Research Paper in Speech Recognition Technology

ورقة بحث عن نظام التعرفة الصوتية

By:

Ayla Zaid Mohammed Harbi

Student ID: 100064

Dissertation submitted in partial fulfilment of the requirements for degree of Master in Information Technology Management

Faculty of Engineering and IT

Dissertation Supervisor

Dr. Sherief Abdallah

March 2014

الجامعة البريطانية في دبي

The British University in Dubai

DISSERTATION RELEASE FORM

Student Name	Student ID	Programme	Date
Alya Zaid Mohammed Harbi	100064	ITM	01 March 2014

Title
Research Paper in Speech Recognition Technology

I warrant that the content of this dissertation is the direct result of my own work and that any use made in it of published or unpublished copyright material falls within the limits permitted by international copyright conventions.

I understand that one copy of my dissertation will be deposited in the University Library for permanent retention.

I hereby agree that the material mentioned above for which I am author and copyright holder may be copied and distributed by The British University in Dubai for the purposes of research, private study or education and that The British University in Dubai may recover from purchasers the costs incurred in such copying and distribution, where appropriate.

I understand that The British University in Dubai may make that copy available in digital format if appropriate.

I understand that I may apply to the University to retain the right to withhold or to restrict access to my dissertation for a period which shall not normally exceed four calendar years from the congregation at which the degree is conferred, the length of the period to be specified in the application, together with the precise reasons for making that application.

Signature

Abstract

Introduction of electronic medical records (EMR) documentations is burdensome for physicians and clinicians in terms of time spent in typing, and adaption of this technology might be hard. Recent advances in speech recognition technology (SRT) had added positive impact to increase the user's adaption toward EMR and to be as an alternative to the transcription services. The aim of this paper is to provide comprehensive details of a research study related to the SRT implementation in Ministry of Health (MOH) and to measure it is benefits in terms of productivity, quality of the reports and patient care, time efficiency, and cost saving, and to further investigate the level of satisfaction among the users. A combination of two study designs were utilized for this study which are cross-sectional study and cohort study and as a tool, self-administrated questionnaire was distributed on a random sample. Results of the study showed that the SRT was beneficial. It increased their productivity by 14%, increased number of patients seen for each user by 10%, quality of care increased by 10%, saving almost 20% of the users time in writing their reports, reducing the time between seeing the patients and validating the reports by 18%, and showed potential cost saving of about 22%. Overall user satisfaction was fairly acceptable as only 16% of the users were unsatisfied. Our results suggested that the benefits of SRT in healthcare field are tangible and that this innovative technology can be considered a tool to increase the adaption of EMR technologies and can add great value on the quality of work and productivity of healthcare providers.

ملخص

يعتبر البعض أن توثيق السجلات الطبية بإستخدام نظم المعلومات الصحية عملية مرهقة و قد يكون من الصعب على مستخدمين السجلات الإلكترونية التكيف مع هذه التقنية. إن التطورات الحديثة التي لحقت نظام التعرفة الصوتية قد أضافت أثر إيجابي لتحسين تأقلم المستخدم مع السجل الطبي الإلكتروني وأيضا كبديل لخدمات النسخ. تهدف هذه الدراسة التي تتعلق بنظام التعرفة الصوتية المطبقة من قبل وزارة الصحة على قياس نسبة رضى المستخدمين و قياس المنافع العائدة من هذه التقنية من حيث الإنتاجية و جودة الرعاية الصحية و إستهلاك الوقت و توفير التكاليف. و عليه تم تطبيق تصميم للدراسة الذي يعتبر مزيج من تصميم الدراسة المقطعية و تصميم دراسة الاتراب و قد تم إستخدام الإستبيان كأداة لهذه الدراسة. أظهرت نتائج الدراسة أن هذه التقنية لها فوائد من كل الجوانب حيث إرتفعت الإنتاجية بنسبة 14% فيما يتعلق بكمية التقارير الصادرة خلال يوم واحد، زيادة عدد المرضى الذين تم معاينتهم بنسبة 10%، احتمالية توفير التكاليف بنسبة 22%، تحسين نوعية الرعاية للمريض بنسبة 10% حيث يتم توفير تقارير مفصلة أكثر عن حالة المريض مما ينتج عن تشخيص أفضل و أدق، تقليل الوقت المستهلك بنسبة 20% لكتابة تقرير واحد، و تقليل الوقت المستهلك بنسبة 18% بين فحص المريض و الإمكانية لكتابة التقرير و التحقق من صحته. كما أثبتت الدراسة أن نسبة رضى المستخدمين مقبولة إلى حد ما حيث ان نسبة الغير راضين فقط 16%. إن هذه النتائج تثبت ان نظام التعرفة الصوتية لها أثر ملموس و إيجابي على نوعية و إنتاجية الرعاية الصحية

Acknowledgment

Table of Contents

Introduction

Information and Communications Technologies have great impact on the ways Governments and businesses interact with each other. It is evident nowadays how pervasive the technology has become especially with all the rapid diffusion of the network and mobile telephony. From this perspective United Arab Emirates (UAE) adapted the idea of e-government strategic framework and to make it official His Highness Sheikh Mohammed Bin Rashid Al Maktoum, Vice President and Prime Minister of the UAE and Ruler of Dubai, announced the post e-Government era to speed up the revolutionary pace toward smart and great utilization of the technology. From this perspective, Ministry of Health (MOH) in UAE has incorporated the advance technology in their strategies and embraced many changes not only to fulfil the UAE overall e-Government strategy, but to enhance the quality of patient care and facilitate their services. Speech Recognition Technology (SRT) was one of these initiatives that MOH adopted. SRT can be described as the process of converting the spoken work into electrical signals as inputs for a computer program, which will allow the user's hand to be free. This can be used in the healthcare fields to facilitate documenting the medical reports and eliminate the need of transcription services.

The introduction of Health Information Systems (HIS) and Electronic Medical Records (EMR) documentations might be considered burdensome for physicians and clinicians in terms of the time consumed to type their documentations. From this perspective, MOH put a lot of investments in implementing SRT with the hope to increase quality of services provided to patient care, increase productivity of the user, reduce the time consumed in documenting medical reports and save the cost.

There are some studies that were conducted which prove the benefits of SRT not only in healthcare fields but in all market segments such as telecommunication industry and warehouses. However, there is no study conducted for SRT in healthcare field in UAE or even in the Gulf Region. Therefore, after the

implementation of SRT in MOH, a study was needed to prove that the investments given for SRT resulted in some benefits. The aim of this paper is to provide comprehensive details of a research study related to the SRT implementation in MOH and to measure its benefits. Hence, the research question for this study is, will SRT provide the benefits that MOH is seeking for?

In order to answer this research question a study was carried out using a combination of two study designs which are cross-sectional study and cohort study. As a study tool, a self-administrated questionnaire was distributed on a random sample of physicians, histopathologists, radiologists and surgeons who are considered to be heavy users of EMR in terms of documentations.

This paper provides literature review about SRT technology and its types, and it discusses the usage of SRT in different market segments and then concentrates in the SRT's impact in healthcare fields, followed by the study details and its methodology, analysis and results.

Literature Review

Background and Definition

Last few years, SRT has emerged into the market to be one of the essential tools for facilitating different kind of services. Different segments of the market witnessed the increasing demand of this technology mainly as to be in line with the demand increase of the customers and the fact that there is a need to boom up their performance to jump ahead of their competitors. Healthcare industry for example forced to adopt these cutting edge innovations and technologies just to be ahead and control their market.

The definition of SRT can be simply described as the process of converting the spoken work into electrical signals as inputs for a computer program, which will allow the user's hand to be free (Baumann, 1993). Though this is being adopted

now a day, it is worth to mention that the fact of using human voice as an input for the computer seemed to be a challenge as computer programs depend on precise inputs and human voice are anything but precise. Each spoken word can be different and can have different meaning in different contexts. From this point of view, several approaches to create SRT have been taken to overcome these difficulties and the degree of the success and accuracy vary depending on its approaches.

Descriptions and Types

The approaches of SRT can be classified into two main classes which are Speaker-Dependent or Speaker-independent. Of those two techniques studies showed that Speaker-Dependent is more accurate and simplest than Speaker-independent, as the accuracy can reach up to 98% while in Speaker-independent vary between 90% and 95% (Baumann, 1993). Speaker-Dependent which is also known as "template Matching" technique is when the software is required to learn the unique characteristics of the user's voice (Joshi, 2010). Hence, new users must first train the tool by speaking some words and phrases to be as inputs to a microphone or the software. The design behind this technique is through utilizing the concept of analog to digital converter, where the analog signal from the microphone is getting digitalized by the converter and stored in a memory. Then a matching process between the input and digitalized voice sample "template" is taking place. From this point of view, we can conclude that the program cannot have a template for each potential user, thus it must be trained with new user's voice inputs to be understandable by the program. The training session can be simply described as taking the users input of a word and computes it into a statistical average of the multiple samples and then storing it in data structure "template".

In the other hand, Speaker-independent which can be referred as "Feature Analysis Technique" depends mainly on the Linear Predictive Coding (LPC). Therefore, instead of working in finding an exact or near exact match of a voice and previous stored template, it processes the user's voice input through LPC and

works in finding characteristics similarities between the actual voice input and the expected inputs. Wide range of speakers can be using this and no need to train the program. However, due to having different accents, different speed of delivering the inputs, different volume and inflections of the voice from person to person, the accuracy of the Speaker-Dependent technique might be less.

