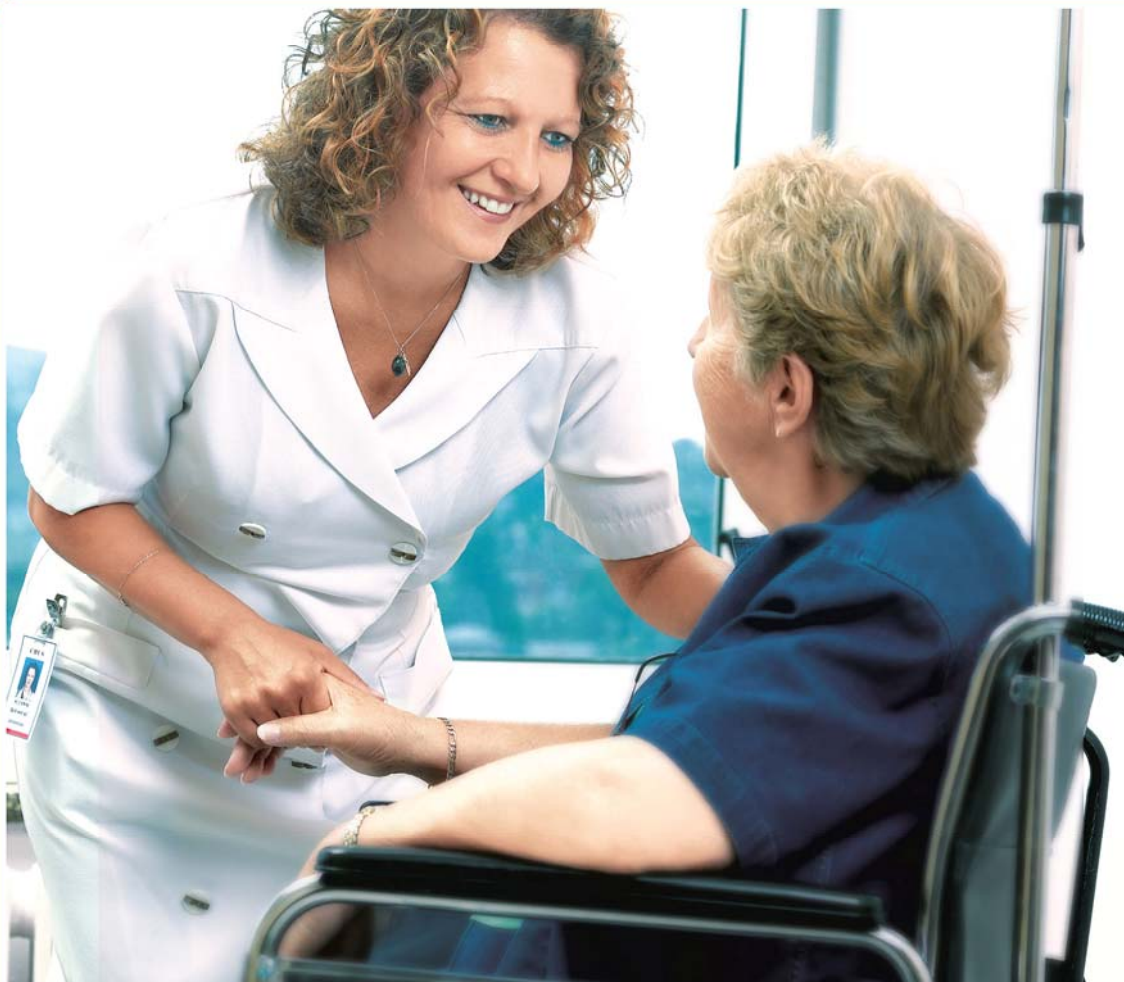


EFFICIENCE DE LA RECONNAISSANCE VOCALE POUR LE CHUS

UÉT MIS

UNITÉ D'ÉVALUATION DES
TECHNOLOGIES ET DES MODES
D'INTERVENTION EN SANTÉ



CHUS

Centre hospitalier
universitaire
de Sherbrooke

© Unité ETMIS 2008

www.chus.qc.ca

EFFICIENCE DE LA RECONNAISSANCE VOCALE POUR LE CHUS



Mai 2008

© Unité ÉTMIS 2008

LA MISSION

L'Unité d'évaluation des technologies et des modes d'intervention en santé (ÉTMIS) du Centre hospitalier universitaire de Sherbrooke (CHUS) est un regroupement d'experts dont les avis sont susceptibles d'influencer les décisions prises par l'administration hospitalière concernant les investissements en technologie de la santé, l'implantation des technologies émergentes, les changements dans la pratique des soins et les modes d'intervention en santé (dispensation des soins et organisation des services). Le créneau privilégié par le comité directeur de l'Unité ÉTMIS est « L'évaluation des pratiques et des modes d'intervention en santé ». Les évaluations tiennent compte de plusieurs volets, dont l'efficacité, la sécurité et l'efficience des technologies, ainsi que les impacts éthiques, légaux, sociaux et économiques liés à l'implantation et à l'administration desdites technologies. L'approche globale de l'Unité ÉTMIS est de développer l'évaluation des technologies en respectant les priorités établies dans la planification stratégique et les projets conjoints avec le Centre de recherche clinique Étienne-Le Bel du CHUS.

LA DIRECTION

Maurice Roy, M. D., LL. M., M. A. P.

Directeur général adjoint, directeur des Services professionnels et directeur médical à la Direction interdisciplinaire des services cliniques du CHUS

Renald Lemieux, M. Ing., Ph. D., M.ETS.

Directeur, Direction de l'évaluation des technologies et des modes d'intervention en santé, CHUS - Sherbrooke

Christian Bellemare, M. Sc.

Coordonnateur, Unité d'évaluation des technologies et des modes d'intervention en santé, CHUS – Sherbrooke

Pour tout renseignement sur ce document ou sur les activités de l'UETMIS-CHUS, veuillez vous adresser à :
Unité d'évaluation des technologies et des modes d'intervention en santé
Hôpital Fleurimont
3001, 12e Avenue Nord
Sherbrooke (Québec) J1H 5N4
Téléphone : (819) 346-1110 13802

© Unité ÉTMIS, Direction de l'évaluation des modes d'intervention et des technologies en santé CHUS 2008.

La reproduction totale ou partielle de ce document est autorisée à condition que la source soit mentionnée.

AVANT-PROPOS

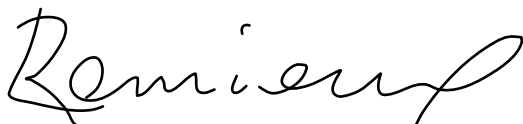
EFFICIENCE DE LA RECONNAISSANCE VOCALE POUR LE CHUS

Depuis longtemps déjà le CHUS cherche à gérer ses ressources avec efficacité et efficience. Dans le contexte actuel de forte pénurie de ressources, le CHUS se doit donc, plus que jamais, de renforcer sa culture d'innovation organisationnelle et se doter, de façon proactive, des outils nécessaires. Dans ce cadre, chaque idée nouvelle doit être confrontée aux données probantes de la littérature et le cas échéant faire l'objet d'un projet d'évaluation avant toute nouvelle implantation.

L'objectif de ce rapport est de répondre à ce besoin d'évaluation à un moment où le Service des archives du CHUS ne parvient plus à répondre aux demandes qui lui sont faites dans un délai raisonnable. Une des tâches du Service des archives est de procéder à la transcription sur fichier informatique des rapports médicaux des médecins à partir d'une bande sonore. Au cours de ces dernières années, la demande de transcription de rapports médicaux a fortement augmenté par rapport aux ressources humaines disponibles, ce qui a conduit à allonger le délai de mise en disponibilité de ces rapports médicaux sur le réseau informatique du CHUS.

Afin de revenir à des délais plus raisonnables et de réaliser de potentielles économies de coûts, il est envisagé de doter le CHUS d'un système de dictée numérique avec reconnaissance vocale pour la transcription des rapports médicaux. Le système de dictée numérique avec reconnaissance vocale est une technologie basée sur un modèle mathématique de reconnaissance des mots. Ce modèle mathématique calcule les probabilités d'occurrence des différentes formes d'ondes puis utilise une base de données des formes d'ondes connues afin de transcrire le mot parlé en un mot écrit. L'utilisation d'un tel système devrait en théorie permettre des gains de productivité pour le CHUS, soit en réduisant le nombre de rapports à transcrire par le Service des archives grâce à une utilisation interactive de ce système par les médecins, soit en réduisant le temps de transcription par rapport médical pour les transcripteurs du service des archives.

Sous cet angle, le Comité directeur des technologies de l'information, par l'entremise de Mr. Réjean Gouin, a mandaté l'unité ÉTMIS du CHUS pour évaluer les gains potentiels associés à la mise en place d'un système de dictée numérique avec reconnaissance vocale avant d'investir plus avant dans ce dossier. Le type d'évaluation demandée s'apparente ici à une analyse des données probantes tirées de la littérature scientifique.



Renald Lemieux, M. Ing., Ph. D., M.ETS.
Directeur, Direction de l'évaluation des technologies et
des modes d'intervention en santé, CHUS - Sherbrooke

REMERCIEMENTS

Ce rapport préliminaire a été préparé à la demande de la Direction des services de soutien de la technologie par **Thomas Poder**, Ph.D., conseiller en évaluation des technologies et par **Christian Bellemare**, M. Sc., coordonnateur de l'Unité ÉTMIS du CHUS.

Nous remercions par avance Mr. **Réjean Guin** du Comité directeur des technologies de l'information pour sa collaboration.

DIVULGATION DE CONFLIT D'INTÉRÊTS :

Aucun conflit à signaler.

RÉSUMÉ

Au cours de ces dernières années, les systèmes de reconnaissance vocale ont évolué à un point tel qu'ils seraient désormais devenus un moyen pratique de générer des rapports médicaux. De fait, de plus en plus de départements d'hôpitaux commencent à utiliser cette technologie et cette dernière est de plus en plus présente dans les foires commerciales, à la fois comme thème de séminaire et comme produit commercialisé.

L'objet de ce rapport est d'évaluer si cette technologie permet d'augmenter la productivité des hôpitaux dans la transcription des rapports médicaux par la réalisation d'une analyse systématique des données probantes de la littérature scientifique. À cet égard, trois critères sont étudiés : la performance du système de reconnaissance vocale, l'efficacité des ressources et les économies potentielles de coûts pour un centre hospitalier. Notre recherche conclue que le système de reconnaissance vocale 1) ne permet pas de diminuer le nombre d'erreurs de transcription par rapport à une transcription humaine ; 2) connaît une forte capacité d'amélioration après une première phase d'adaptation, mais que cette capacité diminue par la suite ; 3) le temps initial d'apprentissage pour les médecins est relativement court, toutefois, une phase d'entraînement de plusieurs semaines est nécessaire avant de parvenir à des résultats satisfaisants ; 4) le temps de production avec interaction du médecin semble être significativement plus long qu'avec le système traditionnel de transcription ; 5) les études sur le temps de production sans interaction du médecin ne sont pas en nombre suffisant pour parvenir à une conclusion fiable ; 6) avec interaction du médecin, l'efficacité du système de reconnaissance vocale n'est pas démontrée, bien au contraire ; 7) sans interaction du médecin, cette solution pourrait être efficace, toutefois, nous ne disposons pas de suffisamment d'éléments pour conclure.

Finalement, nous formulons deux recommandations. Notre première recommandation est de n'utiliser le système de reconnaissance vocale que dans les situations de pénurie de transcripteur ainsi que dans les situations où le médecin souhaite un rapport immédiatement disponible. Notre deuxième recommandation est de réaliser une étude pilote pour tester l'efficacité et l'efficacité de la reconnaissance vocale dans le contexte du CHUS. Si les résultats de cette étude pilote sont suffisamment positifs, une troisième recommandation serait de commencer à implanter le système de reconnaissance vocale dans les départements où la volonté est la plus forte, pour ensuite étendre le système à d'autres départements par effets de débordement.

TABLE DES MATIÈRES

LA MISSION	I
AVANT-PROPOS	III
REMERCIEMENTS	VII
RÉSUMÉ	IX
TABLE DES MATIÈRES	XI
CHAPITRE 1	1
1. PROBLÉMATIQUE	1
1.1 Présentation générale	1
1.2 Objectifs d'évaluation	1
CHAPITRE 2	3
2. RECENSION DES ÉCRITS	3
2.1 Méthodologie	3
2.2 Résultats	4
2.2.1 Études randomisées et études prospectives	5
2.2.1.1 Système de reconnaissance vocale avec interaction du médecin	6
2.2.1.2 Système de reconnaissance vocale sans interaction du médecin	9
2.2.2 Autres études	11
2.2.2.1 Système de reconnaissance vocale avec interaction du médecin	11
2.2.2.2 Système de reconnaissance vocale sans interaction du médecin	14
2.3 Discussion	14
2.4 Conclusion	17
CHAPITRE 3	19
3. RECOMMANDATIONS	19
BIBLIOGRAPHIE	21
Grille de synthèse	24

CHAPITRE 1

1. PROBLÉMATIQUE

1.1 Présentation générale

La région de l'Estrie avait élaboré en 2004 un projet régional d'implantation de la dictée numérique avec reconnaissance vocale dans ses centres hospitaliers. Un processus d'appel d'offres avait dès lors été lancé et un fournisseur avait été choisi. Il s'agissait du système G2 Speech de Philips distribué par la compagnie Comdic.

