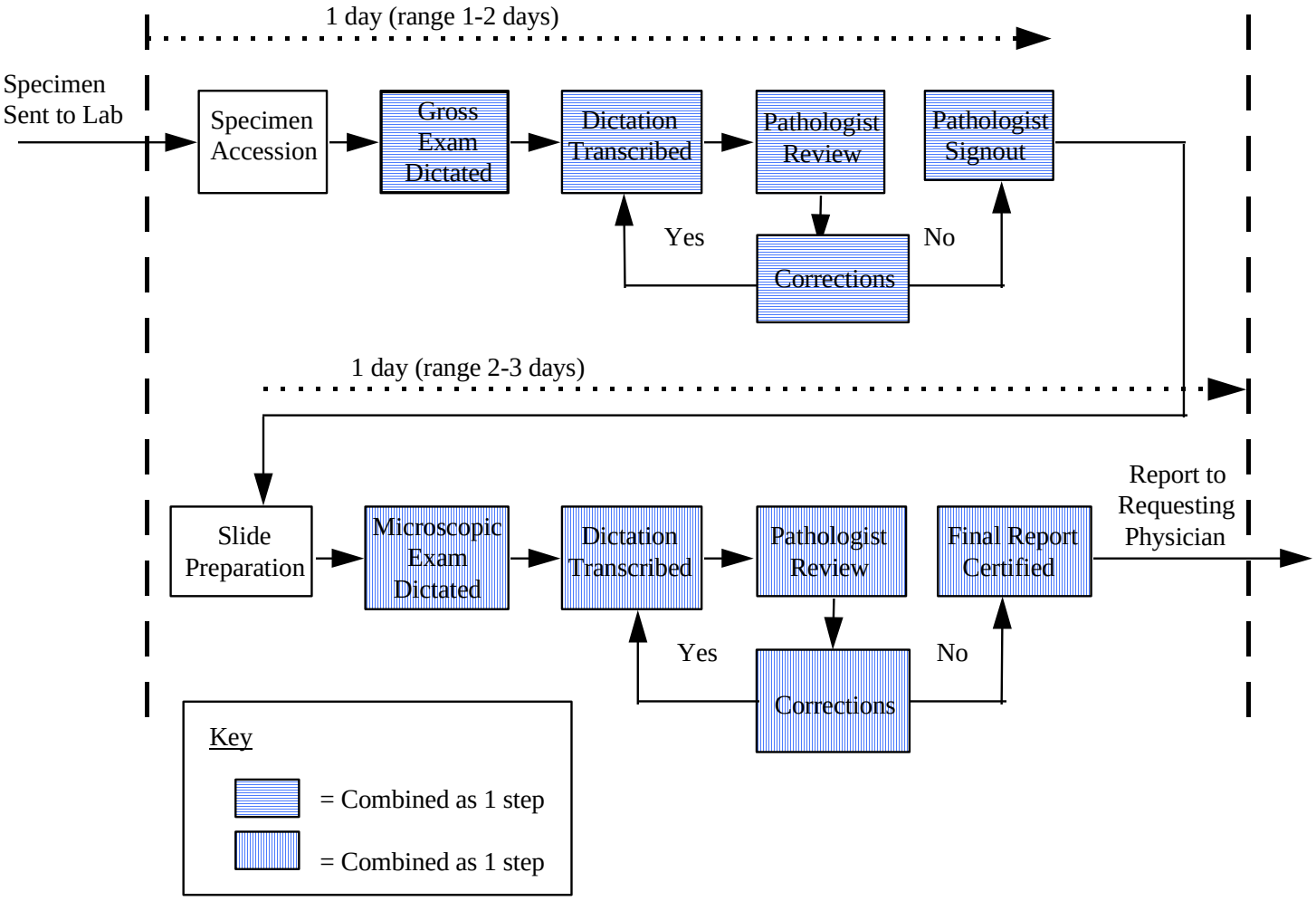
Table 7 Continued

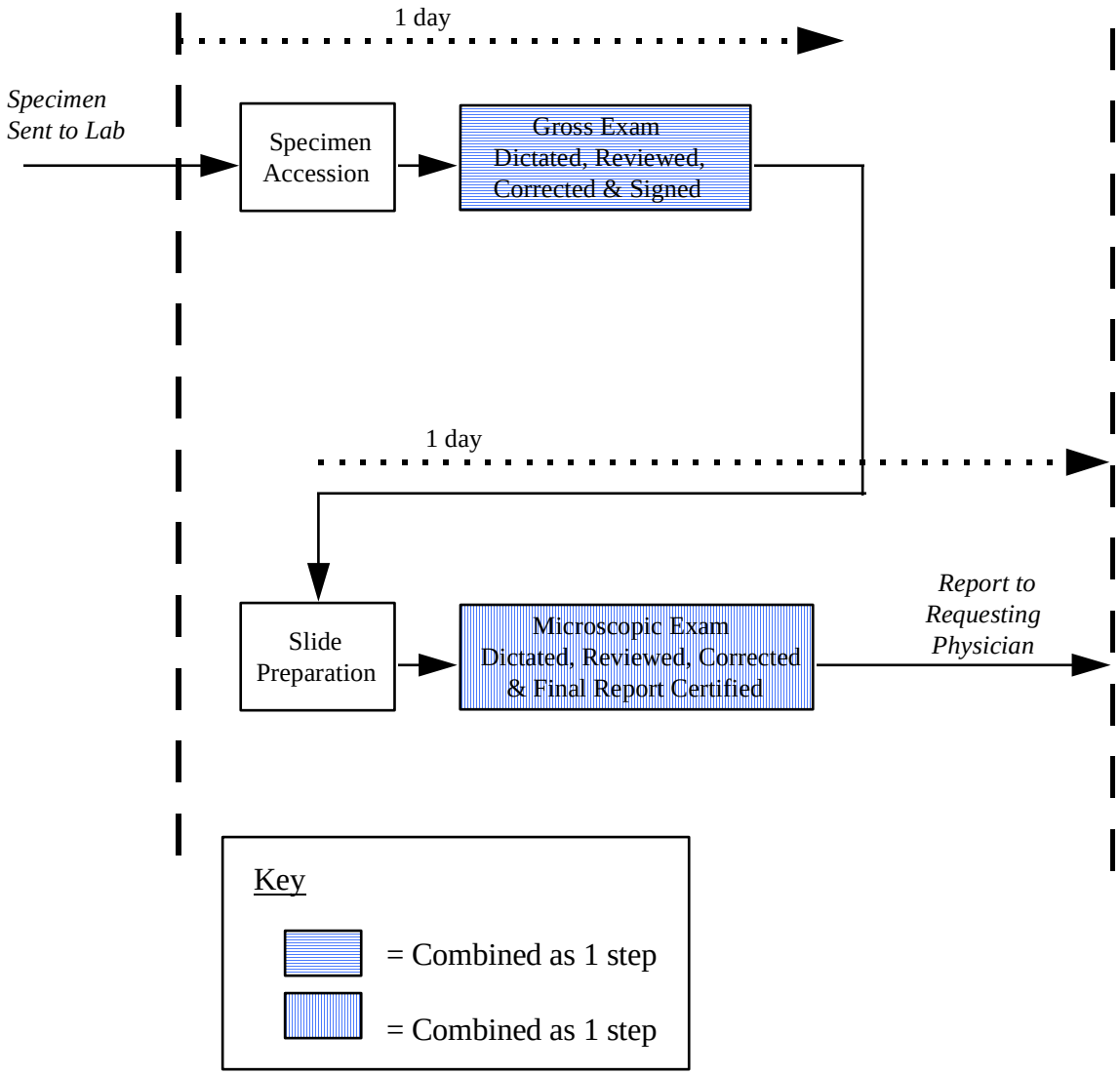
Benefit Summary		
Average Rate of Return	ROI	119%
Internal Rate of Return	IRR	146%
Net Present Value	NPV	$520,350
Payback (months)	P	10.05

Figure Captions

Figure 1. Transcription Services Work Process.

Figure 2. Voice Recognition Technology Work Process.


