Sept
2013

WHITE PAPER

Dr Anna Bayes

SPEECH RECOGNITION – WORTH A SECOND LOOK

WINSCRIBE

Speech Recognition – Worth a Second Look

WITH HEALTH SECRETARY, JEREMY HUNT'S CALL FOR THE NHS TO BE PAPERLESS BY 2018, DR BAYES LOOKS AT THE ROLE THAT SPEECH RECOGNITION CAN PLAY AND ENCOURAGES HEALTHCARE ORGANISATIONS TO CONSIDER TAKING ANOTHER LOOK AT THE TECHNICAL ADVANCES THIS SOFTWARE CAN OFFER.

16-18 Empress Place • London SW6 1TT • United Kingdom
Phone +44 (0) 207 471 0100 • Fax +44 (0)20 7610 1806
Email sales@winscribe.com • Web www.winscribe.co.uk

Table of Contents

Summary

NHS trusts are experiencing increasing financial pressure. In parallel, the commissioning bodies are requiring better quality care and evidence of timely delivery. Improved technology can help deliver some of these required efficiencies. Speech recognition is one such example, which has been used sparsely in healthcare for many years. Recent advances mean that speech recognition is now more relevant to a wider range of clinicians.

This paper examines the main types of speech recognition and its deployment. It covers the viability of its place in the working practice of hospital doctors, nurses, medical secretaries, and general practitioners. It also reviews the benefits for patients, trust IT departments, clinical coding and trusts as a whole.

What is Speech Recognition?

Speech recognition utilises computer software to translate spoken words into text. The technology behind this software has been around for forty years and there has been significant incremental improvement over this time. Now, speech recognition is becoming commonplace and is available on smart phones, for voice dialling or for voice routing of phone calls to the correct recipient. Historically, speech recognition as a form of transcription has been used for many years in healthcare but this generally has been restricted to a few discrete areas. Recent advances however, are enabling the technology to be applied to wider healthcare user groups with concurrent cost savings, time saving and most importantly, reduction in clinical risk, as clinical documentation is transcribed more efficiently and is therefore available to other downstream clinicians to act upon.

The accuracy of speech recognition software improves with each use, as the software "learns" from both the speech patterns of the author and the manual corrections that are made by the author or a medical secretary. Training and the use of advanced features, such as macros, can further enhance both the output quality and the experience of using the software. However, whilst accuracy in speech recognition is important, far more important is the turnaround time of document production and the ability to pinpoint any bottle necks in the production process when integrated to a workflow management tool.

Types of Speech Recognition

Front-end speech recognition

This technology enables the author to view the words on screen as they are transcribed in front of him. He can then make any edits as he sees fit. The document is then usually forwarded to a medical secretary who may add demographic details or format the manuscript for printing. This method usually reduces the need for multiple editing/reviewing cycles as the clinician sees his mistakes and learns to dictate more accurately.

Back-end speech recognition

To the author, this method does not appear to be any different to using a digital dictation system. He dictates as usual, however before the speech reaches the transcriber, it is passed through a speech recognition engine to turn it into text. The medical secretary can then review the text whilst listening to the voice file, allowing quality checks and revisions to be made before formatting and returning to the author for editing and signing.

Historical users in Healthcare

Speech recognition has been successfully used for many years by radiologists and histopathologists in acute trusts. They tend to be able to dictate their long wordy reports in quieter parts of the hospital, turning their speech directly into text. Their working patterns also allow them to set up their headphones and microphones at a single workstation minimising the set up time for hardware.

Their practice also lends itself to the use of many macros or "canned text" that enable them to dictate commonly used phrases and sentences very quickly and accurately. The use of speech recognition by these specialties has enabled them to have faster turnaround times for report writing, transcription and most importantly for passing critical results on to clinicians for urgent action.

In parallel, they have seen excellent return on investment and have been able to redeploy some of their secretarial staff to other duties.

The software has been improved considerably over the years; accuracy and speed of acceptance have led to its increasing adoption amongst even the most Luddite radiologists. To date, however, the widespread use of speech recognition across other hospital specialties has not been seen. There are many reasons for this: firstly, most patient-facing hospital doctors do not have a fixed workstation which they use for most of their dictation. Secondly, they tend to dictate in fairly noisy environments such as busy outpatient clinics.