Another way to classify SRT is through its ability to handle the types of words such as discrete words, connected words or continuous speech. The easiest to implement, is the SRT systems that can only handle discrete words, where users must pause between each word. This can be used where it is required to give only one command or one responses but it would be not liable for multiple word inputs. SRT with connected word inputs allows the user to speak multiple word phrases, but user must be careful in expressing each word and not to mix the end of a word to the beginning of the next one. Continuous speech SRT, on the other hand, can be described as natural SRT system where words can be run together smoothly without pauses. This type is one of the most difficult types to implement.

Making an additional point, SRT might involve difference technology processes such as the below:

- Digitalization: as mentioned earlier converting the analogue signal into digital representations.

- Signal Processing: which is referred to the separation of background noise from the speech

- Phonetics: which is the linguistics that deals with the variability in human speech

- Phonology: which is referred to the recognition of the individual sound distinctions and the use of that sound to encode useful meaning to the spoken human language

- Lexicology and Syntax: which is the linguistics that deals with the words and its elements, nature, it meaning and the relationship between words and the whole lexicon

- Syntax and Programtics: where syntax tells all about the meaning and programtics filling the explanatory gap between sentence meaning and user's meaning..

Technology, Usage and Applications of SRT

According to B.H. Juang & Lawrence Rabiner (2004) the evaluation of statistical modeling enhance the SRT nowadays to invade widespread applications in tasks that would need a human machine interfaces (see figure 1 which illustrates the evolution of SRT). SRT can be used in many ways such as providing updated travel information through query-based information systems and processing in the telephone networks. As well dictations, system control/navigation, commercial/industrial applications or voice dialing (Joshi, 2010) are other fields that SRT dominating on. In this section, we will navigate the usage of SRT and provide examples of where and how it was implemented.
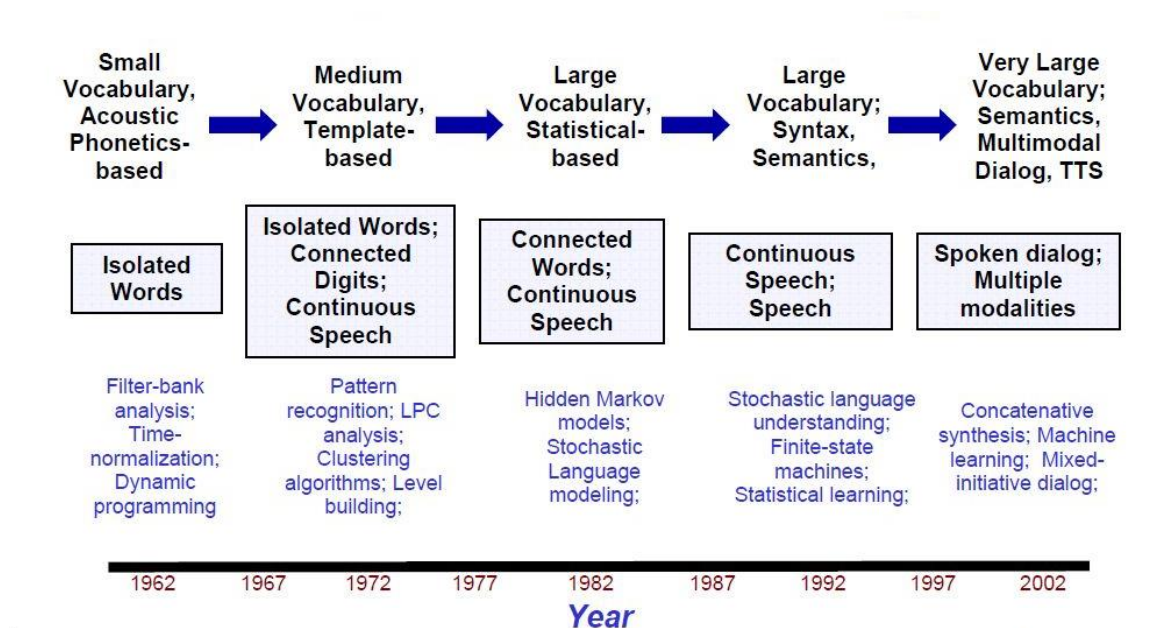

Figure 1. Milestones in speech and multimodal technology research   From (B.H & Rabiner 2004)

The technology of SRT evolved during the last decades. For example the concepts of Linear Predictive Coding (LPC) was approached in the 60s which basically simplify the estimation of the vocal tract response from speech waveforms, however by the next decade the concept of applying fundamental pattern recognition technology to SRT.based on LPC was adopted by Rabiner and Levinson (Rabiner & Levinson, 1979, pp. 336-348), and Itakura (Itakura, 1975, pp. 68-62).The methodology was shifted in the 80s from intuitive template based approach towards more rigorous statistical modeling framework where the idea of utilizing the artificial neural networks took place (B.H & Rabiner, 2004, pp. 1-20; Lowerre & Reddy, 1990,pp. 567-586).). Nowadays, even advanced technology are utilized for developing SRT such as Multi Net ANN (MN-ANN) which uses several parallel and standalone ANNS to learn the views which can as well provide noise robust system without relying on the noise processing, and recent studies showed that MN-ANN provide better performance than Multiple View Single Learner (MVSL) which employ only one ANN to process the entire vocabulary (Shahamiri & Binti Salim, 2014, pp. 199-207).

SRT is proven to be very successful for the usage of call center and customer services. During the 90s, the use of SRT in call center emerged and was adopted by many agencies. For instance, the AT&T Operator line helped a caller to place calls, conduct credit card transactions and arrange payment methods through the use of SRT which reduced the large operating cost of their call center through the reduction of agent positions (Suendermann, 2011, pp. 11-53). Another study confirmed the same result that SRT lower their operational cost which showed that companies using SRT for their incoming customers call have managed to cut off their cost and managed to automate between 40% and 85% of their calls which reduced the number of employees needed for handling these calls. One of the telecom companies actually managed to improve customer services and utilized their knowledge based on their Frequently Asked Questions (FAQs) and one of these telecom company actually managed to routes the caller to one of their 35 automated services modules based on the caller's need (TechTarget, 2009). Because of this valuable fact, many vendors considered investing on the

development of this technology, such as Nuance, Microsoft, IBM, Emerging Technology and Loqueendo.

The industrial market witnessed the evolution of speech technology during the last decades. One of the segments of this market is the warehouse industry. For instance, the DC incorporates have adopted the warehouse voice applications to facilitate the day to day work and to reach a hand free labor. They implemented what is called LUCAS system which is basically SRT. It is worth to mention that warehouse SRT applications would demand far better recognition accuracy than a call center basically because of the following challenges:

- Nonstandard accents

- Variable background noise provide

In contrast to call center application that has a live person as a default if the recognizer could not understand the user, warehouse application must have no fall back and work for all users all the time. Hence, developers of SRT for the warehouse are considering the different processes such as Signal Processing to separate the speech from the background and Phonetics to control variability in human speech. LUCAS system actually evolved during the last 20 years, as their first generation of speaker-dependent often required users to perform a second training of the system as usually users speaks very clearly the first time they train their tool to create their template, so with time users would eventually get very comfortable as they would speed up and their natural speed patterns of the words would get combined which would leads the system to omit them. However, with the evolution of technology, the concept of adoption voice modelling was carried out where the SRT is automatically familiarizes the user's voice template while it is being used, so the template is continuously built up and getting trained as the users works, so the accuracy of SRT actually increases with use (Yurick, 2011, pp. 1-8).

In addition, the usage of SRT has expanded to cover the educational fields. A sample work was developed using SRT to enforce the basic language education

and allow the user pronounce and learn pronunciation of language (from Gazi Teknik)

Moreover, the vision of SRT was as well adopted by the manufacture of smart devices in the 80s. For instance, back in 1988 Apple Computer had announced their vision regarding the SRT for 2011, which was titled "Knowledge Navigator" that defined the concept of SRT (B.H & Rabiner, 2004, pp. 1-20), and just by the end of 90s and early 2000s we witnessed how Automated Speech Recognition comes to the masses for these particular field, for example the usage of Siri by Apple and Google Voice Search App for iPhone (Pinola, 2011). It is becoming possible nowadays to use mobile terminals supported with SRT for many services such as making reservations, paying for holiday accommodation and transporting or controlling the mobile phone and calls while driving a car. The use of speech function call application is very common now with smartphone as mentioned above and recent applications are the "Voice Action" and "Voice it" which are used in android operation systems (Lizuka, et al, 2012, pp. 79-84). Below are some figures that demonstration those applications:
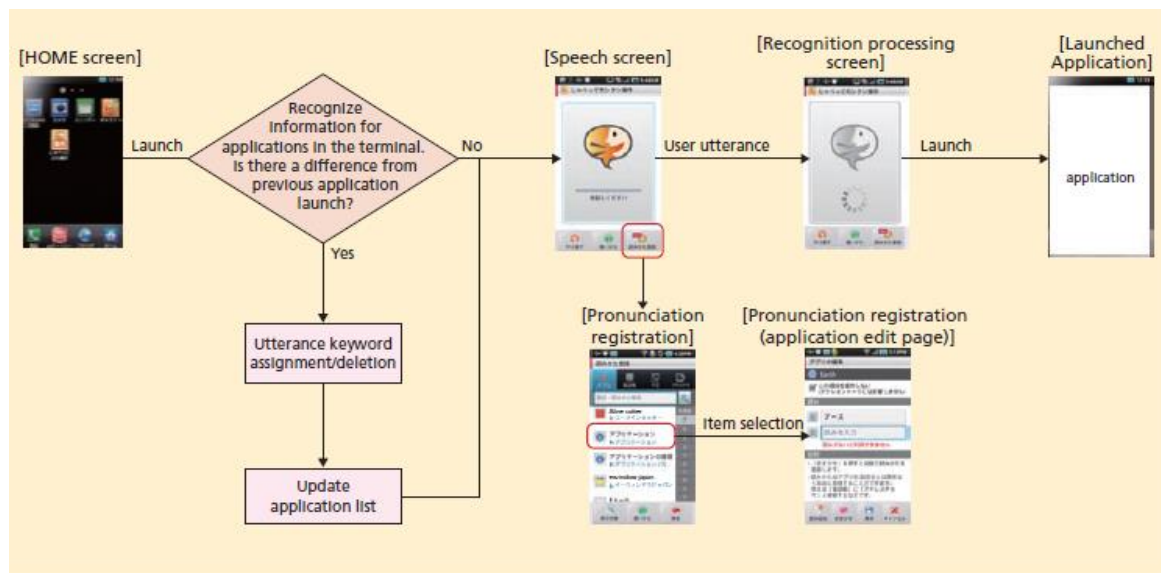