D'une part, la compagnie Comdic s'engageait à fournir au CHUS les licences requises de reconnaissance vocale et de transcription pour permettre la tenue d'un projet de recherche (10 licences « médecins » et 2 licences « secrétaires »). D'autre part, le CHUS s'engageait à évaluer et à publier les résultats du projet de recherche par l'entremise de son unité ETMIS. Cependant, le système n'a jamais été déployé compte tenu du coût d'interface nécessaire au transfert des textes vers le dossier clinique informatisé Ariane (évalué à 10K\$)¹.

Toutefois, l'installation d'un tel système pourrait avoir d'importantes répercussions sur l'efficacité des services. La méthode traditionnelle comporte les étapes suivantes : dicter le rapport au dictaphone, le transcrire par un spécialiste en transcription, le réviser

par le médecin et le signer. L'installation d'un système numérique pourrait considérablement modifier le mode de fonctionnement de certains départements. Par exemple, au département de radiologie, des automatismes pourraient être intégrés dans le processus, tels que : le démarrage de la dictée numérique lors de l'affichage des images radiologiques d'un patient sur le PACS, le transfert à une tierce personne (secrétaire ou autre) pour la vérification de la transcription du rapport, le transfert du rapport au dossier clinique informatisé, etc.

Sous cet angle, le Comité directeur des technologies de l'information, par l'entremise de Mr. Réjean Gouin, a mandaté l'unité ETMIS du CHUS pour évaluer les gains potentiels associés à la mise en place d'un système de reconnaissance vocale avant d'investir davantage dans ce dossier. Le type d'évaluation demandée s'apparente ici à une analyse des données probantes de la littérature scientifique.

1.2 Objectifs d'évaluation

L'objectif premier de cette évaluation est de discerner si les données en provenance de la littérature sont suffisantes afin d'émettre des conclusions et des recommandations en lien avec l'utilisation du système de

¹ Dollars canadiens de 2004.

dictée numérique avec reconnaissance vocale en milieu hospitalier.

Plusieurs sous-objectifs sont ensuite à évaluer :

1. La performance du système de reconnaissance vocale :

- la capacité à reconnaître ;
- la capacité d'apprentissage du système aux divers modes de lecture des utilisateurs (amélioration du dictionnaire, etc.) ;

➤ le temps d'apprentissage (courbe d'apprentissage du système et des utilisateurs) ;

➤ le temps de production du rapport médical (réduction du délai entre la dictée et la disponibilité du rapport final aux médecins).

2. L'efficacité des ressources :

➤ comparaison des coûts et de l'efficacité du système traditionnel avec ceux du système de reconnaissance vocale.

3. Les économies potentielles de coûts.

CHAPITRE 2

2. RECENSION DES ÉCRITS

2.1 Méthodologie

Les moteurs de recherche utilisés pour cette revue systématique sont Embase, CINALH, AMED, OVID Healthstar, OVID Medline, Mantis, Pubmed, Sciencedirect, British Medical Journal, Cochrane Database et Center for Reseach and Dissemination. Les sites web de l'Agence d'évaluation des technologies et des modes

d'intervention en santé et du NIHR Health Technology Assessment Programme ont également été consultés.

La recherche des articles se fait sur une période de référence allant de janvier 1992 jusqu'à mars 2008. Les langues utilisées pour cette recherche sont principalement l'anglais et le français.

Les mots clés utilisés dans les différents moteurs de recherche sont :

#1	SPEECH and RECOGNITION and REPORT	#9	VOICE and RECOGNITION and DICTATION
#2	SPEECH and RECOGNITION and RECORD	#10	VOICE and RECOGNITION and DICTATING
#3	SPEECH and RECOGNITION and SYSTEM	#11	VOCAL and RECOGNITION and REPORT
#4	SPEECH and RECOGNITION and DICTATION	#12	VOCAL and RECOGNITION and RECORD
#5	SPEECH and RECOGNITION and DICTATING	#13	VOCAL and RECOGNITION and SYSTEM
#6	VOICE and RECOGNITION and REPORT	#14	VOCAL and RECOGNITION and DICTATION
#7	VOICE and RECOGNITION and RECORD	#15	VOCAL and RECOGNITION and DICTATING
#8	VOICE and RECOGNITION and SYSTEM	#16	RECONNAISSANCE and VOCALE

Les résultats fournis par les moteurs de recherche et les deux sites web d'évaluation des technologies ont permis de lister plus de 1100 articles issus de revues possédant un comité de lecture². Parmi ces 1100 articles, un grand nombre ne correspondaient pas à

l'objet de notre recherche. Ainsi, tous les articles qui ne traitent pas de l'utilisation d'un système de reconnaissance vocale ayant pour but de retranscrire des rapports médicaux ont été écartés. Afin de déterminer quels articles étaient à écarter, nous avons lu les résumés de chacun de ceux-ci et discuté de leur adéquation avec l'objet de notre recherche. Le nombre d'article résultant de cette

² Aucun rapport réalisé par une agence d'évaluation n'a ici été trouvé.

première sélection est ainsi tombé à 52. Notre volonté étant de baser notre analyse et nos recommandations sur les données les plus probantes possibles, nous avons décidé, dans un premier temps, de nous focaliser sur les études randomisées ainsi que sur les études prospectives dans lesquelles les mêmes rapports sont utilisés pour tester les deux systèmes (traditionnel vs. reconnaissance vocale) ; soit les études dont le niveau de preuve selon la grille de Hailey et al. (2002) est supérieur ou égal à IV. Par ailleurs, la qualité de chacune de ces études a été évaluée au regard de la grille de Oxman et Guyatt (1993). Les études ne parvenant pas à un score minimal de six sur neuf selon la grille de Oxman et Guyatt (1993) ont ainsi été rétrogradées au même rang que les études dont le niveau de preuve est strictement inférieur à IV (Hailey et al., 2002). Les études ne satisfaisant pas à ce double critère ont ainsi été étudiées dans une section à part nommée « Autres études ».

2.2 Résultats

Cette revue de la littérature nous a permis de recenser deux grandes façons d'utiliser le système de reconnaissance vocale : soit le médecin agit en interaction avec celui-ci et réalise la dictée, la transcription et la correction avant de signer son rapport ; soit il n'y a aucune interaction et le médecin réalise la dictée puis effectue, avant signature, la révision de la transcription qui a été exécutée sous la supervision d'une tierce personne (secrétaire ou autre). La grande majorité des études que nous avons ici analysées se sont avérées être

réalisées dans le cadre d'une interaction entre le médecin et le système de reconnaissance vocale. Ainsi, seules six études se focalisent de façon spécifique sur l'utilisation d'un système de reconnaissance vocale au sein d'un département de transcription des rapports médicaux. La Figure 1, à la page suivante, décrit les principaux modes de transcription des rapports médicaux au sein d'un hôpital.

Par ailleurs, cette revue de la littérature indique que presque toutes les études que nous avons analysées ont été réalisées à partir d'un système de reconnaissance vocale en continu (i.e. pas de pause « marquée » entre les mots). En effet, seules les études de Threet et Fargues (1999) et de Smith (2002) se sont focalisées sur l'analyse des avantages et des inconvénients d'un système de reconnaissance vocale en discret.

Finalement, il est ici important de noter que les études que nous avons analysées ont utilisé des systèmes de reconnaissance vocale de différentes langues, telles que l'allemand, l'anglais, le chinois, le français, l'italien et le japonais.

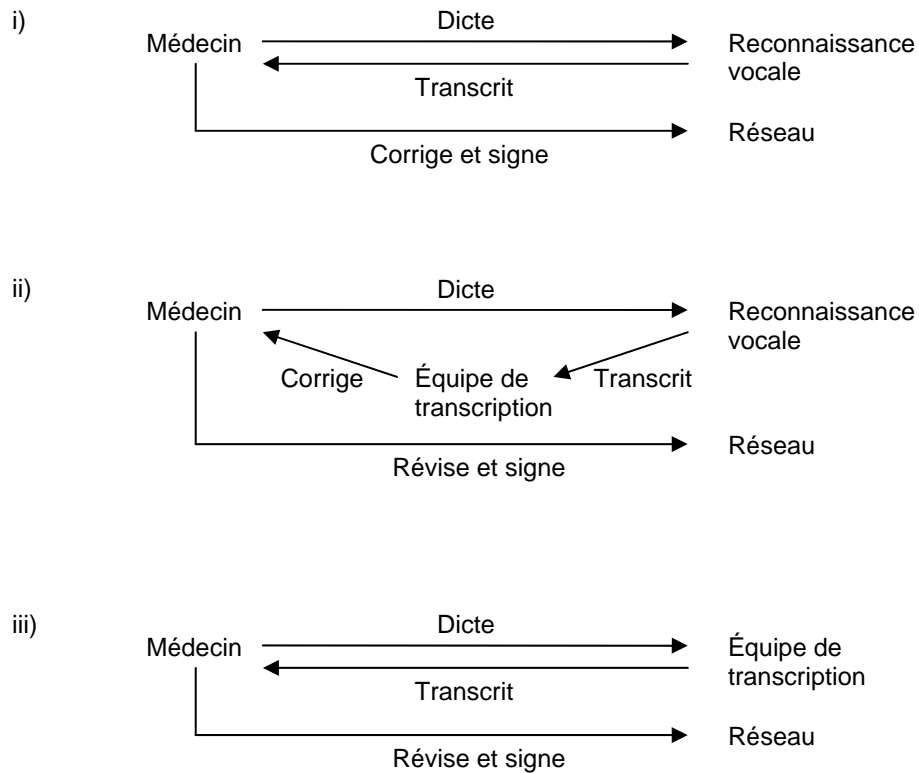


Figure 1 : Principaux modes de transcription des rapports médicaux au sein d'un hôpital : i) le médecin interagit avec un système de reconnaissance vocale ; ii) le médecin n'interagit pas avec un système de reconnaissance vocale ; iii) système traditionnel sans système de reconnaissance vocale.

2.2.1 Études randomisées et études prospectives

L'analyse détaillée de notre présélection de 52 études nous indique un total de 23 études ayant procédé, soit à une répartition aléatoire des rapports médicaux entre les deux systèmes que l'on cherche à comparer (i.e. système traditionnel de transcription utilisé au sein de l'hôpital vs. système de transcription avec reconnaissance vocale), soit à une comparaison des deux systèmes avec l'utilisation des mêmes rapports médicaux.

Dans cette section, nous allons dans un premier temps nous focaliser sur les résultats d'études décrivant une interaction entre le médecin et le système de reconnaissance vocale, et dans un second temps, nous focaliser sur les résultats d'études où il n'existe aucune interaction entre le médecin et le système de reconnaissance vocale.

2.2.1.1 Système de reconnaissance vocale avec interaction du médecin

Capacité à reconnaître

Dans le but d'harmoniser les résultats de notre recherche, dans cette section et dans les sections à suivre, la notion de taux « global » ou de performance « globale » du système de reconnaissance vocale vis-à-vis du nombre d'erreurs de transcription fera référence à l'ensemble des différents types d'erreurs de transcription pouvant être commises (i.e. omission, substitution, insertion de mots, orthographe et grammaire). Lorsque certains types d'erreurs spécifiques seront abordés, cela sera mentionné comme tel dans notre travail.

Par rapport au système traditionnel de transcription de la dictée du médecin par une secrétaire ou par du personnel spécialisé dans la transcription, le système de dictée numérique avec reconnaissance vocale est en règle générale moins performant pour ce qui a trait au nombre d'erreurs de transcription. Après une phase de « rodage » de une à plusieurs semaines du système de reconnaissance vocale, sa performance globale est comprise entre 85 et 98,5% de reconnaissance exacte avec une moyenne globale calculée de 93,9% (Al-Aynati et Chorneyko, 2003 ; Capel et al., 2004 ; Devine et al., 2000 ; Happe et al., 2003 ; Hundt et al., 1999 ; Ichikawa et al., 2005 ; Issenman et Jaffer, 2004 ; Kanal et al., 2001 ; Lui et al., 2001 ; Pezzullo et al.,

2008 ; Rana et al., 2005 ; Takahara et al., 2002 ; Teichgräber et al., 1999 ; Zick et Olsen, 2001)³. En comparaison, le taux global de reconnaissance exacte d'une transcription humaine est compris entre 97 et 99,97% avec une moyenne globale calculée de 99,2% (Al-Aynati et Chorneyko, 2003 ; Capel et al., 2004 ; Issenman et Jaffer, 2004 ; Pezzullo et al., 2008 ; Teichgräber et al., 1999 ; Zick et Olsen, 2001).