However, the flexibility of the speech recognition software has improved significantly over the years. In parallel, scalability, speed, ease of use and ease of deployment are factors that now make speech recognition a viable option for transforming the way documentation is produced.

The following diagram illustrates the different processes currently used in the production of clinical correspondence. Where manual processes are used, production times and costs are greater, but with the deployment of integrated digital dictation speech recognition Trusts can make considerable savings by moving away from paper processes and turnaround times of letters are significantly lower.

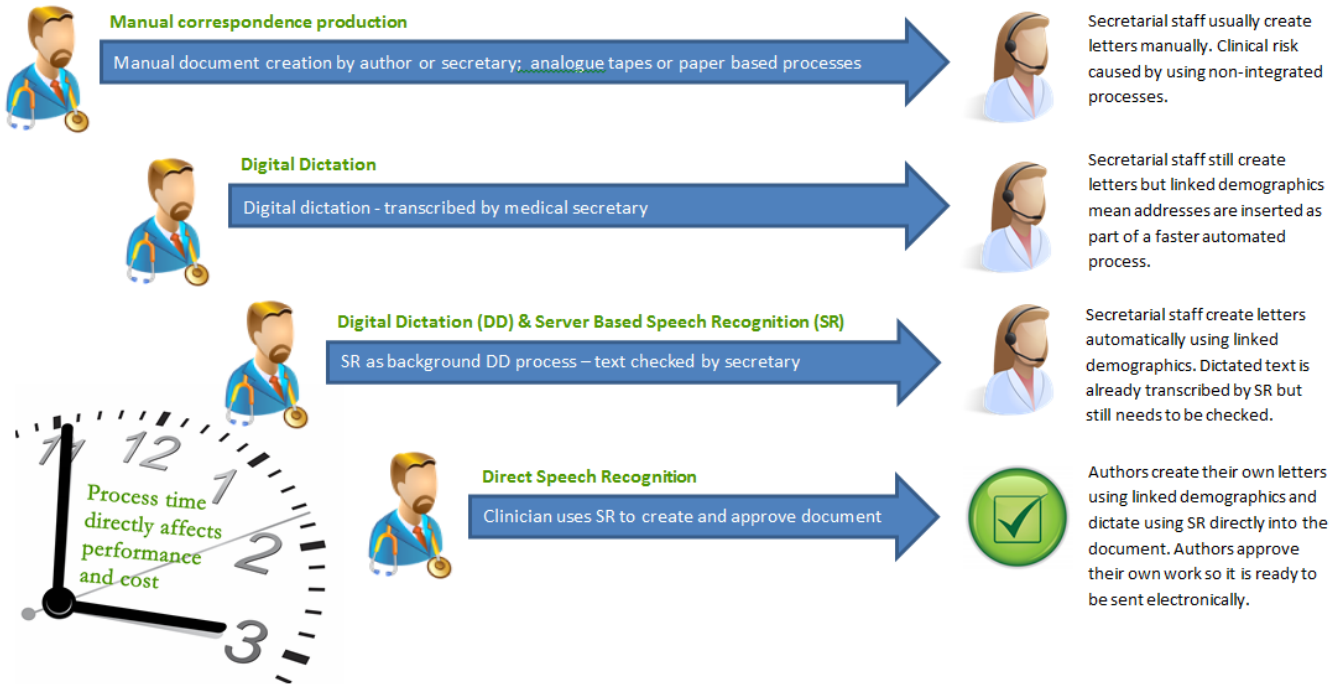

Figure 1: Clinical Correspondence Process - From Manual to Speech Recognition Enabled; Source: Winscribe 2013

Transformation of services

Managing change within the NHS is no easy task and technological implementations are no exception. The King's Fund's report listing the main barriers to delivering NHS technology-based change includes:

> *"The availability of resources, both financial and organisational, affects the ability of the health service to change across the range of its activities, including the use of technology"*[1]

[1] Technology in the NHS-Transforming the patient's experience of care, Alasdair Liddell, Stephen Adshead and Ellen Burgess, The King's Fund, 2008

Building in the appropriate level of transformational support into technology programmes is key to delivering the benefits.