Figure 2: Operation of Voice Action   From (Lizuka, et al 2012 p. 82)
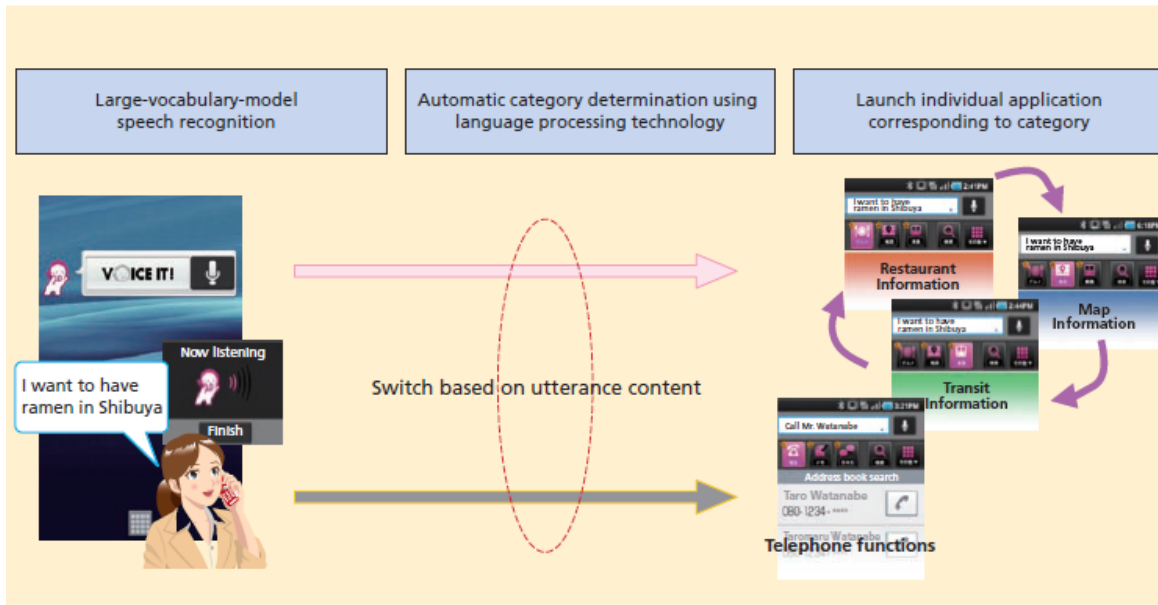
Figure 3. VOICE IT! application screen transitions

From (Lizuka, et al 2012 p. 83)

SRT in Healthcare

The Healthcare industry has witnessed dramatic changes since the late of 20th century and many changes regarding the process of delivering the healthcare services have been incorporated to increase the quality of care. With these changes and advances in healthcare and with the increase rate of competition within the healthcare market many healthcare facilities relied on the new technologies to improve their services and increase their effectiveness and efficiency through a cost effective methodology. A study in 1997 (Rosenthal,, 1997, pp. 436-441) showed that the management and production of clinical reports caused a time and cost burden on those facilities. Hence, many healthcare facilities are now moving to SRT as an alternative of the traditional methods and a solution to increase their productivity through a cost effective way.

There are several studies that have conducted to evaluate the use of SRT in Healthcare field and the results of these studies vary somehow but the majority indicates that the benefits of SRT in healthcare are tangible. One of the studies that was conducted in Walter Reed Army Medical Center showed that the use of

SRT would potentially save the center 520000 Dollar over five years through eliminating the operational cost of the usage of transcriptional positions (Novakoski, 1999). This came in consistent with another study that was carried out in an emergency department that provide services to more than 51000 patients annually that showed that the financial benefits of the SRT usage reached up to more than 645000 Dollar during five years (DMR Group, 1992). Moreover, a survey was conducted on emergency departments of two hospitals which showed the positive financial impact through a five year cost saving of around 472000 Dollar, while the same study showed no impact on saving the time (Novakoski, 1999). Therefore, all those studies confirm that the SRT has a financial advantage for healthcare sectors.

In addition, there was a study that was conducted for the purposes of evaluating the SRT implementation for documentation the outpatients' encounters of the electronic medical records at one of the military hospitals and its 12 clinics which showed many indicators of the SRT benefits. Such results were published by (Hoyt & Yoshihashi, 2010) as below:

- clinicians are producing the clinical reports within the same day of patient visit where they use to take several days before the implementation of SRT

- almost 70% of the clinicians gave the impression that SRT showed more accuracy and is faster than typing.

- 90% of the clinicians believed it improved the quality of their clinical notes

However, the same study as well showed that some of the clinicians by almost 30% discontinued using the SRT system due to several factors such as personal experience and technical or logistical reasons but most importantly inadequate training. According to Ronaldo Parente (2004) who conducted a qualitative study to evaluate the factors that affect the adaption of the SRT among its users, the two main factors that hold back the use of the SRT properly rely on the degree of clinicians' exposure to technology prior the implementation of SRT and the social status, background, beliefs and education of those clinicians (Parente, et al 2004).

As discussed above there were many studies that show the positive impact of the use of SRT. However, a recent study which was conducted by Riso National Laboratory Technical University of Denmark showed negative feedbacks from the physicians of Danish Hospitals. The study showed that the physicians expected more from SRT and they felt that more time was spent on producing the medical records after the implementation of the SRT and that the overall quality of the reports has decreased with the introduction of this technology as only 33% of those physicians believed that SRT was a good idea and the rest disagree with this technology. Some of the rational that was speculated from this study that the factor which has main influence to turn the acceptance into negative results were mainly around the fact that this system was forced and introduced to those clinicians to be as mandatory rather than voluntarily. So the study suggested forcing some kind of change management techniques to turn the acceptance level to its positive side (Alapetite & Andersen, 2009, pp. 36-49).

Ministry of Health and Its SRT Implementation

MOH of UAE is adopting a transformational program to enhance the healthcare delivery system in order to provide safe, effective, patient-centered, efficient and equitable care to the patients. In order to fulfil this transformation, MOH took many initiatives such as implementing Health Information Systems (HIS) that unified and automate the medical records across all the hospitals and clinics under MOH which consist of 15 hospitals, 68 clinics, and 18 other facilities. This initiative was very beneficial however, during the implementation of this HIS project, many of the users struggled with typing the medical reports and documentation, especially that they used to have this process to be proceed manually. As well, transcriptional services and typists to assist the physicians were not available for the majority of MOH hospitals causing increasing the time to produce the medical reports and less time spent with the patients.

From this perspective, MOH moved on to incorporate those efforts of implementing the HIS project with the SRT aiming to ease the daily work of the users and increase the adaption to the systems while lowering the cost of patients documentations and increasing users' productivity.

In November 2011, MOH released a Request for Proposal (RFP) for Medical Speech Dictation System, and Emerging Technology (ET) Company won that RFP by utilizing Dragon Medical from Nuance Healthcare Solutions and SpeechMagic. This SRT provides 99% accuracy according to ET and has built-in medical dictionary with over 80 specialists and subspecialties. Hence the users can easily dictate in to HIS or any other Microsoft System office or windows by placing the cursor into the area where the reports is required to be filled and simply dictate. The SRT used is dependent speech technology and relies on certain stimulation to enhance the learning of the software throughout the use of the SRT.

Aim of the Study

After the implementation of the SRT, MOH wanted to evaluate if this technology really improved the services and could really be used to reduce the cost. After about a month of implementing the SRT, a study was conducted using a questionnaire mainly to investigate the following:

- The level of satisfaction among the users

- Estimate the productivity

- Quality of the reports and patient care

- Speed of producing the reports

- Impact on patient's care and services

- Measure the overall benefits that this technology brought.

Justification of the Importance of this Research

Although many research studies were carried out to study this technology in the healthcare field, none of these studies were conducted in UAE. This research is considered the first conducted in this region. Such studies are needed so other health organizations whether in UAE or even the Middle East region would realize the benefits of the technology and could benchmark with them. It will give them the knowledge of SRT's impact in quality of care, productivity and cost of the organization and will provide them the opportunities to investigate further if there is inconsistency of the results, which will definitely enhance utilizing the most out of this technology.

Methodology

Subjects

The subject for this study is obtained from the staff working in MOH Hospitals where SRT was implemented. The age group of those staff is ranged between 30-50 years old.