Les erreurs les plus fréquentes liées au système de reconnaissance vocale sont dues à des substitutions ou à des insertions inadéquates de mots (Hundt et al., 1999 ; Teichgräber et al., 1999). Le système de reconnaissance vocale est également responsable d'un plus grand nombre d'erreurs majeures pouvant conduire à une interprétation erronée du rapport médical ; ce problème est toutefois beaucoup plus fréquent parmi les médecins ayant peu d'expérience du système de reconnaissance vocale (Rana et al., 2005). À l'inverse, en ce qui concerne les fautes d'orthographe et de grammaire, le système de reconnaissance vocale permet d'atteindre de meilleurs résultats (La Fianza et al., 1993 ; Teichgräber et al., 1999). On peut aussi remarquer des différences selon la personne qui utilise le système de reconnaissance vocale : les résultats sont ainsi moins bons si le médecin utilise une langue qui n'est pas sa langue maternelle (Kanal et al., 2001) ; de même pour les médecins résidents qui ne sont pas encore très habitués à faire des dictées (Capel et al., 2004) ou pour les médecins ayant

³ On a ici considéré les taux les plus élevés fournis par ces études.

moins d'expérience avec les systèmes de reconnaissance vocale (Rana et al., 2005 ; Zick et Olsen, 2001). Finalement, après correction par le médecin de la transcription effectuée par un système de reconnaissance vocale, les rapports médicaux sont considérés dans certaines études comme étant globalement plus lisibles et plus précis pour les autres médecins que les rapports réalisés par transcription traditionnelle (La Fianza et al., 1993 ; Threet et Fargues, 1999).

Capacité d'apprentissage du système

Le plus souvent, il est demandé lors de la phase d'initialisation du système de reconnaissance vocale de télécharger un certain nombre d'anciens rapports médicaux sous format Microsoft Word ou texte afin de lui permettre d'enrichir son vocabulaire et sa syntaxe (Capel et al., 2004 ; Takahara et al., 2002 ; Threet et Fargues, 1999). Le système de reconnaissance vocale demande par la suite au médecin de lire un certain nombre de phrases afin d'associer la voix du médecin à tous les mots utilisés. De plus, la plupart des systèmes de reconnaissance vocale sont conçus de manière à ce que les corrections réalisées par le médecin soient enregistrées et associées aux phrases prononcées par ce dernier afin de ne plus répéter ces erreurs (Hundt et al., 1999 ; Threet et Fargues, 1999).

La plupart des études indiquent ainsi une réelle capacité d'apprentissage des systèmes de reconnaissance vocale. Toutefois, très peu d'études se sont attachées à évaluer cette capacité d'apprentissage.

À cet égard, nos recherches indiquent qu'après plusieurs adaptations, le système de reconnaissance vocale peut connaître une très forte amélioration de son taux global de reconnaissance (de 79,6 à 92,5% après 3 adaptations réalisées en 7 heures ; Hundt et al., 1999 ; de 84,5-86,7% à 95,7% après 4 adaptations ; Lui et al., 2001). L'étude de Happe et al. (2003) montre cependant que la capacité d'apprentissage de ce système est étroitement liée au pourcentage d'erreurs initiales (i.e. suite à la phase d'initialisation du système). Cette étude montre ainsi que le système IBM ViaVoice 8, dans sa version médicale, connaît une très faible capacité d'amélioration de son taux global de reconnaissance exacte en raison de son bon score initial de 98,3%, alors que le système Naturally Speaking 5, dans sa version médicale, connaît une progression de son taux global de reconnaissance exacte de 94,3 à 95,8%.

Période d'apprentissage

La période d'apprentissage pour les médecins est dans la phase initiale de formation comprise entre 30 minutes et 4 heures (Al-Aynati et Chorneyko, 2003 ; Borowitz, 2001 ; Capel et al., 2004 ; Issenman et Jaffer, 2004 ; Takahara et al., 2002 ; Zick et Olsen, 2001). Par la suite, les médecins doivent s'entraîner un certain nombre de semaines avant de pouvoir parvenir à un taux global de reconnaissance vocale exacte qui soit considéré comme satisfaisant pour eux (Issenman et Jaffer, 2004). Cette période d'entraînement peut s'avérer trop longue pour les médecins ayant des

difficultés d'adaptation avec le système (i.e. nécessité de bien prononcer les mots, de faire des phrases synthétiques, de placer convenablement le microphone, etc.) (Issenman et Jaffer, 2004 ; Pezzullo et al., 2008). On doit cependant noter que la plupart des médecins s'adaptent relativement rapidement à ce type de système (Capel et al., 2004 ; Lui et al., 2001 ; Takahara et al., 2002 ; Zick et Olsen, 2001).

Temps de production

On doit ici distinguer deux types de mesures : le temps de réalisation du rapport médical (i.e. délais de transmission non comptabilisés) et le temps de mise en disponibilité du rapport médical sur le réseau informatique de l'hôpital.

En ce qui concerne le temps de réalisation du rapport médical, les résultats de la plupart des études avec interaction des médecins indiquent, comparativement au système traditionnel, une dépense supplémentaire en temps de l'ordre de 16 à 190% pour les médecins (Al-Aynati et Chorneyko, 2003 ; Borowitz, 1999 ; Borowitz, 2001 ; Gale et al., 2001 ; Hundt et al., 1999 ; Issenman et Jaffer, 2004 ; Lui et al., 2001 ; Pezzullo et al., 2008 ; Rana et al., 2005 ; Teichgräber et al., 1999). Cette différence de temps de réalisation s'explique principalement par un temps de correction plus important (Al-Aynati et Chorneyko, 2003 ; Borowitz, 1999 ; Borowitz, 2001 ; Gale et al., 2001 ; Hundt et al., 1999 ; Issenman et Jaffer, 2004), et dans certains cas par un temps de dictée légèrement plus long (Borowitz, 1999 ; Borowitz, 2001 ; Hundt et al., 1999). On doit cependant noter que certaines

études obtiennent des résultats contraires avec des gains en temps de 3 à 61% (Capel et al., 2004 ; La Fianza et al., 1993 ; Takahara et al., 2002 ; Trumm et al., 2006 ; Zick et Olsen, 2001). Toutefois, à l'exception de l'étude de La Fianza et al. (1993) et de Trumm et al. (2006), ces études ont été réalisées avec un nombre de rapports médicaux inférieur à 50. Dans un autre registre, l'étude de Threet et Fargues (1999), qui compare les temps de réalisation des rapports médicaux entre système de reconnaissance vocale et rédaction à la main sur papier ou sur ordinateur faite par le médecin, indique que le système de reconnaissance vocale obtient le deuxième meilleur temps derrière la rédaction à la main sur papier.

Pour ce qui est du temps de mise en disponibilité du rapport médical sur le réseau de l'hôpital, compte tenu du fait que l'interaction avec le médecin élimine les délais de transmission et de prise en charge de la dictée vers le service de transcription, les résultats des études indiquent tous un gain de temps considérable compris entre 53,6 et 99,5% (Borowitz, 1999 ; Borowitz, 2001 ; Hundt et al., 1999 ; Pezzullo et al., 2008 ; Takahara et al., 2002 ; Teichgräber et al., 1999 ; Trumm et al., 2006 ; Zick et Olsen, 2001).

Efficience

Les rares études ayant réalisées un calcul de retour sur investissement indiquent que l'achat et l'implantation d'un système de reconnaissance vocale peut être amorti entre quelques semaines ou mois (Capel et al., 2004 ; Zick et Olsen, 2001) ou au contraire coûter deux fois plus cher que le système traditionnel (Issenman et

Jaffer, 2004 ; Pezzullo et al., 2008). Nous devons cependant ici insister sur le fait que ces calculs de retour sur investissement sont rarement détaillés et que les seules études qui indiquent prendre en compte comme un coût le temps supplémentaire passé par les médecins avec le système de reconnaissance vocale sont celles de Issenman et Jaffer (2004) et de Pezzullo et al. (2008), ces dernières indiquant un coût par rapport médical deux fois plus élevé qu'avec un système de transcription humain. Toutefois, certains bénéfices du système de reconnaissance vocale n'ont pas été évalués, notamment les bénéfices associés à la réduction du temps de mise en disponibilité du rapport médical sur le réseau informatique de l'hôpital (i.e. permet aux médecins d'être payés plus rapidement, de déclencher des traitements plus tôt, etc.) (Kopach et al., 2005).

Économies potentielles

L'étude de Gale et al. (2001) indique une économie en secrétaire de 100.000 USD par an, mais une plus faible productivité du médecin, et celle de Rana et al. (2005) une économie en secrétaire de 20.000 £ par an. De la même manière que dans le paragraphe précédent, des éléments importants ne sont pas pris en considération, notamment le temps du médecin, ce qui empêche de déterminer s'il existe des économies potentielles de coûts pouvant être réalisées avec le système de reconnaissance vocale.

Autres éléments d'intérêt

Dans l'étude de Threet et Fargues (1999), le fait que le médecin

utilise un système de reconnaissance vocale au lieu de faire son rapport sur clavier ou sur papier ne change en rien de façon significative, ni la qualité de l'évaluation du patient, ni la qualité d'écoute du médecin. Par ailleurs 60% des patients faisant partie du groupe avec reconnaissance vocale indiquent que la dictée du médecin permet de mieux comprendre la nature de leur problème. En outre, selon cette même étude, le système de reconnaissance vocale permettrait de réaliser des rapports plus complets. Ce dernier résultat est toutefois à nuancer par les résultats de certaines études indiquant que le plus grand nombre d'erreurs de transcription générées par le système de reconnaissance vocale peut conduire certains médecins à réduire la taille de leurs rapports afin de minimiser leur temps de correction (Pezzullo et al., 2008).

2.2.1.2 Système de reconnaissance vocale sans interaction du médecin

Dans cette sous-section, nous décrivons les résultats d'études où la dictée du médecin est transmise au service de transcription de l'hôpital où une transcription avec un système de reconnaissance vocale est réalisée.

Capacité à reconnaître

Le taux global de reconnaissance sans erreurs du système de reconnaissance vocale est dans ces études compris entre 84,5 et 94,5% (Mohr et al., 2003 ; Vorbeck et al., 2000 ; Zimmel et al., 1996), contre 99,6%

avec la transcription humaine (Vorbeck et al., 2000).

Capacité d'apprentissage du système

Dans l'étude de Morh et al. (2003), le taux global de reconnaissance sans erreurs du système de reconnaissance vocale ne progresse quasiment plus après 4 semaines d'entraînement et reste, toutes erreurs confondues, près de 84,5%.

Période d'apprentissage

L'étude de Zimmel et al. (1996) indique que les médecins qui feront la dictée doivent lire 270 phrases en 1 heure et 30 minutes, puis effectuer une période d'entraînement de 3 heures 30 minutes. De son côté, l'étude de Morh et al. (2003) indique que les secrétaires et les transcripateurs ont reçu 6 heures de formation afin d'optimiser leur utilisation du système de reconnaissance vocale.

Temps de production

Les études montrent, de la même manière que dans la section 2.2.1.1., des résultats contradictoires. D'un côté, nous avons une étude montrant que l'utilisation d'un système de reconnaissance vocale conduit à une baisse de la productivité du service de transcription comprise entre 12,7 et 44,2% (Morh et al., 2003). D'un autre côté, les études de Trumm et al. (2006) et de Vorbeck et al. (2000) indiquent que le système de reconnaissance vocale permet une augmentation moyenne de la productivité du service

de transcription comprise entre 19 et 23%. Ce taux pouvant même progresser jusqu'à 28% dans certaines spécialités médicales comme la sonographie (Vorbeck et al., 2000). Ces différentes études indiquent également que la productivité du service de transcription liée à l'utilisation d'un système de reconnaissance vocale est plus forte quand les médecins sont habitués à la dictée et que les secrétaires et transcripateurs sont relativement « lents » avec la méthode traditionnelle de transcription à l'oreille.

L'étude de Trumm et al. (2006) indique également que la hausse de la productivité du service de transcription liée à l'utilisation d'un système de reconnaissance vocale permet une importante amélioration du temps de mise en disponibilité du rapport médical sur le réseau de l'hôpital (de 10,47 heures à 6,65 heures).

Efficiences

Dès lors qu'il existera un nombre suffisamment important d'études démontrant que le système de reconnaissance vocale permet d'améliorer la productivité du service de transcription, son utilisation sera efficace car elle permettra de réduire les temps de mise en disponibilité des rapports médicaux dans le réseau de l'hôpital, de transcrire davantage de rapports avec le même nombre de transcripateurs, et n'aura aucun impact sur le temps de correction des médecins. Pour l'instant, nous n'avons malheureusement que trop peu d'éléments pour conclure.

Économies potentielles

Aucune information substantielle.

Autres éléments d'intérêt

Certaines spécialités, telle que l'urgence, semble être peu appropriées à l'utilisation d'un système de reconnaissance vocale en raison de l'utilisation d'un vocabulaire assez large, du positionnement du microphone qui peut être davantage négligé, ainsi que de l'importance des bruits de fonds (Zemmel et al., 1996).