A study of the use of speech recognition technology in the USA, showed that continued use of the software post deployment was associated with a positive training experience and a positive perception of how speech recognition improved note quality and clinician productivity.[2]

The Hospital Doctors' View

Excluding histopathologists and radiologists, there have been a few pockets of hospital-based clinicians who have dipped their toes into speech recognition, most commonly on an individual basis. Many trusts have explored the benefits of digital dictation as part of their need to cut costs. However, despite a reduction in turnaround time for transcribing documents (in many cases this has been shortened by 50%), few of these trusts have seen the desired financial benefits in deploying digital dictation.[3]

Hospital doctors use dictation for different aspects of documentation. The two main areas that involved dictation in the past were the production of the clinic letter and the discharge summary. Both of these documents provide a mechanism for the hospital doctor to share details of an episode of clinical care with their colleagues in other settings or departments. It also becomes the primary record of that episode of care since it is usually legible, relatively easy to locate and particularly in the case of discharge summaries, is presented in a consistent format.[4] The need to produce these documents in an accurate, reliable and efficient manner has become imperative as commissioners are demanding specified turnaround times for their receipt and quality of content. To keep up with these demands, trusts

[2] Lessons Learned from Implementation of Voice Recognition for Documentation in the Military Electronic Health Record System; Hoyt et al; Perspect Health Inf Manag. 2010 Winter.

[3] How to utilise digital dictation technology for cost and time efficiencies; Alison Moore: HSJ;26 April 2012

[4] Kusnadi, Kusnadi (2012) Moving From Paper Based To Electronic Hospital Discharge Summaries: A Mixed Methods Investigation. Doctoral thesis, University of Huddersfield.

are seeking new solutions to facilitate the documentation process and also to integrate these into their strategies for moving to paper-lite or paper-free environments.

Clinic

It can be argued that clinic letters provide the ideal opportunity to introduce speech recognition. If speech recognition software is integrated with the trust's Patient Administration System (PAS), patient demographics can be added to the transcribed document without the need for dictation of these data. In addition, if the integration also provides a link to a patient list from a clinic, the clinician can then quickly see which patients' letters have been dictated and which have not. Both of these features add to clinical risk reduction. For clinicians opting for front-end speech recognition, deployment of well-designed software with a user interface that allows the clinician to edit text as they dictate or at the end of dictating, depending on their preference, facilitates user acceptability. This can then be coupled to a workflow engine to facilitate the transfer of the file to the medical secretary (for back-end speech recognition this will be a speech file, for front-end this will be transcribed text). Speech recognition software has the ability to learn from repeated dictation patterns of individual authors and improve accuracy over time. The flexibility for end users to select front-end or not (i.e. default to back-end) optimises the ability to cater for the widely differing requirements of clinicians, which is often based upon their familiarity with the use of IT in clinical practice and in their personal use.

The following diagram shows the typical Clinic letter workflow process from dictation in clinic through Transcription, Approval to E-Distribution using integrated Digital Dictation with linked demographics from PAS for faster letter creation. This integration helps remove dependency on paper processes and improves clinical governance by removing the needs to manually enter or copy and paste patient data.

The enhanced process using server based speech recognition is highlighted in green. Turnaround times are improved by using SR to transcribe text. Secretaries check text for errors – this is faster than typing the text from scratch and frees up secretarial time for other duties.

The third and fastest way to create letters is to allow authors to create and approve letters themselves using speech recognition directly into templates with linked demographics (Highlighted in red). This dramatically reduces the number of steps required to create a letter and speeds up turnaround times noticeably. Less secretarial time is spent on letters and more on the work of their department.