Cluster sampling where the population of interest is divided into clusters and those clusters are selected randomly might be favorable methodology in many cases as it is time effective to choose the sample of study interest (Bowling, 2009, pp. 203-206; Neutens & Rubinson, 2010, pp. 134-141) and many who are convinced that random size might be time consuming, however random sampling was selected as a methodology to select the samples for this particular study. The rationale behind that is simple random size is the best procedure for selecting the subjects as equal representatives of the populations would be accomplished unlike the other procedure. (Bowling, 2009, pp. 203-206; Neutens & Rubinson, 2010, pp. 134-141).

Systematic sampling is also impractical for this study as all the subjects were limited to those who use SRT. Therefore, if a systematic sampling was conducted, time would be a problem as the investigator would redo the process until a reasonable number is allocated.

Piloting the Data Collection Tool

The sample for the pretest was conducted from fellow coworkers and the pretesting was administrated in the same fashion/method for the actual study; distributing self-administrated questionnaire (see appendix). Doing this step will give accurate reflections and better adjustments to amend the actual questionnaire. The size of the pretesting group was four coworkers and after conducting the pretesting, the coworkers provided a critical analysis of the questionnaire. The following table.1 illustrates their suggested improvements:

Table.1 Feedbacks of the Coworkers Regarding the Questionnaire	
Criteria of the Questionnaire	**Comments of the staff & my observation**
Sensitivity of the questions	The questions are accepted (not offending)
Question orders and sequencing	The order of the questions are very consistent and logical
Physical layout and spacing	The layout of the questionnaire is very clear and easy to follow, and the spacing is very appropriate
Question wording & language	Language was clear.
Response categories	One of the coworkers recommended that the responses of the questions that have Before and After should be put in parallel instead of having them in sequence. This is to help illustrating the difference and visualize their comparison.
Length of time for answering the questions	The length of time was very appropriate "about 10 minutes or less"
Reliability	The data collected in a precise way, so it is reliable
Validity	The data collect the information I need, so it is valid

The pretesting was conducted before one week of the initiation of the fieldwork/ data collection phase and revision of my questionnaire was conducted during this period. All the recommended changes were added

Measurement & Data Collection

In order to answer the research questions, a data collection tool which is a survey in a form of a self-administrated questionnaire was utilized. By relying on a self-administrated method we will be able to avoid interviewer bias. Interviewer questionnaire method might make the interviewers to ask the questions in a certain way to bias respondents. For instance, they may ask leading questions, or may appear to hold certain values that can lead to a social desirability (Bowling, 2009, pp. 284-286). Besides, when the interviewer asks the questions, sometimes the respondent is not saying the truth to protect his or her confidentiality. However, in a self-administrated questionnaire, this issue was avoided as no one could know that the questions were answered by a particular person "names in the questionnaire was optional" only staff ID. One problem that might emerge regarding a self-administrated questionnaire is that the respondents may not understand the questions; therefore, they may answer the questions incorrectly. To solve this issue, the researcher was present when distributing the questionnaire to clarify any ambiguously of the questionnaire and answer any questions they ask.

The main questions of the survey were covering four areas. The first part of the questionnaire was about the staff's general demographic data. The second part was to investigate the user's view of the SRT. The questions are subjective and written as both close and open questions. The third part of the survey was objective close questions to assess the users of the technology prior and after the implementation of SRT. The final part of the survey was created to measure the overall satisfaction and to provide the user a chance to write his/her general feedbacks.

Study Design

The study design is considered to be a combination of two study designs, which are cross sectional study and cohort study. It can be seen as a cross sectional study as the survey was distributed in one particular time and the research data is allocated at the same point in time (Gordis, 2004, pp. 173-202), which has advantage of having the results obtained relatively quickly. As well a cross sectional study is the best study design for descriptive studies, however, the relationship before and after using the SRT was one of the main aims of our study and cross-sectional study cannot evaluate and assess the relationship. Therefore, cohort study design was incorporated to our cross-sectional survey by asking the participants in the survey retrospective questions. Also, our study is the first study conducted in UAE regarding the aspects of using SRT; hence, it is very important to conduct initial study that will generate the hypothesis and assumption of the topic before we go for a resource consuming study design. This combined design is suitable for our situation when the resources are limited (Gordis, 2004, pp. 173-202).

Analysis

Our study is a descriptive study and all the variables were quantitative and qualitative and depicted as nominal, ordinal and dichotomous measurements. Both mean; the average and median; the middle-most position were measured to find a value about which the observations tend to cluster, for example, finding the median of the staff's age. While mean and median are going to measure the central tendency, standard deviation and variance is used to find the distance of the measurements away from the mean (Neutens & Rubinson, 2010, pp. 240-246). Moreover, to determine the accuracy of the point estimated was possible as the standard error (SE). This was an asset in indicating how precise an estimate is. Furthermore, as most of our data were nominal and ordinal, pie charts were used to demonstrate the data in a visual basis. The statistical tests and graphs were obtained by using Microsoft-Excel software program.

Ethics

All participants in this study had the right to privacy and confidentiality. Several techniques was taken in order to protect their privacy. For instance, all the data were returned anonymously and directly to the research office and unneeded material were destroyed upon completion of the project. Besides, participants had the right of refusing the participation of the study. An approval was taken from the Authorized personnel of MOH to conduct the research in an ethical manner. Regarding the questions in the survey, no offending questions were provided.

Results

Description of the Population

Our population is driven from MOH of UAE, where the staff under study are geographically coming from different part of Dubai and North of Emirates , and composed mainly of the following specialties:

- Physicians
- Radiologist
- Histopathologist
- Surgeons

The range of the age group of these staff is between 32-62 years old. While studying the age distribution (see the table 2) we can see that the standard deviation is 9.45 which means that, on average, data in the distribution are spread by 9.45 years around the mean.

Table 2: Statistical analysis of the age	
Minimum	32
Maximum	62
Range	30
Mean	47
Average	47

Median	47
Standard Deviation	9.452935
Standard Error	1.286381

Moreover in our study we had 1.29 SE which is not that high. This would mean that our estimation is precise and we would not expect it to vary a lot if we were to take a large number representative samples from the population as it would be vary on average by this amount only (1.29).

User Perspectives of the Speech Recognition Technology

The user perspectives and views of the SRT benefits were investigated and of those users 89% believes that this new technology is beneficial. From those who believed in its benefits, a further analysis was taken to investigate the area in which they find its benefits. The below figure provide is a self-explanatory of their opinions

Chart 1: User Perspective of the SRT Benefits
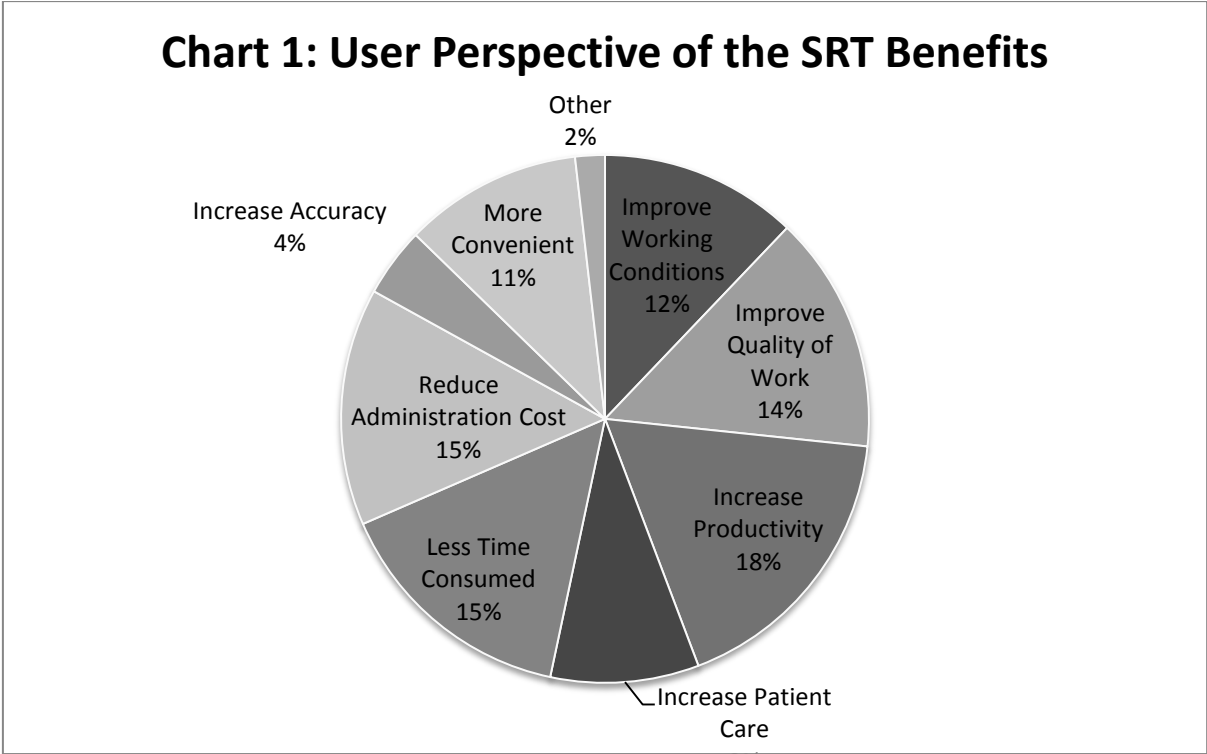
As it is evident that most of them believe that productivity has been increased and administration cost and time have been reduced.