2.2.2 Autres études

Compte tenu du nombre assez limité d'études satisfaisant à un haut niveau de qualité (i.e. sélection aléatoire et/ou utilisation des mêmes rapports pour comparer les différents systèmes de transcription) et de la présence de résultats parfois opposés, nous décidons dans cette section de présenter un bref aperçu des résultats de 29 autres études. Bien que ces 29 études ne satisfassent pas à nos critères de qualité, il est toutefois intéressant de vérifier si leurs résultats sont similaires à ceux des études présentées précédemment. De la même manière que pour les études satisfaisant à un haut niveau de qualité, nous allons ici distinguer les études décrivant une interaction entre le médecin et le système de reconnaissance vocale, des études où il n'existe aucune interaction entre le médecin et le système de reconnaissance vocale.

2.2.2.1 Système de reconnaissance vocale avec interaction du médecin

Capacité à reconnaître

Le taux global de reconnaissance exacte du système de reconnaissance vocale est ici compris entre 70 et 100% (Aprile et al., 1997 ; Arndt et al., 1999 ; Corces et al., 2004 ; Dietz et al., 1996 ; Haug et Booth, 1999 ; Henricks et al., 2002 ; Herman, 2004 ; Langer, 2002a ; Ramaswamy et al., 2000 ; Simmons, 2002 ; Ury, 2007 ; Wheeler et Cassimus, 1999 ; Zafar et al., 1999)⁴, avec une moyenne globale de 94,4%. On peut toutefois faire ici remarquer que l'utilisation de macros et de « gabarits » permet une amélioration significative du taux global de reconnaissance exacte. L'étude de Corces et al. (2004) indique ainsi que l'ajout de « gabarits » permet en moyenne de faire progresser ce taux global de 90 à 96%. Par ailleurs, si le système de reconnaissance vocale est programmé pour ne reconnaître que certains mots clés, le taux global de reconnaissance exacte peut atteindre les 99% (Chang et al., 2008).

Tout comme dans les études satisfaisant à notre critère de qualité, les erreurs les plus communes liées au système de reconnaissance vocale sont les omissions et les substitutions de mots (Ramaswamy et al., 2000 ; Rosenthal et al., 1998). Par contre, le système de reconnaissance vocale ne créerait selon Ramaswamy et al. (2000)

⁴ On a ici considéré les taux les plus élevés fournis par ces études.

que très peu d'erreurs majeures. En outre, le système de reconnaissance vocale permettrait de diviser par plus de trois le nombre d'erreurs d'orthographe par rapport à la transcription humaine (Ramaswamy et al., 2000). Par ailleurs, Rosenthal et al. (1998) indiquent également que le système de reconnaissance vocale commet, toutes catégories d'erreurs confondues, un plus grand nombre d'erreurs avec les médecins utilisant une langue de dictée différente de leur langue maternelle.

Capacité d'apprentissage du système

Les études n'ayant pas atteint notre standard de qualité indiquent elles aussi une capacité du système à progresser. Ainsi, l'étude de Arndt et al. (1999) indique un taux global de reconnaissance exacte progressant en moyenne de 89,2 à 93,5% après une première adaptation du système ayant eu lieu au 9ème jour suivant sa mise en place. Toutefois, cette même étude indique que le taux global de reconnaissance exacte ne progresse qu'assez peu avec les adaptations suivantes. D'un autre côté, l'étude de Zafar et al. (1999) mentionne que des progrès de 5% après chaque deux semaines peuvent être réalisés avec la correction en continue du médecin.

Période d'apprentissage

Le temps d'apprentissage des médecins est, dans la phase initiale de formation, comprise entre 5 minutes et 3 heures (Arndt et al., 1999 ; Bundt, 2002 ; Chang et al., 2008 ; Corces et al., 2004 ; Henricks et al., 2002 ; Korn, 1998 ; Ramaswamy et al., 2000 ; Ury, 2007 ;

Zafar et al, 1999). Par la suite, les médecins doivent poursuivre leur entraînement sur une à plusieurs semaines avant de pouvoir parvenir à un taux global de reconnaissance vocale exacte qui soit considéré comme étant satisfaisant par eux (Corces et al., 2004 ; Rosenthal et al., 1998 ; Simmons, 2002 ; Smith, 2002). Sous un autre angle, Antiles et al. (2000) indiquent que la durée totale de formation de 127 médecins et résidents en radiologie s'est faite sur 6 mois, alors que Sferrella (2003) mentionne que cette durée totale de formation a nécessité seulement 2 semaines pour 50 médecins en radiologie. Les informations fournies par ces deux dernières études sont cependant insuffisantes pour en déduire le temps de formation effectif des médecins.

Temps de production

Par rapport à un service de transcription humaine, le système de reconnaissance vocale ne semble pas permettre de réduire le temps de réalisation des rapports médicaux, mais au contraire l'allonger (Dietz et al., 1996 ; Haug et Booth, 1999 ; Smith, 2002). De fait, selon Hayt et Alexander (2001), les médecins passeraient en moyenne 25% de temps en plus pour faire leurs rapports médicaux. Par contre, le temps de mise en disponibilité des rapports médicaux sur le réseau de l'hôpital serait très fortement réduit grâce à l'implantation d'un système de reconnaissance vocale. Ce temps de mise en disponibilité serait ainsi, en moyenne, réduit de 49 à 99% (Antiles et al., 2000 ; Callaway et al., 2002 ; Hayt et Alexander, 2001 ; Henricks et al., 2002 ; Langer, 2002b ; Lemme et Morin, 2000 ;

Ramaswamy et al., 2000 ; Rosenthal et al., 1998 ; Sferrella, 2003 ; Simmons, 2002 ; Smith, 2002 ; Ury, 2007 ; Wheeler et Cassimus, 1999).

Efficiences

La modélisation budgétaire de Bundt et al. (2002) indique, par rapport à une transcription humaine réalisée au sein d'un service de transcription interne à l'hôpital, que la mise en place d'un système de reconnaissance vocale permettrait une économie de coût de l'ordre de 18%.⁵ Par contre, si la transcription humaine est réalisée par une compagnie de sous-traitance, la mise en place d'un système de reconnaissance vocale représenterait alors un surcoût de 33% (Bundt et al., 2002).

La simulation budgétaire de Kopach et al. (2005) suggère quant à elle que le coût supplémentaire pour réduire d'une journée le délai de mise en disponibilité des rapports médicaux sur le réseau de l'hôpital grâce à l'utilisation d'un système de reconnaissance vocale serait de 0,33 CAD par rapport médical.

Économies potentielles

Dans l'étude de Antiles et al. (2000), l'implantation de 60 postes de système de reconnaissance vocale ainsi que la formation de 127 médecins et résidents en radiologie, aurait permis de

⁵ Il est ici fait l'hypothèse que le médecin passe 3 fois plus de temps à réalisation du rapport médical avec le système de reconnaissance vocale qu'avec un service de transcription humain. Le salaire du médecin est inclus dans les coûts.

réaliser des économies de 530.000 USD par an. De même, selon l'étude de Callaway et al. (2002), la mise en place d'un système de reconnaissance vocale aurait permis une économie de 2,08 USD par rapport médical. De leur côté, Henricks et al. (2002) mentionnent un retour sur investissement de l'implantation d'un système de reconnaissance vocale au bout de 1,9 années, alors que les études de Ramaswamy et al. (2000), de Rosenthal et al. (1998) et de Sferrella (2003) indiquent un retour sur investissement en moins d'un an. De nombreuses autres études indiquent des économies potentielles de coûts (Corces et al., 2004 ; Herman, 2004 ; Simmons, 2002), toutefois, aucune des études présentées dans ce paragraphe n'intègre dans les coûts du système de reconnaissance vocale le temps supplémentaire que le médecin consacre à ce système par rapport à une situation où il utiliserait un service de transcription humaine, ce qui empêche tout calcul réaliste.

Autres éléments d'intérêt

À l'instar de l'étude de Pezzullo et al. (2008), celle de Ramaswamy et al. (2000) indique un changement de comportement de la part de certains médecins faisant face aux erreurs de transcription du système de reconnaissance vocale : ceux-ci se retournent volontairement à établir des rapports plus courts (en moyenne de 37%) afin de limiter leur temps de correction.

2.2.2.2 Système de reconnaissance vocale sans interaction du médecin

Capacité à reconnaître

Les études de Briggs (2004) et de Van Terheyden (2005) indiquent chacune un taux global de reconnaissance sans erreurs de 95% pour le système de reconnaissance vocale.

Capacité d'apprentissage du système

Aucune information substantielle.

Période d'apprentissage

Aucune information substantielle.

Temps de production

Toutes les études indiquent ici des gains de productivité liés à l'utilisation d'un système de reconnaissance vocale. Ces gains seraient compris entre 24 et 30% (Briggs, 2004 ; Spring, 2003 ; Van Terheyden, 2005). Dans l'étude de Briggs (2004), le gain de productivité varie de 5 à 53% selon la rapidité d'exécution des transpositeurs avant la mise en place du nouveau système. De fait, les transpositeurs sont désormais davantage des correcteurs que des transpositeurs.

Efficiences

De la même manière que dans le paragraphe précédent, l'efficiences devrait être comprise entre 24 et 30% (Briggs, 2004 ; Spring, 2003 ; Van Terheyden, 2005).

Économies potentielles

Ibid.

Autres éléments d'intérêt

Aucune information substantielle.

2.3 Discussion

La réalisation de rapports médicaux connaît actuellement une très nette tendance à la hausse (Van Terheyden, 2005). Cette forte demande est alimentée par l'accroissement d'une population vieillissante nécessitant un grand nombre de tests et traitements médicaux, et par le besoin continu des différents acteurs du système de santé de partager ces informations médicales (médecins référents, organismes de remboursements, etc.). L'offre de transcription humaine ne sera bientôt plus suffisante pour couvrir ce flux, ce qui démontre le besoin de faire appel à des technologies de substitution telles que les systèmes de reconnaissance vocale.

Les systèmes de reconnaissance vocale sont basés sur un modèle mathématique de reconnaissance des mots. Ce modèle mathématique calcule les probabilités d'occurrence des

différentes formes d'ondes puis utilise une base de données des formes d'ondes connues afin de transcrire le mot parlé en un mot écrit.

Jusqu'au milieu des années 90, la littérature scientifique sur les systèmes de reconnaissance vocale rendait essentiellement compte de versions anciennes qui ne permettaient pas une dictée fluide. L'ergonomie et le vocabulaire de ces versions étaient limités. Depuis, d'immenses progrès ont été réalisés et les versions les plus récentes de ces logiciels permettent une reconnaissance bien plus fiable sur un vocabulaire très vaste. Par exemple, pour un logiciel de reconnaissance vocale avec un vocabulaire général, le nombre de mots et locutions disponible peut aller jusqu'à 265 000, avec de très nombreux termes médicaux (Capel et al., 2004). De plus, certains logiciels comme celui fourni par IBM (ViaVoice Pro Edition 8) peuvent être enrichis jusqu'à 65 000 mots supplémentaires par procédure de lecture automatique de rapports déjà existants. En outre, il est possible d'augmenter la précision et la rapidité du système en reliant celui-ci avec un système de fichiers Text ou Word pré-remplis (gabarit) (Corces et al., 2004 ; Henricks et al., 2002). Finalement, la plupart des systèmes de reconnaissance vocale seraient actuellement capables de distinguer ce qui est un éternuement ou une toux de ce qui est un mot, réduisant ainsi encore plus les risques d'erreurs (Smith, 2002 ; Spring, 2003). En ce qui concerne le matériel périphérique nécessaire au bon fonctionnement du système de reconnaissance vocale et dont la qualité faisait auparavant défaut, d'importants progrès ont également été réalisés. Des microphones et des cartes de son

adaptés sont désormais disponibles. De plus, l'augmentation de la mémoire vive des nouveaux postes informatiques est devenue un atout indispensable.

Malgré ces progrès, obtenir un bon résultat nécessite toujours de respecter un certain nombre de contraintes auxquelles il est très difficile de se soustraire. De nombreux auteurs indiquent ainsi la nécessité de réaliser des dictées à un rythme qui ne soit pas trop rapide et de parler avec un maximum d'articulation (Corces et al., 2004 ; Pezzullo et al., 2008 ; Zafar et al., 1999). En plus de cela, obtenir un bon taux global de reconnaissance vocale nécessite un positionnement adéquat du microphone et des bruits de fonds limités (Zemmel et al., 1996). Par ailleurs, les utilisateurs potentiels de ce type de technologie ne sont pas tous prêts à réaliser ces différents efforts. Kopach et al. (2005) mentionnent ainsi que les compagnies produisant les systèmes de reconnaissance vocale estiment que près de 50% des médecins ne seraient pas aptes à les utiliser. De surcroît, on remarque que le résultat d'une transcription par un système de reconnaissance vocale est une transcription littérale, ce qui peut poser problème quand les dictées des médecins ne sont pas structurées pour correspondre à un rapport final. De fait, les informations peuvent être retranscrites de façon non linéaire, ou encore sans ponctuation, ce qui nécessite de réaliser un ajustement par le transcripteur ou le médecin et conduit finalement à une réduction des gains potentiels de productivité. En outre, si des mots sont mal corrigés par le transcripteur ou le médecin, ces mots seront enregistrés dans le dictionnaire du système de reconnaissance vocale

et apparaîtrons par la suite comme des mots exacts dans les futures transcriptions.