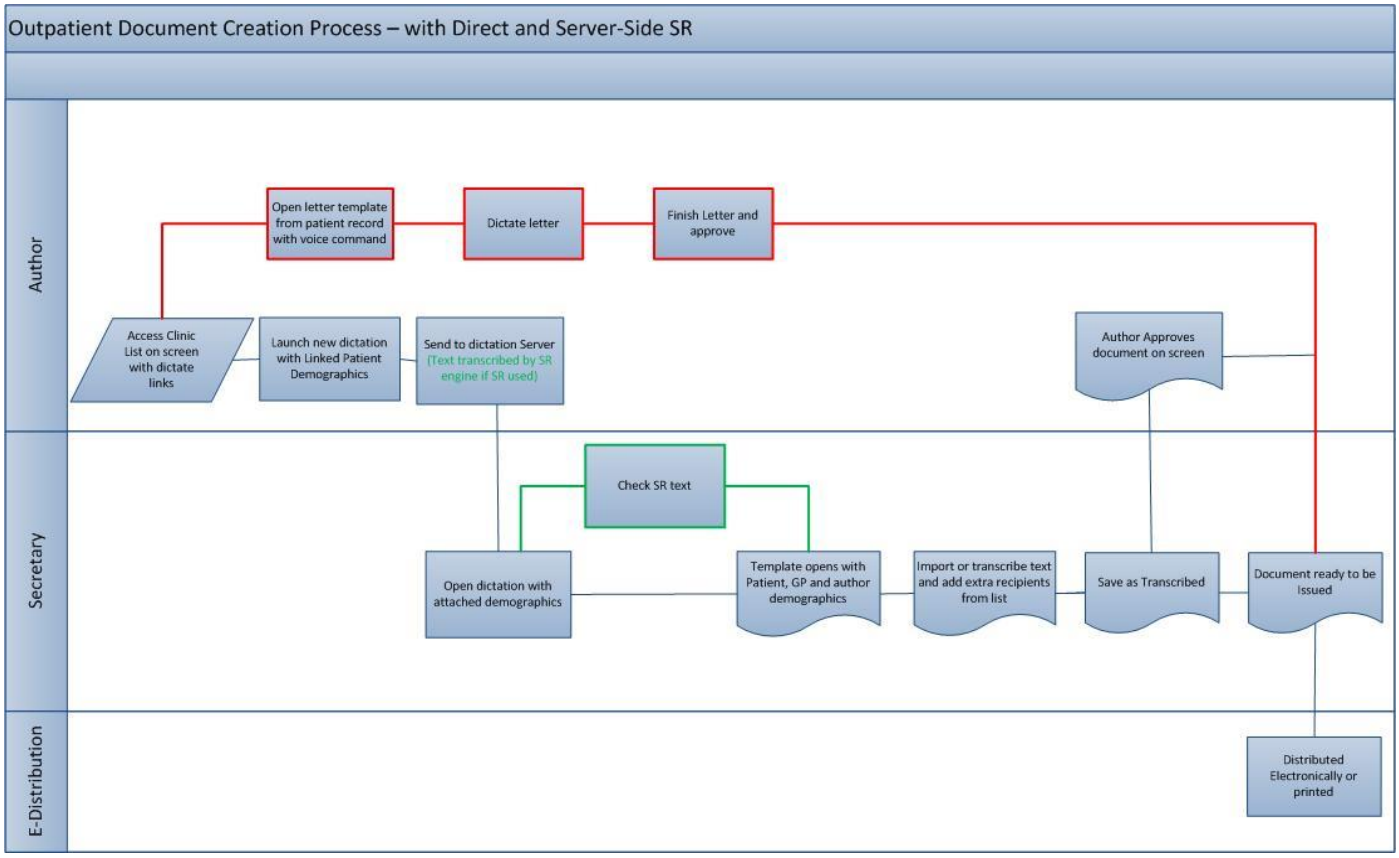

Figure 2: Outpatient Document Creation Process with Direct and Server-Side Speech Recognition; Source: Winscribe 2013