Measuring the Benefits of SRT

After investigating the view of the users, actual objective measurements were taken to answer our research question:

- Investigate productivity through comparing how many reports are produced before and after the implementation of SRT, and how many patients were seen.
- Look into the potential cost saving through investigating the need for transcriptionist after the use of SRT
- Investigate the quality of care through comparing the comprehensively of the report before and after SRT implementation
- Investigating the effect on the time consumed through measuring the time spent in writing a report and the time spent on examine the patient and to validate the reports

Increase the productivity

As mentioned earlier to investigate the productivity we compared the amount of reports produced prior implementing the new technology and after the implementation and the results where depicted as below. The percentage of the increase is as well measured and it reaches up to 14% increase of the productivity only after one month of SRT implementation:

Chart 2: "Reports per Day" Before The Implementation of the SRT

41 to 50 reports
5%

31 to 40 reports
5%

More than 50 reports
5%

21 to 30 reports
11%

0 to 10 reports
58%

11 to 20 reports
16%
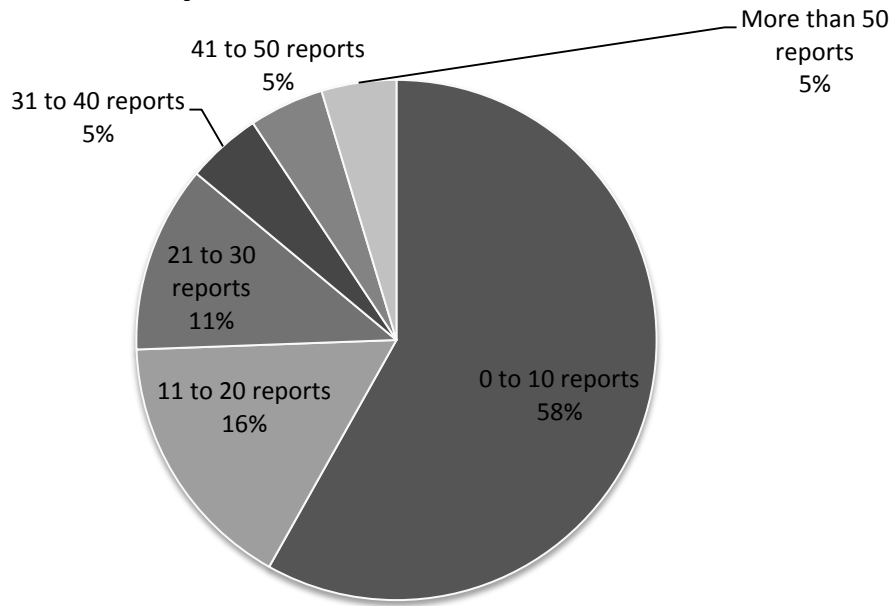
Chart 3: "Reports per Day" After the Implementation of SRT

More than 50 reports
7%

41 to 50 reports
0%

31 to 40 reports
5%

21 to 30 reports
17%

0 to 10 reports
49%

11 to 20 reports
22%
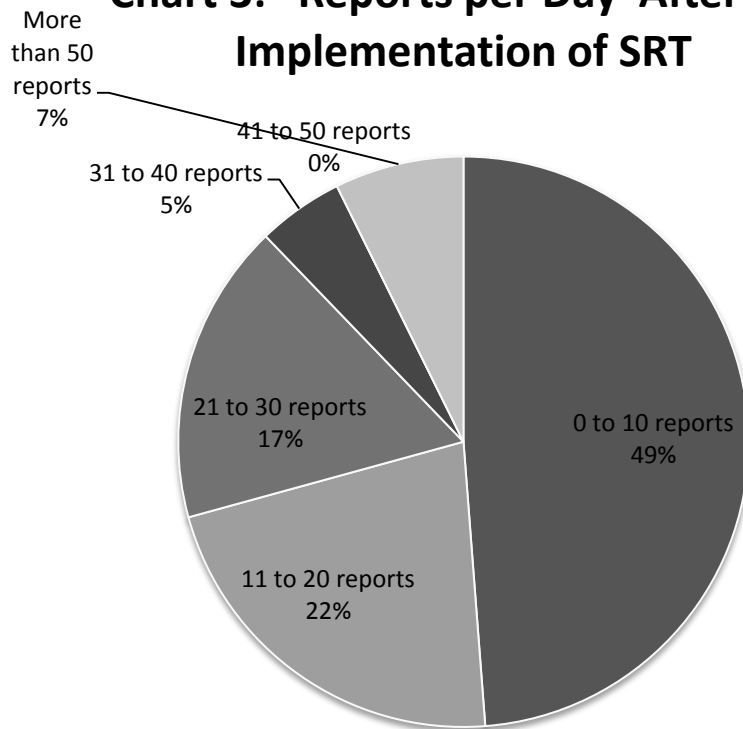
Further investigation into the productivity was taken through analyzing the patient seen number before and after the SRT implementation, once again SRT showed its benefits through 10% increase of patient seen. Below is a more comprehensive view regarding this fact.

Chart 4: "Patient Seen per Day" Before Implementing the SRT

- 0 to 10 patients 20%
- 11 to 20 patients 29%
- 21 to 30 patients 27%
- 31 to 40 patients 2%
- 41 to 50 patients 7%
- More than 50 patients 15%

Chart 5: "Patient Seen per Day" After the Implementation of SRT

Segment	Percentage
0 to 10 patients	20%
11 to 20 patients	29%
21 to 30 patients	22%
31 to 40 patients	7%
41 to 50 patients	2%
More than 50 patients	20%
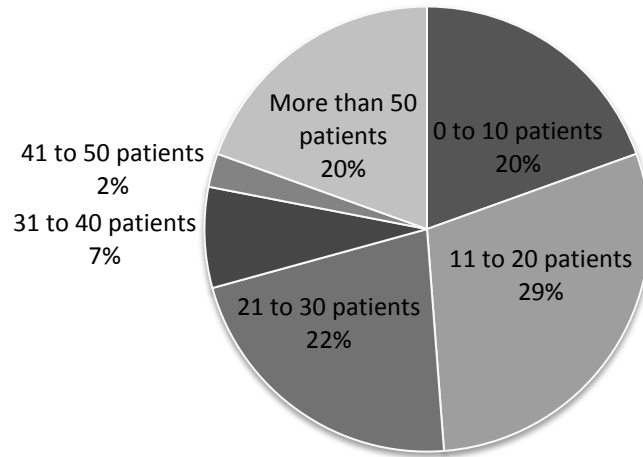
Potential Cost Saving

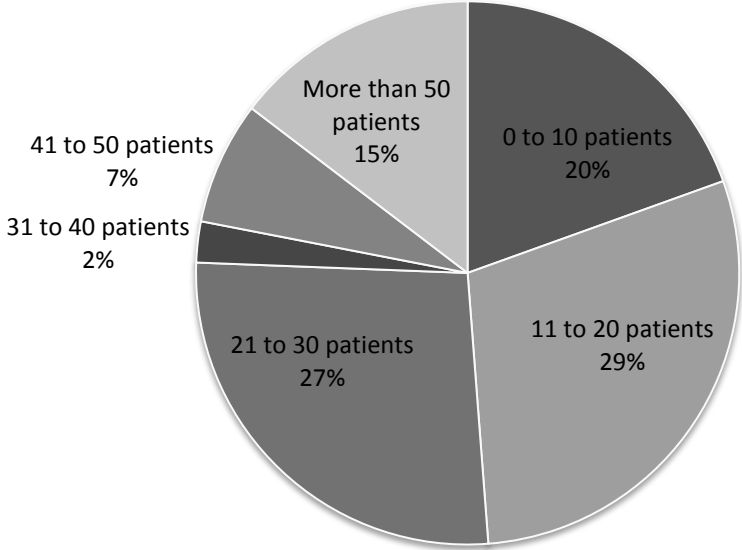The cost saving might emerge from the fact that there are some typists/transcriptionist who are utilized to type cause the physicians used to have difficulty to spare time and type their reports. However, with the use of SRT, typists' positions might be utilized in other professionalism and rotated to have different tasks and job. Therefore, in the survey, the need of the typists even after the SRT implementation was investigated as well the number of typists who they have before and after the implementation. From this point of view, MOH can have potential cost saving that reach about 22% which is almost the quarter of the total typists we have.

Chart 6: Potential Cost Saving

Unrequired Typists
22%

Required Typists
78%
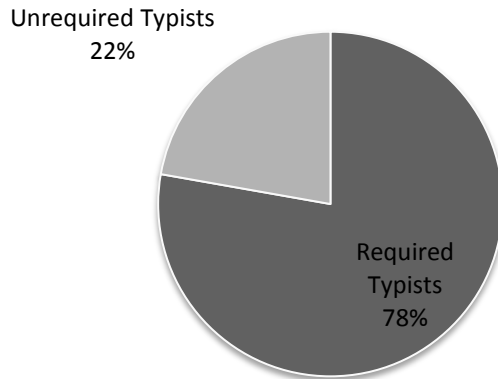
Improve Quality of Care

The more comprehensive the medical reports are the higher quality and patient safely will be obtained. From this perspective, the level of details of the medical reports was examined before and after the implementation of the SRT. While only 9% of the user believed that SRT would improve quality of care, the objective measurements showed us that quality of care is definitely improved as the results showed that the reports are more details and comprehensive by 10%, which evidently will increase the quality of care provided to the patient.