1 day (range 1-2 days)
Specimen Sent to Lab
Specimen Accession
Gross Exam Dictated
Dictation Transcribed
Pathologist Review
Pathologist Signout
Yes
No
Corrections
1 day (range 2-3 days)
Slide Preparation
Microscopic Exam Dictated
Dictation Transcribed
Pathologist Review
Final Report Certified
Report to Requesting Physician
Yes
No
Corrections
Key
= Combined as 1 step
= Combined as 1 step


1 day
Specimen
Sent to Lab
Specimen
Accession
Gross Exam
Dictated, Reviewed,
Corrected & Signed
1 day
Slide
Preparation
Microscopic Exam
Dictated, Reviewed, Corrected
& Final Report Certified
Report to
Requesting
Physician
Key
= Combined as 1 step
= Combined as 1 step

Appendix A

WRAMC Transcription Services Satisfaction Survey

WRAMC Transcription Services Satisfaction Survey

This questionnaire was designed to determine the level of satisfaction with the dictation services. Your ratings will provide valuable feedback on the capability of the dictation services in meeting your dictation needs. The personal identification information is intended to enable linking of initial evaluations to follow-up evaluations completed at a later date. All personal identification information and demographic information will be kept confidential. Thank you for your participation.

Please rate the following items on the scale provided. Circle the number that corresponds to your answers (for example, Very dissatisfied = 1, Somewhat dissatisfied = 2, No opinion/Neutral = 3, Somewhat satisfied = 4, Very satisfied = 5).

	Very satisfied	Somewhat satisfied	No opinion/ Neutral	Somewhat dissatisfied	Very dissatisfied
How satisfied were you with the accuracy of the transcriptions?	5	4	3	2	1
How satisfied were you with the timeliness of completing the report (from time of dictation to time when the report was ready for signature)?	5	4	3	2	1
How satisfied were you with the ease of using the dictation services?	5	4	3	2	1
How satisfied were you with the effect of the dictation services on GME, QI and RM?	5	4	3	2	1
Overall, how satisfied were you with using the dictation service?	5	4	3	2	1

Please provide any additional comments on the lines below.

_____
_____
_____
_____
_____

Please answer the reverse side of this survey.

WRAMC Transcription Services Satisfaction Survey (continued)

Demographic information

First name: _____Last name: _____ SSN: ___-__-____
Age: __ Gender: M / F Ethic origin: _____ Accent: Y / N

Military: Y / N Rank: O3 O4 O5 O6
Civilian: Y / N Grade: GS12 GS13 GS 14

Number of years at WRAMC: ___
Position in the Department of Pathology: Staff / Resident

Appendix B

WRAMC Voice Recognition Technology Satisfaction Survey

WRAMC Voice Recognition Technology Satisfaction Survey

This questionnaire was designed to determine the level of satisfaction with the Kurzweil voice recognition technology. Your ratings will provide valuable feedback that will be used to assess the utility of this technology in your service and whether this technology should be used in other clinical services. The personal identification information is intended to enable linking of initial evaluations to follow-up evaluations completed at a later date. All personal identification information and demographic information will be kept confidential. Thank you for your participation.

Please rate the following items on the scale provided. Circle the number that corresponds to your answers (for example, Very dissatisfied = 1, Somewhat dissatisfied = 2, No opinion/Neutral = 3, Somewhat satisfied = 4, Very satisfied = 5).

	Very satisfied	Somewhat satisfied	No opinion/ Neutral	Somewhat dissatisfied	Very dissatisfied
How satisfied were you with the accuracy of the transcriptions?	5	4	3	2	1
How satisfied were you with the timeliness of completing the report (from time of dictation to time when the report was ready for signature)?	5	4	3	2	1
How satisfied were you with the ease of using the voice recognition technology?	5	4	3	2	1
How satisfied were you with the effect of the voice recognition technology on GME, QI and RM.	5	4	3	2	1
Overall, how satisfied were you with using this voice recognition technology?	5	4	3	2	1

Please provide any additional comments that you wish to make on the lines below.

_____

_____

_____

Please answer the reverse side of this survey.

WRAMC Voice Recognition Technology Satisfaction Survey (continued)

Demographic information

First name: _____Last name: _____ SSN: ___-__-____
Age: __ Gender: M / F Ethic origin: _____ Accent: Y / N

Military: Y / N Rank: O3 O4 O5 O6
Civilian: Y / N Grade: GS12 GS13 GS 14

Do you have a sore throat, cold, hoarseness or any other condition at this time which alters the tone of your voice?
Y / N
If yes, please describe. _____

Did you go through the Kurzweil enrollment process of reading to the computer before dictating pathology reports into the system?
Y / N
If yes, how long did you read to the computer?
15 min 30 min 45 min 60 min 75 min 90 min

Excluding the enrollment process, how many times have you used this system prior to this session?
1 2 3 4 5 6 >6