Ward

Use of speech recognition to aid the production of the discharge summary is a more complex process. Ideally the discharge summary should no longer be dictated at the end of a patient's hospital stay, but should be iterated throughout the discharge planning process by multiple authors, starting at the time of admission. Discharge summaries are highly structured documents and their relevance goes beyond the mere communication of clinical progress: they are also, in most cases, the main input into the clinical coding process which trusts rely upon for revenue generation. For this reason, applying speech recognition technology is not as straightforward as it is in the clinic letter scenario. However, most recently, there have been big advances in the use of voice command options in addition to speech recognition. These allow the speech recognition software to become fully integrated into the discharge summary creation tool. In some cases this may be a stand-alone document production tool, in others it may be the hospital's electronic patient record (EPR). The benefit of the latter is that it has been shown to not only speed up the ability of clinicians to enter structured data into the appropriate field of an EPR but also to greatly enhance user acceptance of the EPR. KLAS research in the USA identified that 76% of the clinicians using "desktop" speech recognition (directly controlling an EPR system via speech) report faster turn-around time as the largest benefit, together with better service to patients.[5]

In a study across 80 Emergency Departments (ED) in New South Wales, Australia, 2000 doctors were trained in the use of speech recognition to facilitate the production of ED discharge summaries. They used voice command to navigate the ED system that was already in situ and had been the subject of huge user frustration to the extent that many clinicians had refused to use it. The advent of voice command reduced the need to navigate with mouse clicks and when coupled with voice recognition, allowed the clinician to record the necessary structured clinical data accurately and quickly, producing the ED discharge summary in parallel. More importantly for the hospitals, the quality of the data collected improved so much that revenue from emergency department visits increased. For clinicians, the benefits of the EPR began to accumulate, as the record became data rich and hence more useful. The consequent freeing up of clinician time in a

[5] Dr. "Multi-Task": Using Speech to Build Up Electronic Medical Records While Caring for Patients; Shagoury, chapter in Advances in Speech Recognition 2010, pp 247-273

busy ED also resulted in a faster throughput of patients and allowed more time to be spent on caring for patients.[6]

As new hardware becomes available to clinicians and "bring your own device" is more prevalent, technology suppliers are seeing the benefit of making their solutions available on multiple platforms. Clinical apps for inputting and reviewing EPR are now available, and some clinicians are even using their own tablets and smartphones to dictate directly into the EPR. The mobility of these devices means they are flexible for diverse clinical environments and potentially will allow clinicians to use their time more efficiently e.g. whilst waiting in operating theatres. For trusts, this potentially provides another route for deploying speech recognition to clinicians wishing to use their own portable devices. Trust-owned or clinician-owned mobile devices will also facilitate the viability of deploying and the user acceptance of speech recognition to mobile workers. As more care is being delivered in the community, peripatetic workers have a growing need to efficiently record their clinical activity whilst away from their base.

Opportunities for Discharge Planning

As speech recognition software has improved, the ability to use it successfully in relatively noisy environments (such as the EDs in New South Wales) has been shown. It has also been easier to train more users more quickly. Busy clinical areas were previously thought of as too noisy for deployment of speech recognition. In addition, clinicians tend to move around wards and clinical areas during their working day and in the past, it would have been too costly to provide adequate microphones in all clinical areas. The high turnover of junior staff was also seen as barrier. The success of the programme in New South Wales provides evidence that it is possible to train many junior staff quickly to use a speech recognition system to generate elements of the discharge planning process that culminate in a multi-discipline multi-authored discharge summary. The use of speech recognition should reduce the discomfort of many clinicians who are impeded by their poor keyboard skills and hence

[6] Manly Emergency Department Voice Recognition Evaluation; NSLHD/New South Wales Ministry of Health; April 2012

chose to type poor-quality scant documents. The freedom of using speech recognition embedded into a smart template which pulls in the relevant demographic details whilst prompting population of the relevant fields should improve the quality of documentation. This will not only enhance information delivery to the recipients but also aid the trust in reaping appropriate compensation through a better-informed coding process.

The GP's View

Primary care in the UK leads the world in paperless EPR healthcare adoption.[7] Different practices have different methods of producing documentation, which most commonly takes the form of referral letters. In many cases, the referral letter is often copied and pasted directly into Choose and Book. For some practices, the responsibility for typing these referrals is with the doctor, for others, the ability to dictate exists. The volume of dictation is generally far smaller than that in secondary care; however, the risks of delay or items getting lost still have the same potential clinical consequences.