Chart 7: "Reports Comprehinsivily" Before Implementing the SRT

Two pages
2%

A page and a half
10%

Half a page
31%

A page
57%
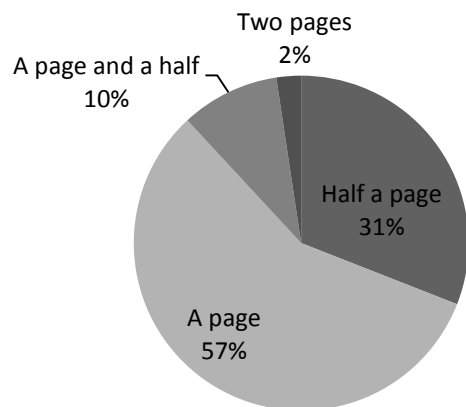
Chart 8: "Reports Comprehensivily" After the Implementation of the SRT

A page and a half
17%

Half a page
24%

A page
59%
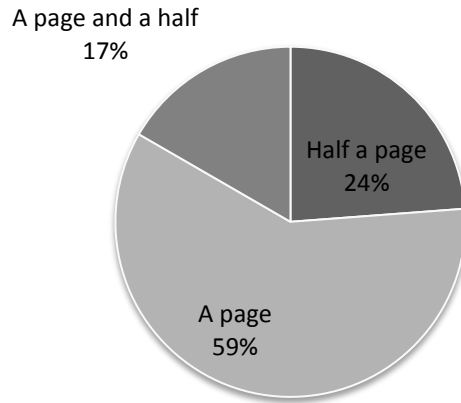
Less Time Consumed

15% of the users believed that this technology actually save them time, and when we conducted the actual objective measurements by analyzing the time spent prior implementing and after implementing the SRT, it was clearly evident that almost 20% of their time spent in writing the reports was saved by this technology.

Chart 9: "Time Consumed" Before the Implementation of the SRT

Average of one Hour
22%

Less than 5 Min
20%

Less than 30 Min
58%
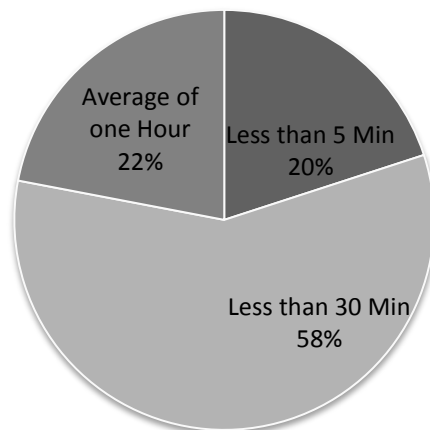
Chart 10: "Time Consumed" After the Implementation of the SRT
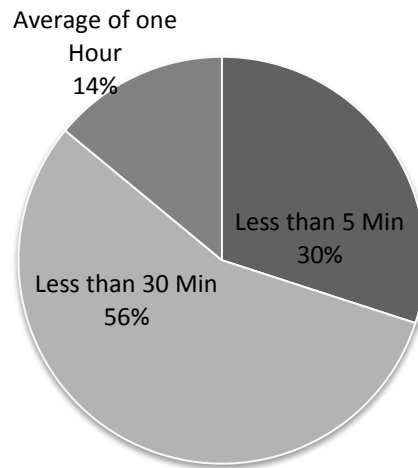
In addition, it was believed that SRT will reduce the time between examining the patient and validating the reports, as before users will have to write the reports later on after seeing the patient but with SRT the reports would be ready to be validated immediately after seeing the patient as the input for the reports were inserted directly as voice. Results showed that the time consumed between seeing the patient and validating the reports is less than before by 18%. Below are two pie charts to show you the comparison before and after implementing the SRT.

Chart 11: "Time consumed" Before the Implementation of the SRT for Validation
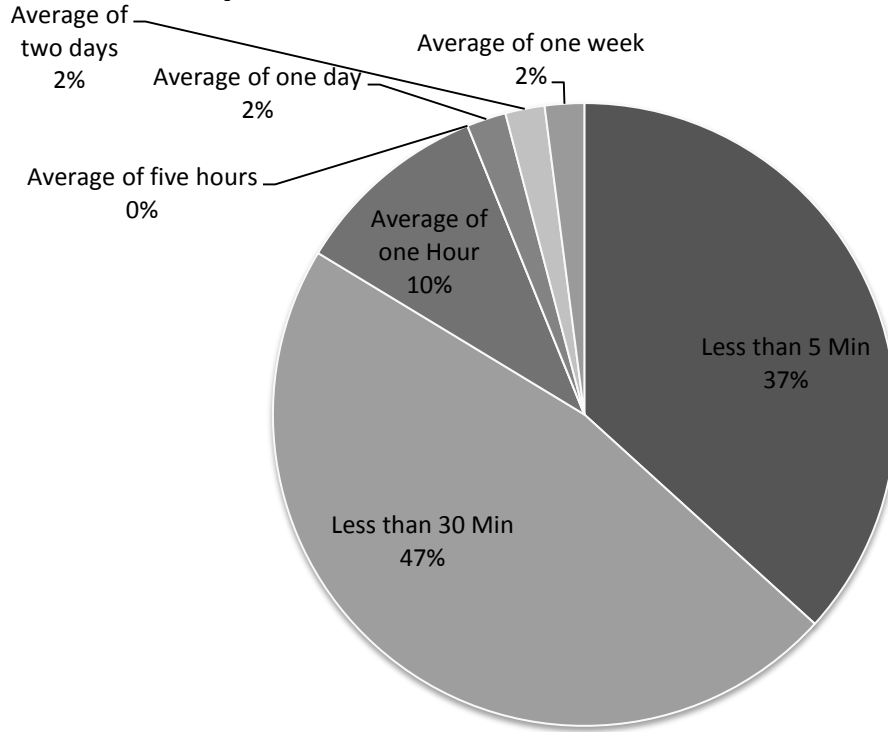
Chart 12: "Time Consumed" After the Implementation of the SRT for Validation

Average of five hours 4%
Average of one day 0%
Average of two days 2%
Average of one Hour 0%
Less than 5 Min 37%
Less than 30 Min 57%
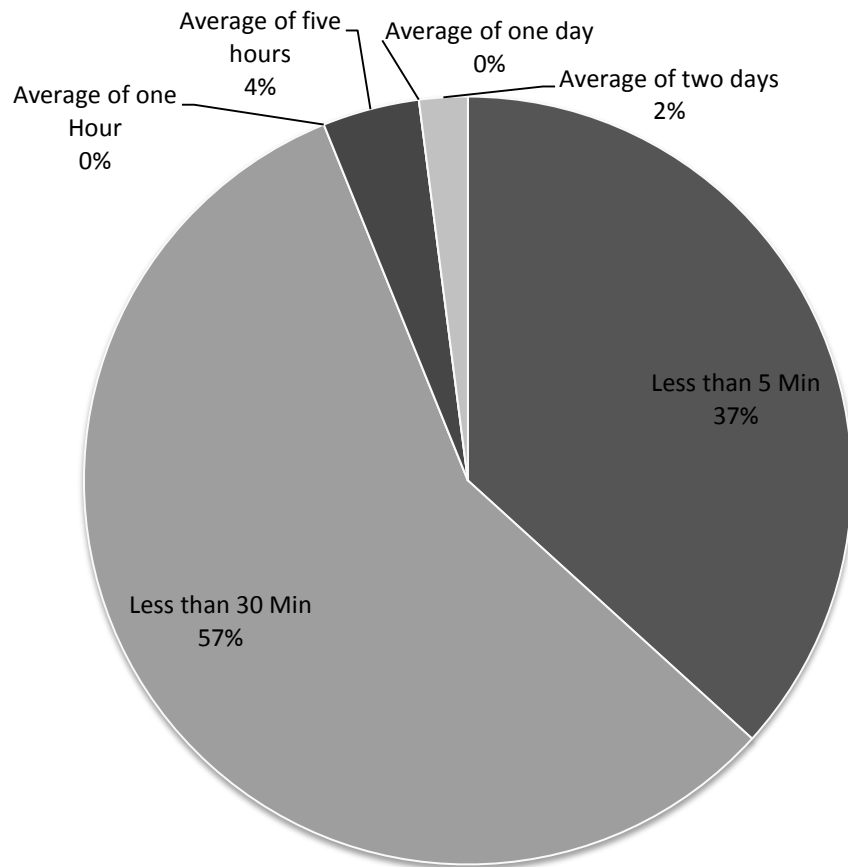
Measuring User Satisfaction

The final part of survey was intended to examine the general satisfaction of the all users who participated in this study, and it was evident that almost more than three quarters of the users are satisfied and less than quarters of the users resisted this technology more likely because of the time needed for the users to adapt to the change and the new technology especially considering that the survey was conducted only after one month of implementing it. Below charts are a clear illustration regarding the satisfaction level of those users.