Avec un système de reconnaissance vocale, arriver à un résultat sans erreurs de transcription peut donc s'avérer être long et coûteux (Pezzullo et al, 2008). Par ailleurs, ne procéder à aucune correction ou à une correction minimale n'est pas acceptable. En effet, même si les erreurs réalisées par ce système s'avèrent être rarement en mesure de conduire à une mauvaise interprétation de la part de tierces médecins, laisser des erreurs mineures dans le rapport peut être considéré comme inacceptable dans la mesure où ce rapport est le premier moyen de communication entre le médecin et la communauté et que cela pourrait être interprété comme un signe de travail mal fait ou d'un manque d'intérêt de la part du médecin. En outre, d'un point de vue médico-légal, chaque rapport médical doit être vérifié en détail par son auteur ou un correcteur médical humain. Par contre, si l'on compare les temps de dictée et de correction du médecin dans une situation où le médecin a recours à un service de transcription humaine avec une situation où le médecin interagit avec un système de reconnaissance vocale, cette obligation médico-légale conduit dans le second cas à des temps de travail supplémentaires bien plus importants pour les médecins. Par conséquent, si un système de reconnaissance vocale avec interaction du médecin est mis en place, le temps de travail supplémentaire passé par le médecin doit être compensé (pas de patient(s) examiné(s) pendant ce temps là) afin d'inciter ce dernier à utiliser le système.

En effet, les médecins ne devraient pas avoir à supporter seuls la double responsabilité (physique et financière) du service de transcription.

Une solution pour éviter aux médecins d'avoir à dépenser du temps supplémentaire dans le processus de transcription des rapports médicaux serait de confier la charge de corriger les erreurs de transcription du système de reconnaissance vocale à des secrétaires ou aux membres du service de transcription de l'hôpital. Sachant qu'une minute de dictée équivaut à environ 2-4 minutes de transcription humaine, d'important gains de productivité pourraient ainsi être réalisés grâce à un système de reconnaissance vocale (Spring, 2003 ; Vorbeck et al., 2000 ; Van Terheyden, 2005). Toutefois, si on demande à une secrétaire ou à un transcrivoteur de corriger le texte créé par un système de reconnaissance vocale, cela n'est plus le même métier, ce qui n'est pas forcément une transition aisée à réaliser (Spring, 2003).

En dépit de ces différents inconvénients, les systèmes de reconnaissance peuvent néanmoins se révéler être un outil utile à la production des rapports médicaux. En effet, avec un service traditionnel de transcription, les délais de retour des rapports aux médecins pour révision et validation peuvent être très longs. Dans de tels cas, le médecin ou professionnel de santé peut ne plus se rappeler des détails du cas examiné, ce qui implique pour ces derniers de se « replonger » dans le dossier du patient et parfois à réaliser une nouvelle analyse. L'utilisation par ces professionnels d'un système de reconnaissance vocale a ainsi l'avantage d'être instantané. Par ailleurs, si à partir du moment où les

patients ont été vus, les délais de réalisation et de mise à disponibilité des rapports dans le réseau informatique de l'hôpital sont trop longs, cela peut empêcher les médecins de soigner leurs patients en ayant toute l'information nécessaire. Cette situation peut ainsi conduire à des retards dans le début du traitement et dans de rares cas à des erreurs médicales (Kopach et al., 2005). Au contraire, le système de reconnaissance vocale pourrait conduire à ce que le patient commence son traitement plus tôt. De même, disposer d'un rapport plus tôt permettrait de libérer le patient plus tôt et de « libérer » des lits d'hôpitaux. Ce système peut également permettre des rapports plus précis et limiter le nombre de rapports qui se perdent dans le réseau informatique (i.e. meilleur contrôle du processus de documentation car tout est dans le système). En outre, la rapidité du système devrait permettre aux médecins de se faire payer plus rapidement. Finalement, certains aspects doivent être mentionnés, comme le besoin d'immédiateté et de contrôle du médecin sur la réalisation de son rapport (Reinus, 2007). Par conséquent, même si les systèmes de reconnaissance vocale peuvent s'avérer être plus dispendieux sur une base budgétaire que le système traditionnel de transcription, ils pourraient cependant être appréciés pour tout un ensemble d'aspects qui ne sont pas évalués de façon quantitative et monétaire.

Une stratégie intéressante d'utilisation du système de reconnaissance vocale pourrait être d'identifier les départements médicaux où ce système serait le plus efficace. Dans la mesure où la moitié des articles

que nous avons étudié ont été élaboré au sein de départements de radiologie, cela semble suggérer que les systèmes de reconnaissance vocale sont surtout implantés en des lieux utilisant un vocabulaire très spécifiques et assez répétitif. En effet, en radiologie, certaines phrases se répètent très souvent dans les rapports médicaux. Une telle situation permettrait au système de reconnaissance vocale d'être très efficace. Par contre dans d'autres spécialités, telles que la pédiatrie ou la psychiatrie, cela est beaucoup moins le cas et les gains à attendre du système de reconnaissance vocale sont bien plus faibles (Borowitz, 1999, 2001 ; Issenman et Jaffer, 2004 ; Morh et al., 2003).

De manière générale, le système de reconnaissance vocale s'avère surtout efficace dans les cas de figure où les personnes en charge de faire la transcription traditionnelle sont « lentes » et/ou quand les rapports sont dictés de façon claire, avec des termes répétitifs d'un rapport à un autre (Vorbeck et al., 2000 ; Mohr et al., 2003), avec des mots simples portant peu à confusion et d'une durée moyenne supérieure à 3 minutes (Mohr et al., 2003).

2.4 Conclusion

Les résultats de certaines études tendent à montrer que l'utilisation d'un système de reconnaissance vocale visant à retranscrire des rapports médicaux pourrait se révéler être une option judicieuse pour le CHUS. Toutefois, les résultats de notre revue systématique de la littérature ne sont

pas assez probants dans la mesure où un nombre non négligeable d'études démontrent que les inconvénients d'un tel système sont supérieurs à ses avantages (Issenman et Jaffer, 2004 ; Pezzullo et al., 2008). Par ailleurs, nous aurions pu nous attendre, avec l'amélioration progressive de cette technologie, à ce que les études les plus récentes indiquent de meilleurs résultats. Ce n'est pas ce que nous avons constaté : la proportion d'études indiquant une plus grande efficacité et/ou efficacité du système de reconnaissance vocale sur le système traditionnel restant assez stable dans le temps. Notre recherche littéraire ne permet donc pas de conclure sauf dans certains domaines.

En résumé, le système de reconnaissance vocale 1) ne permet pas de diminuer le nombre d'erreurs de transcription par rapport à une transcription humaine ; 2) connaît une forte capacité d'amélioration après une première phase d'adaptation, mais cette capacité diminue par la suite ; 3) le temps initial d'apprentissage pour les médecins est relativement court,

toutefois, une phase d'entraînement de plusieurs semaines est nécessaire avant de parvenir à des résultats satisfaisants ; 4) le temps de production avec interaction du médecin semble être significativement plus long qu'avec le système traditionnel de transcription ; 5) les études sur le temps de production sans interaction du médecin ne sont pas en nombre suffisant pour parvenir à une conclusion fiable ; 6) avec interaction du médecin, l'efficacité du système de reconnaissance vocale n'est pas démontrée, bien au contraire ; 7) sans interaction du médecin, cette solution pourrait être efficace, toutefois, nous ne disposons pas de suffisamment d'éléments pour conclure.

Bien que réalisées avec moins de rigueur méthodologique, les études que nous avons classées comme étant de moindre qualité aboutissent à des résultats très similaires à ceux des études que nous avons classées comme étant d'une plus grande qualité. Ces données peuvent dès lors renforcer le poids de nos conclusions.

CHAPITRE 3

3. RECOMMANDATIONS

Comme mentionné dans notre revue de la littérature scientifique, il existe plusieurs façons d'utiliser un système de reconnaissance vocale dans un milieu hospitalier : soit le médecin fait tout et interagit avec le système de reconnaissance vocale, soit le médecin transmet la bande sonore de sa dictée au service de transcription qui utilise le système de reconnaissance vocale. Dans le premier cas, le temps passé par le médecin à réaliser son rapport médical est plus long qu'avec le système traditionnel (i.e. le médecin doit toujours faire la dictée et la correction est plus longue). Dans le second cas, les résultats des études ne sont pas en nombre suffisants.

Recommandation n° 1

Compte tenu des résultats fournis par cette analyse des données probantes de la littérature scientifique, notre première recommandation est de n'utiliser le système de reconnaissance vocale que dans les situations de pénurie de transcripateurs ainsi qu'en situation urgente (i.e. le médecin est pressé de rédiger son rapport médical).

Recommandation n° 2

Notre deuxième recommandation est de réaliser une étude pilote pour tester l'efficacité et l'efficacité de la reconnaissance vocale dans le contexte

du CHUS. Compte tenu des différences de prix et des avantages et inconvénients de chacun des systèmes de reconnaissance vocale offerts sur le marché, nous recommandons, en dehors du système G2 Speech de Philips, que d'autres logiciels soient également testés et comparés: IBM ViaVoice Pro (210 CAD), Dragon Dictate Naturally Speaking (220 CAD)⁶. À noter que G2 Speech et Dragon Dictate Naturally Speaking nécessitent l'ajout d'un dictionnaire médical (spécialisé), alors que IBM ViaVoice Pro dispose d'un programme performant d'auto-apprentissage qui permet d'éviter un tel ajout (Capel et al., 2004). Néanmoins, le système IBM ViaVoice Pro ne bénéficiant plus depuis 2005 d'une actualisation de sa base de données, nous ne recommandons pas son utilisation.

Si une étude pilote est réalisée, il faudra prévoir l'adaptation des médecins et professionnels au système de reconnaissance vocale lors de journées « creuses » afin d'éviter une surcharge de stress. Il faudra également déterminer si le système de reconnaissance vocale retenu est un système en continu ou en discret. Dans le cadre d'une interaction avec le système de reconnaissance vocale, le système en discret est pour certains médecins plus intéressant dans la

⁶ Coût pour une licence, avec frais de transport. Dollars canadiens de 2008.

mesure où les dictées sont discontinues, que les médecins ont souvent besoin de faire une pause pour rassembler leur pensée ou examiner leurs documents ; par ailleurs le système en discret peut dans un tel cas de figure s'avérer faire moins d'erreurs de transcription (Smith, 2002). Le système de reconnaissance vocale en discret permet également de corriger immédiatement les erreurs en revenant sur le dernier mot dicté, ce qui n'est pas toujours le cas avec le système en continu. Finalement, le système de reconnaissance vocale en discret incite les médecins à ne délivrer que des mots clefs et ainsi à réduire de façon considérable la taille du rapport sans pour autant y perdre en lisibilité et en informations. De plus, certains mots sont associés à des codes permettant de retranscrire dans le rapport un ensemble de mots préprogrammés. Finalement, si une telle étude est réalisée, il faudra faire bien attention à distinguer les erreurs majeures, qui changent la compréhension du rapport médical, des erreurs mineures

(orthographe, grammaires, erreurs de syntaxes, etc.).

Par la suite, si les résultats de cette étude pilote sont suffisamment positifs, nous recommandons de commencer à implanter le système de reconnaissance vocale dans les départements où la volonté est la plus forte pour ensuite – si les résultats sont positifs – étendre le système à d'autres départements par effets de débordement. Par ailleurs, si cela augmente le temps de réalisation des rapports pour certains professionnels de santé ou médecins, il faudra faire en sorte que les différents bénéficiaires de ce système soient au courant de la contrainte de temps supplémentaire que cela engendre pour ces professionnels de santé ou médecins. De même, il faudra publiciser les économies de coûts et de temps générées afin de valoriser les différentes personnes impliquées dans l'utilisation de ce système.