Dr Owen Uprichard, a General Practitioner in Leicestershire, is very aware of the latent hazards associated with document production in general practice. As a Clinical Assistant at Leicester Royal Infirmary, he also works in a hospital clinic environment once a week where he has experienced the benefits of a workflow-enabled digital dictation system provided by Winscribe.

> *"Prior to the digital dictation system being available, I had trouble finding a functioning Dictaphone and blank tapes for my clinic. I was also concerned that the tapes would get lost or damaged. Most importantly, I worried that the delay in typing would lead to letters slipping through the net with clinical consequences. Now I find that my dictation is always transcribed quickly and I am confident that any letters that I mark as urgent will be treated appropriately."*

[7] Matthew Swindells, Solution or Distraction? BCS ITNOW,2012, Volume 54, Issue 1, Pp. 27-29

Dr Uprichard feels that such a system would be advantageous in general practice so that one could easily track the production stage the letter had reached if a patient returns to the surgery. Currently, Dr Uprichard is responsible for typing his own referral letters and then messaging them to other staff in the surgery for copy and pasting into Choose and Book. There is no secretarial support in his small surgery to transcribe voice files and implementation of such a system might be considered heavy-handed for the relatively small volume of paperwork his practice generates. The use of speech recognition software, together with some workflow management tools in this environment could save Dr Uprichard time in manually typing his letters. It could be fully integrated with his GP system to facilitate accurate and timely transcription of his documents whilst providing the rigorous audit trail to minimise the risk of lost letters.

The Nurse's View

The role of nurses is changing throughout many areas of practice and the prevalence of the Nurse Consultant and Nurse Practitioner as well as Consultant Midwife roles mean increasing autonomy and responsibility for patient care. Invariably this is associated with the need to produce timely, structured documentation. The use of speech recognition software in these situations is not too dissimilar from its use by doctors. However, for the majority of nurses working in busy departments, there is a daily struggle to balance patient care with the ever-growing requirement to document activity. The potential to use a combination of speech recognition with voice command to allow recording of structured clinical data directly into an EPR could potentially make nurses more efficient as they dictate on the go into small smart phone size devices or tablets. This could particularly be a boon for the subset of users with minimal keyboard skills.

One recent deployment of tablets to community nurses in Bristol had a number of unexpected benefits. One of these was that the nurses used latent features of the tablet PC to augment the functionality they were trained to use. It was found that nurses started using the speech recognition functionality on the tablets whilst out and about to enter data directly into the clinical software more efficiently.

The Medical Secretary's View

The dawn of digital transcription caused concern for many medical secretaries who feared their role would not survive. This has not been the case for many reasons. Where digital dictation has been implemented most successfully, the role of the medical secretary has evolved from transcriber to editor. In addition, valuable time has been released to allow them to assist their clinical teams more effectively. For those who have been involved in pilots or roll outs of speech recognition, they have seen their role transform into that of many early adopters of new technology – they have become super-users. For some who have embraced the opportunity, they have gone on to become trainers for their colleagues – both secretarial and clinical - in optimal use of the solution, thus enhancing their role.

The speech recognition software also enables improved balancing of workloads through the ability to report transcription turnaround time. Full introduction of an integrated workflow management system with speech recognition allows managers to see any delays in document production and actively manage workloads to ensure timely creation. This of course is becoming more important as both discharge summaries and clinic letters are now incorporated into most trust CQUIN frameworks and delays in production may result in loss of CQUIN revenue.

The IT View

Whilst primarily a technology solution, deployment of speech recognition should be sponsored by the business in order to reap the benefits. This is the single most important lesson that has been learned from many hospital IT implementations. Change management, clinical engagement and executive sponsorship will facilitate the adoption of technology across the stakeholder groups. In particular, it has been shown that training is best provided via a multi-channel approach that is on-going post-deployment to allow users to refresh and build on their knowledge as their confidence in using the new tools increases.