Chart 13: Satisfaction Level

Extremely Not Satisifed 2%
Extremely Satisified 5%
Not Satisfied 14%
Neutral 27%
Satisfied 52%
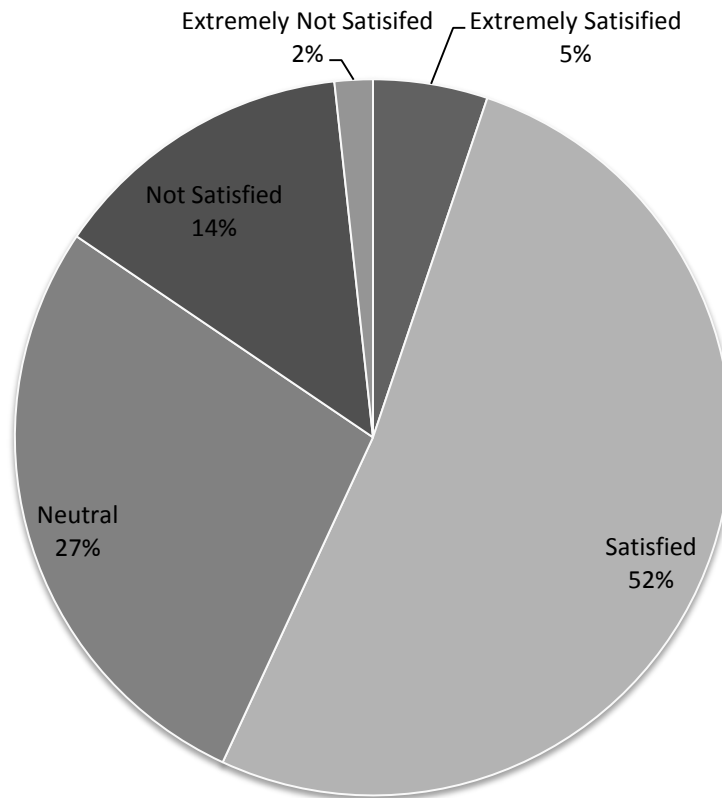
Discussion

With the introduction of EMR to the field many physicians find it a bit difficult to adapt to the new technology, especially those physicians who used to write their documentation manually. An US study showed that the EMR adaption is very slow and only 20% out of 900,000 clinicians use EMR. Reasons behind it rely on the typing speed and the different templates that they need to follow. However, with the introduction of SRT, the adaption use of those technologies is enhanced as it releases them from the typing and transcription burdens (Nuance Communications, 2008). Here in UAE it was not different from US, as with the introduction of the HIS and EMR, physicians and clinicians had

difficulties in adapting it, especially the fact that those physicians are not used to the computer base environments so typing medical records using the keyboard was a problem. However, with SRT the workload was decreased and this burden has been eliminated.

 As mentioned earlier, 89% of the users believed that this technology is beneficial in how it increased their accuracy, increased productivity, decreased the time consumed, made their work more convenient, improved their quality and reduced administrative work which confirms that our users share the same perspectives of the clinicians who were on the study conducted by Hoyt and Yoshihashi (2010) that indicates the same impression of clinicians toward SRT. While this was subjective to their point of view, when the study was carried out, it proved that their point of view is true, as we witnessed how SRT added more value to the daily work of those users.

Results showed that the MOH users' productivity has increased by 14% after only one month of the implementation, and clinicians can see more patients than what they used to have by 10%. If we are examining the Hoyt & Yoshihashi (2010) study, we will find that the productivity of the clinicians in the military hospital under study and its 12 clinics was increased dramatically as well which confirms the point that our study was aiming to prove.

As well, there was another study that confirmed the results of this research paper. The study was by Koivikko, et al in 2008 when they measured the SRT effect on the radiological dictation process and estimated the report turnaround times which is the time between performing the imaging for the patients to the time when the reports are available to the physicians and clinicians. Koivikko and his collegues followed up 14 radiologists periodically for 2 years in a university hospital and were able to have a sample size of more than 20000 examinations. Their study confirmed that their productivity as the amount of reports available within one hour rose from 26% to 58%. Not to mention that the report turnaround time was cut down by more than 80% allowing the radiologist to produce more comprehensive reports for clinicians.

Furthermore, one of the largest benefit that most of these studies have described is reducing the cost as the financial benefits of SRT can reach between 400,000 – 650,000 Dollar (DMR Group, 1992; (Novakoski, 1999). Our study did not come against those studies, it only confirmed them as it showed 22% of potential cost saving. The introduction of SRT has released administrative personnel and typists from the task of transcribing reports. Hence, those personnel can dedicate their time to other tasks that contributes to an enhancement in patient care and allow those staff to be put to better use elsewhere.

In addition, quality of care is increased with the introduction of SRT, in our study the comprehensively of the reports were enhanced by 10%. It is known that the more comprehensive the report is the more quality of healthcare would be obtained as it enables the physicians to take better decisions concerning the patients. Moreover, the report delivery times are considerably shortened and quality of care is increased by faster results delivery which is very much needed especially in Accident and Emergency section as physicians would need to act fast to provide best quality of care to the patients and save their lives and this might be difficult if the radiological reports and laboratory findings reports reported late to them. In addition, these reports are saved in the same system of the HIS used in the hospitals and the same standards, which means that problems of comprehension and loss are avoided.

Further to our discussion, it is inevitable that the use of SRT would reduce the time consumed to provide the reports and provide the medical care services. Our study confirmed this point of view by showing 20% reduction of the time needed to write a report as many of those users used to spend about an hour but now they need just less than five minutes to produce a fully comprehensive report. As well, our study showed that the time between examining the patient and validating the reports is reduced by 18%. To support this, a study was conducted in US showed that the by using SRT the average of documenting the patients encounters is reduced by 50% providing the clinicians and physicians more quality time to spend with the patients ((Nuance Communications, 2008).

While most of MOH users showed their satisfaction, still 16% of the users were dissatisfied with the SRT. The Reasons might be several and from the questionnaire some of those users requested for more training as they did not grasp the perfect way to use it and other clinicians indicated that the time of the training should be associated with the time the software is being installed so they can practice it right away of the training, however the technique used by MOH is to provide the training then install the software by several days. So the way they were trained and inadequate training was one of the main factors why those clinicians were not happy with the SRT. This came in consistent with Ronaldo Parente (2004) study in which this study showed about 30% of users discontinued the use of SRT due to the same reason mentioned here. Many of the clinicians who showed their dissatisfaction complained about the noisy environments that would make dictating their reports a bit difficult, which might add a burden to modify their dictating process. As mentioned previously in the literature review, that the DC incorporates which used LUCAS system experienced some of the difficulties in term on nonstandard accents and variable background noises

Despite the fact that some of the users expressed their dissatisfaction, still many of them were satisfied and the study proved objectively that SRT brought many valuable benefits to MOH and its clinicians.

Strengths

The study was conducted in an ethical manner, as the approval was granted from the organization and all the users had the choice to participate or refuse the participation in this study. Furthermore, as it was indicated before, a questionnaire was used for this study as a tool for measurement, and this questionnaire was pre-tested before collecting the data. This makes the tool more valid as any ambiguously of the questionnaire was fixed according to the feedbacks obtained from the pre-testing phase.

Limitations

It is a fact that the larger the sample size is the confident we would be of an estimate. Unfortunately, as this survey was voluntary the responses to it were limited and out of

200 surveys that distributed to all users who used the SRT at that time, only 64 personnel agreed to participate, however we still got 32% representation of all the population under study which is fairly enough to produce valid results.

It should be noted that my measurement tool (questionnaire) is subjective way to undertake a study and especially that we are relying on the credibility of the participants. It is undeniable that the objective tools are more valid than subjective tool as they depend on the actual results from the tools rather than relying on the participants' credibility. However, measuring this study by an objective tool such as having a third party reviewing the reports before implementing SRT and after, as well estimating the time to produce the reports would be difficult and none practical to be applied on the field and might need more than a year to have the results ready. Therefore, although it would be better to use objective tools, a questionnaire was a reasonable and practical choice for this kind of study especially that it is the first of it is kind in UAE.

In terms of generalizing the study, it can only be generalized to clinicians aged from about 30-60 years, working in UAE, as the knowledge level of computer skills and the user of technology differs from country to country. Therefore, for any future research it is suggested to conduct the study using different samples and participants from different health organizations to be able to generalize the results.

The accuracy level as well were not measured in this study, as it requires a software tool to estimate the accuracy of reports prior and after the SRT implementation. Due to time and resource limitation this was overlooked, and it is suggested that further study should be conducted in this field.

Conclusion

In conclusion, we can see that our results answered our research question and MOH did actually gain the benefits expected from SRT as it showed the following results:

- Increased their productivity by 14%
- Increased number of patients seen for each user by 10%
- Quality of care increased by 10%

- Saving almost 20% of the users time in writing their reports
- Reducing the time between seeing the patients and validating the reports by 18%
- Showed potential cost saving of about 22%.

It is clearly noticeable that this study proved that the benefits of SRT in healthcare field are tangible and that this innovative technology could be considered a tool to increase the adaption of EMR technologies and can add great value on the quality of work, saving cost and time, and productivity of healthcare providers. All in all, it can be said that the contribution of this study was valuable as it is the first study conducted in UAE in terms of the benefits emerged from the usage of SRT in healthcare field. That is why it would be considered the foundation of any future research. As it is a combination of a cross-sectional study and cohort study it would give the researchers the opportunities of generating hypothesis regarding the concept and benchmark their results with the results of this study which established the temporal nature of association being under study. It is recommended for any future study to be conducted in a larger sample size in order to generalize the results on a larger proportion of a population. As well, because this study was performed after the implementation of the SRT by about a month only, a follow-up study is recommended.

References

Alapetite, A., & Andersen, H. B. (2009). Acceptance of speech recognition by physicians

    : A survey of expectations, experiences, and social influence. *International*

    journal of human-computer studies, *67*, 36-49.

    http://dx.doi.org/10.1016/j.ijhcs.2008.08.004

Baumann, J. (1993). Voice Recognition. Retrieved February 21, 2014, from Human

    Interface Technology Laboratory website:

    http://www.hitl.washington.edu/scivw/EVE/I.D.2.d.VoiceRecognition.html

B.H, J., & Rabiner, L. R. (2004, August). *Automatic speech recognition: a brief history of*

    the technology development. Rutgers University and the University of California.