BIBLIOGRAPHIE

- Al-Aynati, M.M. et Chorneyko, K.A. (2003). Comparison of voice-automated transcription and human transcription in generating pathology reports. *Archives of Pathology and Laboratory Medicine*, 127 (6), pp. 721-725.
- Antiles, S., Couris, J., Schweitzer, A., Rosenthal, D. et Da Silva, R.Q. (2000). Project planning, training, measurement and sustainment: the successful implementation of voice recognition. *Radiology Management*, 22 (1), pp. 18-31.
- Aprile, I., Tommasini, G., Iaiza, F., Biasizzo, E., Lavaroni, A., D'Agostini, S. et Fabris, G. (1997). Use of a speech recognition system in neuroradiology. *Rivista di Neuroradiologia*, 10 (suppl. 2), pp. 251-252.
- Arndt, H., Petersein, J., Stockheim, D., Gregor, P., Hamm, B. et Mutze, S. (1999). The use of automated speech recognition in diagnostic radiology. *Rofo: Fortschritte auf dem Gebiete der Rontgenstrahlen und der Nuklearmedizin*, 171 (5), pp. 400-404.
- Borowitz, S.M. (1999). Abstract of Section Scientific Presentations: Sections on Computers and Other Technologies. *Pediatrics*, 104 (suppl. 3), part 3 of 3, p. 670.
- Borowitz, S.M. (2001). Computer-based Speech Recognition as an Alternative to Medical Transcription. *Journal of the American Medical Informatics Association*, 8 (1), pp. 101-102.
- Briggs, B. (2004). *Speech Recognition is Finding its Voice*. Health Data Management and SourceMedia, Inc.
- Bundt, T.S. (2002). Transcription Cost Analysis Reveals Opportunities for Savings, Efficiencies. *Journal of the American Health Information Management Association*, 73 (8), pp. 55-60.
- Callaway, E.C., Sweet, C.F., Siegel, E., Reiser, J.M. et Beall, D.P. (2002). Speech recognition interface to a hospital information system using a self-designed visual basic program: initial experience. *Journal of Digital Imaging*, 15 (1), pp. 43-53.
- Capel, D., Soltner, C., N'Guyen, J-L. et Beydon, L. (2004). Logiciel de reconnaissance vocale pour les comptes rendus d'hospitalisation de réanimation. *Annales Françaises d'Anesthésie et de Réanimation*, 23, pp. 344-348.
- Chang, P., Sheng, Y-H., Sang, Y-Y., Wang, D-W. (2008). Developing a Wireless Speech- and Touch-Based Intelligent Comprehensive Triage Support System. *CIN: Computers, Informatics, Nursing*, 26 (1), pp. 31-38.
- Corces, A., Garcia, M., Gonzalez, V. et Varela, M. (2004). Word Recognition Software use in a Busy Orthopaedic Practice. *Clinical Orthopaedics and Related Research*, 421, pp. 87-90.
- Devine, E.G., Gaehde, S.A. et Curtis, A.C. (2000). Comparative evaluation of three continuous speech recognition software packages in the generation of medical reports. *Journal of the American Medical Information Association*, 7 (5), pp. 462-468.
- Dietz, U., Rupprecht, H.J., Espinola-Klein, C. et Meyer, J. (1996). Automatic report documentation in cardiology using a speech recognition system. *Zeitschrift fur Kardiologie*, 85 (9), pp. 684-688.
- Gale, B., Safriel, Y., Lukban, A., Kalowitz, J., Fleischer, J. et Gordon, D. (2001). Radiology report production times: voice recognition vs. transcription. *Radiology management*, 23 (2), pp. 18-22.
- Hailey, D., Roine, R. et Ohinmaa, A. (2002). Systematic review of evidence for the benefits of telemedicine. *Journal of Telemedicine and Telecare*, 8 (Suppl.1), pp. 1-7.
- Haug, P. et Booth, E. (1999). Utah Health System Experiments With Voice Recognition Software. *IT Health Care Strategist*, 1 (3), pp. 9.
- Happe, A., Pouliquen, B., Burgun, A., Cuggia, M. et Le Beux, P. (2003). Automatic concept extraction from spoken medical reports. *International Journal of Medical Informatics*, 70, pp. 255-263.

- Hayt, D.B. et Alexander, S. (2001). The pros and cons of implementing PACS and speech recognition systems. *Journal of Digital Imaging*, 14 (3), pp. 149-157.
- Henricks, W.H., Roumina, K., Skilton, B.E., Ozan, D.J. et Goss, G.R. (2002). The Utility and Cost Effectiveness of Voice Recognition Technology in Surgical Pathology. *Modern Pathology*, 15 (5), pp. 565-571.
- Herman, S.J. (2004). Speech recognition and the creation of radiology reports. *Applied Radiology*, May 2004, pp. 23-28.
- Hundt, W., Stark, O., Scharnberg, B., Hold, M., Kohz, P., Lienemann, A., Bonél, H. et Reiser, M. (1999). Speech processing in radiology. *European radiology*, 9, pp. 1451-1456.
- Ichikawa, T., Koizumi, J., Takahara, T., Myojin, K., Yamashita, E., Nasu, S., Yanagimachi, N., Imai, Y. et Tsukune, Y. (2005). Continuous speech recognition system for radiological reporting: comparison with experience of dictation. *Nippon Igaku Hoshasen Gakkai Zasshi – Nippon Acta Radiologica*, 65 (4), pp. 384-386.
- Issenman, R.M. et Jaffer, I.H. (2004). Use of Voice Recognition Software in an Outpatient Pediatric Speciality Practice. *Pediatrics*, 114 (3), pp. 290-293.
- Kanal, K.M., Hangiandreou, N.J., Sykes, A.M., Eklund, H.E., Araoz, P.A., Leon, J.A. et Erickson, B.J. (2001). Initial evaluation of a continuous speech recognition program for radiology. *Journal of Digital Imaging*, 14 (1), pp. 30-37.
- Kopach, R., Sadat, S., Gallaway, I.D., Geiger, G., Ungar, W.J. et Coyte, P.C. (2005). Cost-effectiveness analysis of medical documentation alternatives. *International journal of technology Assessment in Health care*, 21 (1), pp. 126-131.
- Korn, K. (1998). Voice Recognition Software for Clinical Use. *Journal of the American Academy of Nurse Practitioners*, 10 (11), pp. 515-517.
- La Fianza, A., Giorgetti, S., marelli, P. et Campani, R. (1993). Vocal recognition in dental and oral radiology. *Radiologica Medica*, 86 (4), pp. 432-435.
- Langer, S. (2002a). Radiology Speech Recognition: Workflow, Integration, and Productivity Issues. *Current Problems in Diagnostic radiology*, May/June, pp. 95-104.
- Langer, S. (2002b). Impact of Speech Recognition on Radiologist Productivity. *Journal of Digital Imaging*, 15 (4), pp. 203-209.
- Lemme, P.J. et Morin, R.L. (2000). The implementation of speech recognition in an electronic radiology practice. *Journal of digital imaging*, 13 (2 suppl. 1), pp. 153-154.
- Lui, C.Y., Tam, K.F., Lam, H.S. et Chan, L.K. (2001). Continuous speech recognition system in a non-native English speaking radiology department. *Asian Oceanian Journal of Radiology*, 6 (1), pp. 59-64.
- Mohr, D.N., Turner, D.W., Pond, G.R., Kamath, J.S., De Vos, C.B., et Carpenter, P.C. (2003). Speech Recognition as a Transcription Aid: A Randomized Comparison With Standard Transcription. *Journal of the American Medical Informatics Association*, 10 (1), pp. 85-93.
- Oxman, A.D. et Guyatt, G.H. (1993). The science of reviewing research. *Annals of the New York Academy of Sciences*, 703, pp. 125-34.
- Pezzullo, J.A., Tung, G.A., Rogg, J.M., Davis, L.M., Brody, J.M. et Mayo-Smith, W.W. (2008). Voice Recognition Dictation: Radiologist as Transcriptionist. *Journal of Digital Imaging*, 21 (4), pp. 384-389.
- Ramaswamy, M.R., Chaljub, G., Esch, O., Fanning, D.D. et Van Sonnenberg, E. (2000). Continuous Speech Recognition in MR Imaging Reporting: Advantages, Disadvantages, and Impact. *American Journal of Roentgenology*, 174, pp. 617-822.
- Rana, D.S., Hurst, G., Shepstone, L., Pilling, J., Cockburn, J. et Crawford, M. (2005). Voice recognition for radiology reporting: Is it good enough? *Clinical Radiology*, 60, pp. 1205-1212.

- Reinus, W.R. (2007). Economics of Radiology Report Editing Using Voice recognition Technology. *Journal of the American College of Radiology*, 4 (12), pp. 890-894.
- Rosenthal, D.I., Chew, F.S., Dupuy, D.E., Kattapuram, S.V., Palmer, W.E., Yap, R.M., Levine, L.A. (1998). Computer-Based Speech Recognition as a Replacement for Medical Transcription. *American Journal of Roentgenology*, 170, pp. 23-25.
- Sferrella, S.M. (2003). Success with voice recognition. *Radiology Management*, 25 (3), pp. 42-49.
- Simmons, J. (2002). Talking it through: busy physical therapy practice converts from manual transcription to voice recognition – What Works: Speech Recognition. *Health Management Technology*, 23 (2), pp. 38.
- Smith, K.C. (2002). A Discrete Speech Recognition System for Dermatology: 8 Years of Daily Experience in a Medical Dermatology Office. *Seminars in Cutaneous Medicine and Surgery*, 21 (3), pp. 205-208.
- Spring, C. (2003). Productivity gains of speech-recognition technology: organizations that gain the most will be those armed with the facts, tempered by reasonable expectations – Voice Recognition/Transcription. *Health Management Technology*, 24 (1), pp. 54-55.
- Takahara, T., Nakajima, M., Nitatori, T. et Hachiya, J. (2002). Japanese radiological report creation with continuous speech recognition. *Nippon Igaku Hoshasen Gakkai Zasshi – Nippon Acta Radiologica*, 62 (1), pp. 23-26.
- Teichgräber, U.K., Ehrenstein, T., Lemke, M., Liebig, T., Stobbe, H., Hosten, N., Keske, U. et Felix, R. (1999). Automated speech recognition for the generation of medical records in computed tomography. *Rofo*, 171 (5), pp. 396-399.
- Threet, E. et Fargues, M.P. (1999). Economic evaluation of voice recognition for the clinicians' desktop at the Naval Hospital Roosevelt Roads. *Military Medicine*, 164 (2), pp. 119-126.
- Trumm, C.G., Glaser, C., Paasche, V., Crispin, A., Popp, P., Kuttner, B., Francke, M., Nissen-Meyer, S. et Reiser, M. (2006). Impact of a PACS/RIS-integrated speech recognition system on radiology reporting time and report availability. *Rofo*, 178 (4), pp. 400-409.
- Ury, A. (2007). Practices can speed productivity and reduce costs by adding speech recognition to their EHRs. *Healthcare Informatics*, 24 (9), pp. 57-58.
- Van Terheyden, N. (2005). Is speech recognition the Holy Grail? *Health Management Technology*, 26 (2), pp. 42-45.
- Vorbeck, F., Ba-Ssalamah, A., Kettenbach, J. et Huebsch, P. (2000). Report generating using digital speech recognition in radiology. *European Radiology*, 10, pp.1976-1982.
- Wheeler, S. et Cassimus, G.C. (1999). Selecting and implementing a voice recognition system. *Radiology Management*, 21 (4), pp. 37-42.
- Zafar, A., Overhage, J.M. et McDonald, C.J. (1999). Continuous Speech Recognition for Clinicians. *Journal of the American Medical Informatics Association*, 6, pp. 195-204.
- Zemmel, N.J., Park, S.M., Schweitzer, J., O'Keefe, J.S., Laughon, M.M. et Edlich, R.F. (1996). Status of voicetype dictation for windows for the emergency physician. *The Journal of Emergency Medicine*, 14 (4), pp. 511-515.
- Zick, R.G. et Olsen, J. (2001). Voice Recognition Software Versus a Traditional Transcription Service for Physician Charting in the ED. *American Journal of Emergency Medicine*, 19 (4), pp. 295-298.