Users are quick to form an opinion of any new technology. If the opinion is not favourable, it can take a long time to reverse it. It is therefore vital that users feel supported through the early stages of adoption and that there are multiple channels available to obtain appropriate help. In particular floor walkers and a dedicated helpdesk hot line can greatly enhance user acceptability and the long-term embedding of technology into an organisation.[8]

The Patient's View

Since 2004, trusts have been obliged to copy patients into clinical correspondence unless the patient opts out of this. With the DH policy "The Power of Information"[9], we can increasingly expect more of the clinical record to become freely available to the patient. Turnaround of clinical documentation has been prone to hold-ups for many trusts with delays of months not being uncommon. The clinical risk of such waits can result in recipients being unaware of investigation and treatment plans with the possibility of patients receiving suboptimal treatment. With patients now recipients of this correspondence, the reputational damage of such delays can percolate through to patients who may choose to go to other organisations as they are offered greater choice in selecting healthcare options.

The obvious clinical safety benefits of timely letter production and availability are multiplied as patients with complex conditions are more commonly managed in the community with numerous clinicians caring for them.

[8] The Long and Winding Road…An Independent Evaluation of the Implementation and Adoption of the National Health Service Care Records Service (NHS CRS) in Secondary Care in England; Cresswell et al, 2011

[9] The Power of Information - http://informationstrategy.dh.gov.uk/about/the-strategy/

The Trust Management's View

NHS Trusts are operating under financial constraints, mirroring the national economic climate. Consequently, trusts are using cost improvement initiatives including QIPP to drive down costs whilst improving quality and delivery of care. In parallel, there have been significant changes to commissioning in NHS England from April 2013 through the introduction of the Health and Social Care Act 2012. Voice recognition has a part to play in delivering cost-effective, timely documentation to meet cost improvement targets and also revenue attached to CQUINs for delivery of quality documentation to commissioners within the commissioned timeframe. In the past, clinical documentation barely featured on the radar of senior trust executives. However, with the focus on clinical safety, minimising litigation and meeting CQUIN targets, the value of this documentation and the processes that underpin its creation are heading for the spotlight.

Opportunities for Coding

The opportunities for embedding speech recognition into structured documentation or into relevant fields in an EPR provide potential for clinicians to engage in upstream coding. By selecting clinical terms for diagnoses and procedures, the clinician will aid the clinical coders in their task and potentially improve revenue generation for the organisation through more accurate coding. This is particularly important with the inclusion of co-morbidities, which may result in the attainment of a more valuable Healthcare Resource Group (HRG – a classification used by healthcare providers to allocate resource costs to care delivered). This process becomes further enhanced with repeated visits (as is common in those patients with multiple co-morbidities) since the effort of

recording these terms on the first occasion will, in many cases, allow them to be easily repeated in the next documented event.

Speech recognition – the latest advances

Winscribe's most recent advances in speech recognition will revolutionise the deployment effort. Their latest solution is fully web based meaning that the software does not need to be installed and upgraded on individual PCs but can be accessed from any PC. This also makes the solution more scalable as further departments and clinicians become users.

Unlike older speech recognition platforms, it also enables users to be "grouped" together as a team so elements learned from the correction of transcribed text from one team member can be shared across all authors in the team. This feature makes the software quicker to start using – the new user will not need to spend a few hours preparing or priming the software as is the case with most speech recognition platforms. This feature is particularly important in areas where there is high staff turnover – particularly with regular new intakes of junior doctors. Users can also take their profile with them when they move from one organisation to another meaning they do not have to recreate their profile from scratch and the previous effort correcting transcribed data is preserved.

The application of speech recognition software in the production of structured clinical documentation such as clinic letters and discharge summaries will pave the way for the introduction of speech recognition in all elements of clinical record keeping. As users become more familiar with speech recognition – whether used through formal introduction via trust deployments or ad hoc use on their personal smart phones and tablets, clinicians will be demanding further interactive availability with greater use of voice command to move the cursor from field to field of structured data collection screens. Speech recognition has the power to crack one of the biggest dilemmas associated with electronic clinical records – the desire of clinicians to have well populated structured data in clinical records without the effort of manually entering it.