DMR Group. (1992). *A professional benefit study: VoiceEM within a hospital emergency*

    department.

Hoyt, R., & Yoshihashi, A. (2010). *Lessons learned from implementation of voice*

    recognition for documentation in the military electronic health record system.

Itakura, F. (1975). Minimum prediction residual principle applied to speech recognition.

    IEEE transactions acoustics, speech and signal processing, *23*(1), 67-72.

Joshi, C. (2010, April). Speech Recognition. Retrieved February 22, 2014, from Slide

    Share website: http://www.slideshare.net/charujoshi/speech-recognition

Koivikko, M. P., Kauppinen, T., & Ahovuo, J. (2008). *Improvement of report workflow*

    and productivity using speech recognition. Society for Imaging Informatics in

    Medicine. http://dx.doi.org/10.1007/s10278-008-9121-4

Lizuka, S., Tsujino, K., Ogurt, S., & Furukawa, H. (2012). Speech recogntion technology

and applications for improving terminal functionality and services usiblity. *NTT*

DOCOMO Technical Journal, 79-84.

Lowerre, B., & Reddy, R. (1990). [The HARPY speech understanding system]. In

Readings in speech recognition (pp. 576-586). Elsevier Inc.

Neutens, J. J., & Rubinson, L. (2010). *Research Techniques for the Health Sciences*.

Benjamin Cummings.

Novakoski, W. L. (1999). *Using voice recognition to improve medical records*

production. Baylor University.

Nuance Communications. (2008). *Speech recognition: Accelerating the adoption of*

electronic medical records. Nuance Communications.

Parente, R., Kock, N., & Sonsini, J. (2004). An analysis of the implementation and

impact of speech-recognition technology in the healthcare sector. *Perspectives in*

health information management.

Pinola, M. (2011, November 2). Speech recognition through the decades: how we

ended up with Siri. Retrieved February 22, 2014, from TechHive website:

http://www.techhive.com/article/243060/speech_recognition_through_the_decad

es_how_we_ended_up_with_siri.html

Rabiner, L. R., & Levinson, S. E. (1979). Speaker independent recognition of isolated

words using clustering techniques. *IEEE Transactions on acoustics, speech and*

signal processing, *27*(4), 336-348.

Rosenthal,, D. F. (1997). A voice-enabled, structured medical reporting system. *Journal*

of the american medical informatics association, 436-441.

Shahamiri, S. R., & Binti Salim, S. S. (2014). Real-time frequency-based noise-robust automatic speech recognition using multi-Nets artificial neural networks: a multi-views multi-learners approach. *Neurocomputing*, *129*, 199-207.

Suendermann, D. (2011). *Advances in commercial deployment of spoken dialog systems*. Springer.

TechTarget. (2009). Leveraging speech recognition technology in call centers. Retrieved February 22, 2014, from TechTarget website: http://searchcrm.techtarget.com/report/Leveraging-speech-recognition-technology-in-call-centers

Yurick, S. (2011). *Evolution of speech recognition technology in the warehouse*. LUCAS Systems.

Appendix

A Survey for Evaluating the Speech Recognition Technology Implementation and Its Adaption Level, and Measuring its Benefits

We would be very grateful if you would complete this questionnaire, which should only take less than 10 minutes, and is concerned about the overall implementation of the speech recognition or what is called the speech recognition technology. This questionnaire is an important step in this study that aims to investigate the level of satisfaction among the users, and estimate the productivity, quality, speed of producing the reports, impact on patient's care and services and measure the overall benefits that this technology brought. YOUR ANSWERS ARE IMPORTANT TO US, as it will assist for better plans for any future interventions.

Please read through the covering letter before filling in this questionnaire. Participation is entirely voluntary. All responses will be treated as strictly **confidential**.

General Questions	
1) Name (Optional):	
2) MOH ID:	
3) Title	
4) Hospital:	
5) Date of Birth:	_ _ _ _ _ _ _ _ Day Month Year
6) Specialty	☐ Radiologist ☐ Surgeon ☐ Histopathologist ☐ Physician ☐ Other: please specify ……………………
7) Contact Number	Mobile Number: ……………………………………. Office Number: …………………………………….. Email: ………………………………………………..

User's View of the Speech Recognition (SRT) Technology

8) Do you think that the Speech Recognition Technology (SRT) was beneficial?
   - ☐ Yes
   - ☐ No

9) In your view, what are the benefits that the SRT brought to you and your hospital? (You may select more than one answer)
   - ☐ Improve working conditions, workflows and services through better decision support and less stress for the user
   - ☐ Increase quality of work as more complete and comprehensive reports are presented
   - ☐ Increase the productivity as more reports are produced with less time
   - ☐ Better patient care and safety is increased
   - ☐ Less time consumed and less waiting time which would increase efficiency and speed
   - ☐ Reduce administration cost as less paper work, and the need to have typist is eliminated
   - ☐ Increase accuracy
   - ☐ More convenient
   - ☐ Other (please specify :………………………………………………)
   - ☐ No benefits

10) Did the SRT meet your expectations?
   - ☐ Yes it did
   - ☐ Yes and it exceeded my expectations
   - ☐ No

11) Do you think you need typist even after implementing the SRT?
   - ☐ Yes
   - ☐ No

12) Referring to the above question, If yes please specify why?

   ....

13) Do you enjoy working with the system (SRT)?
   - ☐ Yes
   - ☐ No

14) How many reports do you produce per day:

14.1) **Before** implementing the SRT system
- ☐ 0 to 10 reports
- ☐ 11 to 20 reports
- ☐ 21 to 30 reports
- ☐ 31 to 40 reports
- ☐ 41 to 50 reports
- ☐ More than 50 reports

14.2) **After** implementing the SRT system
- ☐ 0 to 10 reports
- ☐ 11 to 20 reports
- ☐ 21 to 30 reports
- ☐ 31 to 40 reports
- ☐ 41 to 50 reports
- ☐ More than 50 reports

15) How many patients do you see per day:

15.1) **Before** implementing the SRT system
- ☐ 0 to 10 patients
- ☐ 11 to 20 patients
- ☐ 21 to 30 patients
- ☐ 31 to 40 patients
- ☐ 41 to 50 patients
- ☐ More than 50 patients

15.2) **After** implementing the SRT system
- ☐ 0 to 10 patients
- ☐ 11 to 20 patients
- ☐ 21 to 30 patients
- ☐ 31 to 40 patients
- ☐ 41 to 50 patients
- ☐ More than 50 patients

16) How many of your reports are processed <u>with SRT</u> Technology each day?

16.1) **Before** implementing the SRT system (if using an older system)
- ☐ 0 to 10 reports
- ☐ 11 to 20 reports
- ☐ 21 to 30 reports
- ☐ 31 to 40 reports
- ☐ 41 to 50 reports
- ☐ More than 50 reports

16.2) **After** implementing the SRT system
- ☐ 0 to 10 reports
- ☐ 11 to 20 reports
- ☐ 21 to 30 reports
- ☐ 31 to 40 reports
- ☐ 41 to 50 reports
- ☐ More than 50 reports

17) How many typists your department has to type and correct your reports before the implementation of SR?

17.1) **Before** implementing the SRT system
- ☐ None
- ☐ One typist
- ☐ Two typist
- ☐ More than two

17.2) **After** implementing the SRT system
- ☐ None
- ☐ One typist
- ☐ Two typist
- ☐ More than two

18) What is the average length of your reports?

18.1) **Before** implementing the SRT system
- ☐ Half a page
- ☐ A page
- ☐ A page and a half
- ☐ Two pages
- ☐ More than two pages

18.2) **After** implementing the SRT system
- ☐ Half a page
- ☐ A page
- ☐ A page and a half
- ☐ Two pages
- More than two pages

19) What is the average turnaround time for one report?

19.1) **Before** implementing the SRT system
- ☐ Less than 5 min
- ☐ Less than 30 min
- ☐ Average of one hour
- ☐ Average of five hours
- ☐ Average of one day
- ☐ Average of two days
- ☐ Average of one week
- ☐ More than a week

19.2) **After** implementing the SRT system
- ☐ Less than 5 min
- ☐ Less than 30 min
- ☐ Average of one hour
- ☐ Average of five hours
- ☐ Average of one day
- ☐ Average of two days
- ☐ Average of one week
- More than a week

20) What is the average time between the examination and the validation of the report?

20.1) **Before** implementing the SRT system
- ☐ Less than 5 min
- ☐ Less than 30 min

20.2) **After** implementing the SRT system
- ☐ Less than 5 min
- ☐ Less than 30 min

☐ Average of one hour		☐ Average of one hour	
☐ Average of five hours		☐ Average of five hours	
☐ Average of one day		☐ Average of one day	
☐ Average of two days		☐ Average of two days	
☐ Average of one week		☐ Average of one week	
☐ More than a week		More than a week	

Overall Satisfaction and Recommendations

21) Please rate your overall satisfaction of the SRT technology that is implemented in your hospital.

    ☐ Extremely satisfied

    ☐ Satisfied

    ☐ Natural

    ☐ Not satisfied

    ☐ Extremely not satisfied

22) What do you recommend to enhance the current SRT technology? (Optional)

   ----

23) Feel free to add any additional comments (Optional)

Thank you for your participation, we really appreciate your time