Liste des abréviations utilisées dans la grille de synthèse :

RV : Reconnaissance Vocale

MD : Médecin

tps : temps

min : minute

h : heure

s : seconde

CAD : Canadian Dollar

USD : United States Dollar

vs : versus

P : P-value

Nx : Nouveau

bcp : Beaucoup

pr : pour

chq : chaque

NO : Notes à l'Ordinateur

ST : Sous-Traitance

TM : Transcription Maison

NS : Naturally Speaking

FOR : Formation

Auteurs	Année	Langue logiciel	Objet	Random	Prospective cas=contrôle	Capacité à reconnaître	Capacité apprentissage	Temps apprentissage	Temps de production	Efficience	Coûts	Remarques
Al-Aynati et Chorneyko	2003	Anglais	Rapports de pathologies Traditionnel vs. IBM ViaVoice Pro version 8 (médical)	Oui	Oui, 206 rapports de 114 mots de moyenne	93.6% pour RV vs. 99.6% (en moyenne)	Oui	30 min	RV en moyenne 2 fois plus long à réviser par le MD	Oui	Implantation du matériel 2250 cad	À ne pas implanter quand il y a assez de secrétaires P<0.001
Antiles et al.	2000	Anglais	Radiologie Traditionnel vs. RV 75% des rapports sont fait avec RV MD fait tout	Non	Non	X	X	6 mois pour former 72 MD, 26 clinician fellows et 29 résidents	En deux ans, turnaround réduit de 4.3 à 2.2 jours	Économie de 530.000 usd par an (pas de détails)	60 postes de RV	Faut une bonne planification, de la formation continue et du soutien technique
Aprile et al.	1997	Italien	Neuroradiologie Traditionnel vs. Voice Type rel 1.32	X	X	3% d'erreurs avec RV. Rapport plus compréhensible	X	X	Réduction significative du temps avec RV	X	X	X
Arndt et al.	1999	Allemand	Radiologie Apprentissage du SP 6000	X	Non	8.4 à 13.3% d'erreurs au début. 2.4 à 10.7% après 1 ^{ère} correction.	La 1 ^{ère} adaptation corrige beaucoup, les suivantes peu	2 à 3 heures (y compris la formation initiale)	SP 6000 est plus rapide que d'écrire le rapport à la main	X	X	Le risque d'erreur est indépendant du médecin
Borowitz	1999	Anglais	Pédiatrie Traditionnel vs. IBM Via Voice vs. Dragon Naturally Speaking	Oui, 729 rapports. 210 Via Voice et 162 Dragon	Non	X	X	X	Tps dictée: 180 vs 178 vs 198 secondes Tps révision: 149 vs 235 vs 225 secondes Tps achèvement 5.94 vs 0.15 vs 0.28 jours	Temps dictée + édition est 25% plus long avec RV mais dossier est plus rapidement dans le système	Coûts de transcription peuvent constituer 11% des frais de visite	Même proportion de visite follow up et nouvelle P<0.0001 Longueur des rapports 10% moins long pour IBM Via Voice
Borowitz	2001	Anglais	Pédiatrie Traditionnel vs. IBM ViaVoice Millenium	Oui 580 vs 549	Non	X	X	50 minutes de la part du MD	Tps dictée : 180 vs. 203 secondes pour RV Tps révision : 147 vs. 177 sec pour RV	Les rapports sont complétés plus tôt avec RV = plus rapidement dans le système : 5.3 vs <1 jour	X	Pas de gabarits Pas de diff significative dans les types de visites Rapports avec RV un peu plus long : 2 lignes P<0.05
Briggs	2004	Anglais	Utilisation de RV (Dictaphone Corp.) dans une firme de transcription	X	X	95%	Oui	X	Gain de productivité de 24% (de 5 à 53%)	Oui	X	Le travail a changé, désormais sont surtout des correcteurs.
Bundt	2002	Anglais	Électromyographie Notes à l'ordinateur (situation initiale) vs. RV vs. sous-traitance vs. transcription maison	Non	Non	X	X	Une seule session pour MD avec RV.	Avant, 30 min par MD pour faire 1 rapport. Hypothèse que MD passe 15 min avec RV et 5 min avec transcription.	Pour 2352 rapports/an, le coût -en \$- moyen par rapport est : NO = 35.00 RV = 18.86, ST = 14.23, TM = 23.03	Besoin de 1.5 trans pour les 2352 rapports des 3 MD.	Analyse de sensibilité sur durée de dépréciation. Nouveaux syst permettraient de voir 20-38.5% plus de patients

Auteurs	Année	Langue logiciel	Objet	Random	Prospective cas=contrôle	Capacité à reconnaître	Capacité apprentissage	Temps apprentissage	Temps de production	Efficience	Coûts	Remarques
Callaway et al.	2002	Anglais	Radiologie Contrat transcription vs. Visual Basic 6.0 de Microsoft	X	X	X	X	X	Le dossier est mis beaucoup plus rapidement dans le système (3-4 fois moins de temps)	10650 usd de coûts de transcription éparpillés pour 5109 rapports	Coût de 3000 usd par implantation	Ont également testé pour la transcription de rapports faits par étudiants (positif)
Capel et al.	2004	Français	Réanimation Traditionnel vs. IBM ViaVoice Pro Edition Version 8	X	Oui, 20 rapports Médiane de 429 mots	97% pour Secrétaire 92% pour Interne 95% pour MD	60 fichiers Word fournis au logiciel pour qu'il améliore vocabulaire et syntaxe. 30 min pour apprendre la voix du MD + nouveaux mots	1 heure pour les MD	Tps dictée : 30-50% plus court avec RV Tps révision : idem pour MD, 30-50% + long pour Interne	Achat amorti en quelques semaines	Micro inclus dans pack 210 cad en 2008	Local silencieux <u>Texte lu</u> Élocution du MD est plus régulière que celle de l'Interne Pas de dictionnaire médical P<0.05
Chang et al.	2008	Chinois	Infirmière <u>mobile</u> au triage en urgence On teste le Speech SDK 5.1 de Microsoft	Sélection de 30 infirmières selon leur disponibilité	Non. 3 scénarios testés ; chacun par 10 infirmières	1% d'erreurs. Accent ou bruit ne semble pas jouer sur le résultat	Pas cette possibilité (ce n'est pas un problème lié à la technologie) 410 mots	30 minutes au max Les inf n'ont jamais utilisé RV avant	108 secondes dans un environnement simulé de 61 dB Pourrait être plus long en situation réelle Sans RV c'est entre 30 et 120 s	X	X	Forte volonté à utiliser ce système. Inf font la correction. <u>Système ne reconnaît que les mots clés.</u> Permet d'avoir les mains libres
Corces et al.	2004	Anglais	Orthopédie Dragon Naturally Speaking 5.0	Non Expérience naturelle avec un nouveau médecin	Non	Clinique : 10% d'erreurs en moyenne sans gabarits, 4% avec. <u>Nx MD</u> : au début il y a plus d'erreurs (10%) et de temps passé. Mais au bout de 12 semaines, les gains sont importants (4% d'erreurs).	Avec l'utilisation de fichiers Word pré-remplis par type de problème (gabarits), la rapidité et la qualité augmentent	5 à 30 minutes + ajustements au cours du 1er mois	<u>Nx MD</u> : Au début : 6.4 minutes par rapport. À la fin : 3.6 minutes par rapport.	Économies de 1500 à 2000 usd par mois, mais on ne connaît pas le volume	199 usd par logiciel La version médicale n'a pas été jugée utile	La notice du système de RV indique 4-5% d'erreurs. Le médecin corrige lui-même. Le système permet d'avoir des documents plus complets. <u>Faut parler lentement et articuler.</u>
Devine et al.	2000	Anglais	Comparaison de 3 versions médicales: IBM ViaVoice 98, Dragon System 3.0 et L&H Voice Xpress 1.2	X	Oui 12 MD	7-9% d'erreurs avec IBM, 13-15% avec L&H et 14-15% avec Dragon System	X	X	X	X	X	X
Dietz et al.	1996	Allemand	Cardiologie Traditionnel vs. IBM Voice Type	X	Non 200 rapports de 301 mots en moyenne	95.1%	X	X	Temps moyen de RV supérieur de 3 minutes	X	X	Moins de faute et aucun besoin de vérification préliminaire, c'est le MD qui corrige direct.

Auteurs	Année	Langue logiciel	Objet	Random	Prospective cas=contrôle	Capacité à reconnaître	Capacité apprentissage	Temps apprentissage	Temps de production	Efficience	Coûts	Remarques
Gale et al.	2001	Anglais	Radiologie Traditionnel vs. RV	X	Oui, 69 rapports (t-test par sous groupe de catégorie de rapport)	X	X	X	Cela prend 162 s au radiologue (qui fait tout) vs. 74 s sans RV	Économies en secrétaire 100.000 usd par an, mais plus faible productivité du radiologue		Bcp prob avec le système de RV (bug, etc.), ce qui augmente le tps réalisation des rapports
Happe et al.	2003	Français	Coronographie Naturally Speaking 5 vs. IBM ViaVoice 8 (version médicale - vm - dans chaque cas)	X	Oui, 28 rapports de 140 mots en moyenne	Sans version médicale, 9% d'erreurs. Avec, 4.2% d'erreurs pour NS vs. 1.7% pour IBM	Faible pour IBM vm + forte avec NS : passe de 5.7 à 4.2% d'erreurs	X	X	X	X	X
Haug et Booth	1999	Anglais	Traditionnel vs. RV	X	X	95-98%	Oui	X	50 à 90 min de + par jour pour 150 rapports	X	X	MD fait tout
Hayt et Alexander	2001	Anglais	Radiologie PACS + RV	X	Non	Quelques erreurs	X	X	86% des rapports produits en - de 12 h contre 3% avant	- d'attente en urgence pour avoir diag. Radiologue passe 25% de temps en + pr faire rapports	X	Bcp moins de discussion entre cliniciens et radiologues : bcp moins de consultation face à face
Henricks et al.	2002	Anglais	Biopsie Traditionnel vs. Clinical Reporter for Pathology Version 4.02	X	Non	70-90% d'exact	X	2-3 heures la 1 ^{ère} fois. 4 semaines pour être efficace	Améliore le turnaround des rapports	Retour sur investissement : 1.9 année 23864 mots par mois = 2625\$ économisés en coût de transcription	120 h de configuration pour mettre les gabarits + 40 h de tests	Développer les gabarits est important. Pas de MD, on a des techniciens qui dictent et corrigent direct
Herman	2004	Anglais	Revue littérature Traditionnel vs. RV en radiographie	X	X	90-100%	X	X	moindre	Fortes économies (ex : 1.7 M sur 5 ans)	X	Radiologue fait la dictée et corrige : prend bcp de tps en +
Hundt et al.	1999	Allemand	Radiologie Traditionnel vs. Phillips SP 6000	Oui, 400 rapports de 135 mots en moyenne	Non	Passe de 79.6% d'exact à 92.5% après 3 adaptations Erreurs surtout dues à des substitutions ou insertions de mots	Intègre nouveaux mots, s'adapte aux corrections en les associant aux voix	Adaptation 1 a eu lieu au bout de 70 min, la 2 à 160 min et la 3 à 430 min.	Tps dictée : 2.5 min vs. 3.4 pour RV. Tps correction : 1.1 min vs. 4.7 Avec RV, le rapport est dans le système en 8.1 min vs. 23.5h	Efficient si on considère le temps mis pour rendre le rapport disponible, sinon le temps d'élaboration est + long	X	Système s'améliore de 5% à chaque adaptation. Radiologue fait tout avec RV. Secrétaire fait aussi correction.
Ichikawa et al.	2005	Japonais	Radiologistes habitués à la dictée pour secrétaire vs. nouveaux radio. RV = AmiVoice	Oui	X	96.4 vs. 95.9% d'exact (RV dans les 2 cas)	X	X	X	X	X	Pas de diff. significative entre les 2 groupes de radiologues

Auteurs	Année	Langue logiciel	Objet	Random	Prospective cas=contrôle	Capacité à reconnaître	Capacité apprentissage	Temps apprentissage	Temps de production	Efficience	Coûts	Remarques
Issenman et Jaffer	2004	Anglais	Pédiatrie Traditionnel vs. Dragon Naturally Speaking 6.0	X	Oui, 42 rapports de 225 mots en moyenne	99.97% d'exact vs. 90.8% pour RV	X	2h orientation + 2h formation + 1 mois entraînement (30 rapports)	MD passe 3 fois + de tps à dicter et corriger avec RV (9.6 vs. 3.3 min) après entraînement (dont 1 min de set up)	RV coûte 2 fois plus sur un an (licence, tps en + pour MD) 15290 vs. 6970 usd	Salaires horaires MD = 100 usd, secrétaire = 11 usd	1 seul des 4 MD sélectionné participe. Autres sont frustrés par le logiciel MD fait correction P<0.001
Kanal et al.	2001	Anglais	Radiologie Différents groupes de personnes testent IBM MedSpeak 1.1	X	Oui, 12 rapports pour 6 personnes	88.4% pour les non natifs et 90.3% pour les natifs	X	X	X	X	X	Pas de diff selon le sexe. P<0.01
Kopach et al.	2005	Anglais	Projection éco Tout l'hôpital Traditionnel vs. RV + signature électronique	Non	Non 50% des MD vont ici utiliser RV	X	X	X Qui fait dictée?	Tps de réalisation est estimé passer de 70 à 14 jours avec RV	Pour réduire le délai d'une journée, le coût par rapport est de 0.33 cents CAD de 2003	5.631.669 avec RV vs 4.794.671 CAD sur 4 ans	En 2000, tps de réalisation d'un rapport est de 70 jours à l'hôpital de Sunnybrook Analyses de sensibilité
Korn	1998	Anglais	Dragon Naturally Speaking	Non	Non	Forte	Faible	15 minutes	Dictée + correction: 5 m	X	X	X
La Fianza et al.	1993	Italien	Radios orthopantomographiques Traditionnel vs. RV	Non	Oui, 760 rapports	Flesh index de 28.9 vs 30.62 pr RV = meilleure lisibilité. Opinion de 4 MD également meilleure pr RV	X	X	2.99 vs 2.24 minutes pr RV	X	X	Moins de faute de grammaire et de transmission P<0.0005 Faut une bonne prononciation pour ce résultat
Langer	2002a	Anglais	MedSpeak V1.3 vs. PowerScribe V2.2a	Non	Non	96% dans salle silencieuse avec MedSpeak après qq semaines ; 95% même avec bruit pour PS	Oui, qq semaines	X	PS trois fois plus efficace que MS pour le turnaround (MD fait tout)	Coût/rapport : MS = 17.12 ; PS = 12.2 ; Human = 11.84 usd de 1996	X	RV marche pas si bcp de logiciels sur ordinateur
Langer	2002b	Anglais	Radiologie Secrétaire (1) vs. RV (2) vs. Secrétaire + PACS (3) vs. PACS + RV (4)	Non, 40 sites, envoi de survey	Non	X	MD fait tout Avec RV	X	En turnaround, par rapport à (1) (2) est 127% + efficace, (3) 137% + et (4) 189% + P<0.15	X	X	25% de réponse complète au survey. Les sites utilisent pas tous à 100% la technologie.
Lemme et Morin	2000	Anglais	Radiologie Traditionnel vs. PowerScribe, L&H	Non	Non	Très bon taux	X	8 jours pour tout le département	Création du rapport passe de 2h à 1 minute	X	X	Radiologue fait la dictée et la correction.
Lui et al.	2001	Anglais	Radiologie MD non natif Traditionnel vs. RV	X	Oui	84.5-86.7% au 1 ^{er} coup, puis 95.7% après 4 fois avec même rapport	Oui	Faible	Transcription - rapide avec RV (MD fait tout)			Le bruit de fond influence bcp la performance. Avantageux en situation urgente

Auteurs	Année	Langue logiciel	Objet	Random	Prospective cas=contrôle	Capacité à reconnaître	Capacité apprentissage	Temps apprentissage	Temps de production	Efficience	Coûts	Remarques
Morh et al.	2003	Anglais	Endocrinologie et psychiatrie Traditionnel vs. RV Secrétaires vs. transcripteurs	Oui, débute après 4 semaines de routine	Non	84.5% Le taux ne progresse pas lors de l'étude; donc 6h de FOR + 4 semaines de routine semble optimal	X	6h pour les secrétaires et transcripteurs	Les personnes avec RV sont 12.7 à 44.2% moins productives.	Secrétaires moins efficaces que transcripteurs, avec RV. Pr personnes lentes, RV est plus efficace.	X	MD fait rien que signer MD sait pas si on utilise RV ou non (blind).
Pezzullo et al.	2008	Anglais	Radiologie privée IRM Traditionnel vs. TalkTechnology 2.1.28, AGFA	Oui, 200 rapports	Non	0.12 vs 5.1 erreurs par rapport avec RV (100 vs 96%). Après signature, 35% des rapports RV ont encore des erreurs contre 3% pr les autres. Pas d'erreurs graves.	X	3 années d'expérience avec RV	Dictée + correct 2.09 (1.59 de dictée) vs 4.11 min (2.02 de dictée??) pour RV. Rapport avec RV 20% plus courts (MD s'adapte à RV). Turnaround : 72.3 vs 33.5 min avec RV.	Surcoût de 6.1 usd de 2003 par rapport avec prise en compte du salaire MD. Coût avec transcription : 2.97 usd.	X	MD frustrés et fatigués. Juin-août 2004. Pas de macro ou de gabarits. P<0.0001 MD fait tout.
Ramaswamy et al.	2000	Anglais	IRM Traditionnel vs. MedSpeak 1.2 pour radiologie	Non	Non	92.7% en dictée spontanée. Erreurs d'omission passe de 0.3 à 1 mot / 1000, d'orthographe de 3 à 0.8. Très peu d'erreurs majeures.	Oui	1h pour MD	Turnaround baisse de 87.8 à 43.6 h. 62.5% disponibles <24h contre 10.5% avant	Retour sur investissement en 10 mois.	10.000 usd d'installation. 5072 rapports en 9 mois	La longueur des rapports a baissé de 95 à 60 mots grâce à effet de RV sur MD (baisse tps dictée) MD fait tout. No diff. selon l'expérience du MD avec RV
Rana et al.	2005	Anglais	Radiologie Traditionnel vs. Dragon NS pro v.5 MD expérimentés vs. non	Oui	Oui?	Plus d'erreurs avec RV, mais idem pour erreurs majeures ; NS avec MD expérimentés, pour autre MD il y a +d'erreurs	Oui	X	Tps total réduit avec RV : -82s pour MD expérimentés, -117s pour autre MD. Tps pour radiologie augmente : +14s	Bénéfice net de 20.000€ par an	X	MD fait tout. Étude pas en situation réelle. P<0.001
Reinus	2007	Anglais	Radiologie Modélisation des économies de coûts (Trop simpliste)	Non	Non	Pas pris en compte	Pas pris en compte	Pas pris en compte	Hyp que MD met entre 30s et 150s de tps en + avec RV	Oui, si pas trop de rapport à faire	X	Considère pas le besoin d'im-médiateité et de contrôle du MD
Rosenthal et al.	1998	Anglais	Radiologie Traditionnel vs. MedSpeak	Non	Non	Substitution de mots. + d'erreurs avec non natifs	X	1 semaine de formation sur le tas	Turnaround réduit de 99% : de 1013 à 13 min	Retour sur inv en 1 an	Coût installation 12.000 usd par unité	1 semaine en situation réelle. Tps interaction logiciel + long
Sferella	2003	Anglais	Radiologie Traditionnel vs. RV	Non	Non	X	X	2 semaines pour former 50 MD	Turnaround : 41% complétés en <24h avant, désormais entre 85-92%	12.3 FTE en moins. Retour sur inv en 12 mois	40 postes de RV	200 FTE et 50 MD pour le département MD fait tout

Auteurs	Année	Langue logiciel	Objet	Random	Prospective cas=contrôle	Capacité à reconnaître	Capacité apprentissage	Temps apprentissage	Temps de production	Efficience	Coûts	Remarques
Simmons	2002	Anglais	TRP Traditionnel vs. TalkNotes	Non	Non	99%	Oui	FOR 3 jours. Performant en 2 semaines	Turnaround réduit de 7-10 jours à - de 24h	50% de réduction des coûts	X	70 macros qui réduisent tps de dictée de 50%
Smith	2002	Anglais	Dermatologie Dragon Dictate en discret	Non	Non	Moins d'erreurs avec RV qu'avec transcripteur	Oui	Quelques jours	Tout est réalisé en moins de 24h Peu de tps passé en plus	De + en + efficient Économie de 100.000 usd en 8 ans	X	Depuis 1994, 40-60 rapports par jour. Faut de la persévérance
Spring	2003	Anglais	Résumé de différentes études de MedQuist	X	X MD fait rien	X	X	X	1 min de dictée = 3-5 min de transcription sans RV	30% de gains productivité pour les transcripteurs	X	D'un pt de vue médico-légal, chq rapport doit être vérifié
Takahara et al.	2002	Japonais	Radiologie Traditionnel vs. AmiVoice	X	Oui, 10 rapports	97.1%	Oui, à partir de 22589 rapports	Rapide	56.2 s (dont 25 s de correction) pr RV vs. 142.8 s pr transcripteur	X	X	Permet de réduire turnaround. MD fait tout
Teichgräber al.	999	Allemand	Tomographie Traditionnel vs. IBM ViaVoice	Oui, 200 rapports	Non	99 vs 97% pour RV. Zéro fautes d'orthographe avec RV, mais bcp plus de termes erronés	X	X	Tps dictée : 9.4 vs 4.5 min pour dictaphone. Turnaround de 47.3 vs 12.7 pour RV.	X	X	Recommandé pour rapports urgents. MD fait tout.
Threet et Fargues	1999	Anglais	Hôpital militaire Notes papiers vs. notes ordinateur vs. Dragon Dictate Août 96 - avril 97	Oui	Non	Précision des rapports évalués par MD externes. RV = le + précis	Introduction de 5 rapports (912 mots)	Nombre d'erreurs décroît après chaque utilisation	RV = 2 nd e meilleur temps derrière notes papiers	RV permettrait de faire face à la hausse des rapports sans employer de transcripteurs en plus	Coûts installation = 1400 usd par MD pr RV vs. 4440 par MD pr transcript	Pas d'incidence sur relation patient-MD quand RV utilisé pdt visite
Trumm et al.	2006	Allemand	Radiologie Traditionnel vs. RV+transcript vs. RV+MD	Oui	Non 669 rapports	X	X	X	Meilleur turn-around avec RV 10.47 vs. 6.65 vs. 1.27h Dictée plus rapide : 0.44 vs. 0.54 vs. 0.62 s par mot/image	X	X	On compare traditionnel avec MD fait tout ou MD fait rien Mode de calcul étrange
Ury	2007	Anglais	Avis d'un MD	Non	Non	97-98%	Oui	Entraînement de 10 min pour MD au début, puis on s'habitue	Meilleur turnaround	Transcription peut coûter 17000 usd par an par MD	X	Possible que le MD ait des liens financiers avec compagnies RV
Van Terheyden	2005	Anglais	Traditionnel vs. Philips	X	X	95%	Oui	X	Tps réalisation devrait passer de 8 à 5.6 min. Turnaround réduit de 24 à 17h avec RV	Permettrait de produire 30% de rapports en plus	X	MD a des liens avec Philips. Paramètres tirés des compagnies. MD fait rien

Auteurs	Année	Langue logiciel	Objet	Random	Prospective cas=contrôle	Capacité à reconnaître	Capacité apprentissage	Temps apprentissage	Temps de production	Efficience	Coûts	Remarques
Vorbeck et al.	2000	Allemand	Radiologie(s) Traditionnel vs. Philips SP6000 MD fait rien	X	Oui, 450 rapports Un MD utilise RV depuis 3 ans, l'autre non	0.4% vs 5.5% d'erreurs avec RV (avant correction par typist). Chq rapport est revu par les 2 MD.	Oui, 20 pages	X	En moyenne 19% plus rapide. Peut aller jusqu'à 28% en sonographie, à 24% pour MD expérimenté	% au traditionnel, transcripateur rapide gagne 14%, un lent gagne 25% avec RV.	X	Performance dépend du transcripateur, du MD (débit et expérience) et de la spécialité. Situation réelle.
Wheeler et Cassimus	1999	Anglais	Radiologie Traditionnel vs. RV	X	X	95-100%	X	X	Création du rapport passe de 30h à <15 min dans 65% des cas, <30 min dans 80% et <1h dans 90%	X	X	Radiologue fait tout
Zafar et al.	1999	Anglais	Critères pour avoir un faible taux d'erreurs avec RV	Non	Non	97% avec version médicale et 30 min de formation	Version de base : faut introduire 50 rapports + 400 autres mots pour atteindre 98% Si correction en continue, 5% de progrès chq 2 semaines	30 min pour MD	X	X	Avec version médicale (349-695 usd en +), pas besoin d'ajouter 400 mots pr avoir 98%	MD doit parler clairement, en continu, en articulant chq syllabe.
Zemmel et al.	1996	Anglais	Urgence. Évaluer la précision du dictionnaire médical pour urgence de IBM VoiceType.	Non	Oui, 1 rapport 7 lecteurs	Réduite si vocabulaire non médical (17 vs 9%), micro mal positionné (22 vs 13%) et bruit de fonds (28 vs 11%). En condition optimale on a 9% d'erreurs.	Oui	270 phrases à lire en 1h30. Puis 3h30 d'entraînement.	X	Trop coûteux	X	Non utilisable en urgence. P<0.05
Zick et Olsen	2001	Anglais	Urgence Traditionnel vs. Dragon Naturally Speaking Medical 4	Non	Oui, 47 rapports Un MD avec expérience, l'autre non	99.7 vs 98.5% pour RV. 1.2 correction par rapport vs 2.5 avec RV. Tout type d'erreur est considéré.	X	10 minutes de dictée + 20 min de configuration	Turnaround de 39.6 vs 3.65 min pour RV Tps dictée + correction : 3.77 vs 3.65 min pour RV.	RV est moins coûteux, mais calculs non crédibles	Transcription = 7.5 usd par rapport	Utilisent des gabarits. P<0.05 MD fait tout. Moins d'erreurs avec MD expérimenté.



ÉQUIPE DE L'UÉTMS

Maurice Roy, M.D., L.L.M., M.A.P.
Directeur général adjoint, directeur des services professionnels et directeur médical à la Direction interdisciplinaire des services clinique du Centre hospitalier universitaire de Sherbrooke (CHUS)

Renald Lemieux, M. Ing., Ph.D., M.ETS.
Directeur à la Direction de l'évaluation des technologies et des modes d'intervention en santé du CHUS

Christian Bellemare, M.Sc.
Coordonnateur de l'Unité d'évaluation des technologies et des modes d'intervention en santé du CHUS

Thomas Poder, M.Sc., Ph.D.
Conseiller en évaluation des technologies

Suzanne K. Bédard, B.A.
Conseillère en évaluation des technologies

COMMUNIQUER AVEC L'UÉTMS

Pour déposer une demande d'évaluation, pour commander un rapport d'évaluation déjà paru ou pour tout renseignement sur les activités de l'Unité, communiquez avec :

Unité d'évaluation des technologies et des modes d'intervention en santé (UÉTMS)

Centre hospitalier universitaire de Sherbrooke – Hôpital Fleurimont
3001, 12^e Avenue Nord
Sherbrooke (Québec) J1H 5N4

Téléphone : 819.346.1110 poste 13802
Courriel : uetmis.chus@ssss.gouv.qc.ca



Centre hospitalier
universitaire
de Sherbrooke

UNITÉ D'ÉVALUATION DES
TECHNOLOGIES ET DES MODES
D'INTERVENTION EN SANTÉ