

REVUE SYSTÉMATIQUE DES REVUES SYSTÉMATIQUES SUR L'UTILISATION DE LA RECONNAISSANCE VOCALE POUR LA DICTÉE MÉDICALE

NOTE DE SYNTHÈSE

Octobre 2017

Auteurs

Thomas PODER, Ph.D., cadre intermédiaire, UETMISSS,
CIUSSS de l'Estrie – CHUS

Jean-François FISETTE, Ph.D., conseiller en évaluation,
UETMISSS, CIUSSS de l'Estrie – CHUS

Messages clés

- La reconnaissance vocale permet généralement de réduire les délais de mise en disponibilité des rapports médicaux;
- En avant-plan, le rapport médical est disponible immédiatement, mais au prix d'une surcharge de travail pour les utilisateurs;
- En arrière-plan, un gain de productivité des transcriptionnistes est possible;
- La progression du taux de reconnaissance a été limitée au cours des deux dernières décennies.

INTRODUCTION

Afin de faciliter la production des rapports médicaux, de réduire leur délai de mise en disponibilité, ainsi que de potentiellement réduire les coûts de transcription, de nombreux systèmes de reconnaissance vocale (RV) ont été commercialisés depuis le début des années 80 (Leeming et al. 1981; Robbins et al. 1987). Très populaires dans les années 90, de nombreux systèmes de RV ont été implantés dans les hôpitaux, notamment aux États-Unis (Metha et McCloud 2003; Hammana et al. 2015). Dans les années 2000, on note cependant un essoufflement de cet enthousiasme, considérant que les attentes n'étaient pas toujours

rencontrées par les utilisateurs (Pezzulo et al. 2008; Alapetite et al. 2009; Ahlgrim et al. 2016).

Le principe général de la RV repose sur un système de dictée numérique combiné à 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 CIUSSS de l'Estrie – CHUS, que ce soit en réduisant le nombre de rapports à transcrire par le Service des archives grâce à une utilisation en avant-plan (Figure 1) de ce système par les médecins ou en réduisant le temps de transcription par rapport médical pour les transcriptionnistes du Service des archives (*i.e.* utilisation en arrière-plan).

CONTEXTE DE L'ÉVALUATION

En 2004, la région de l'Estrie avait élaboré un projet régional d'implantation de la dictée numérique avec RV pour ses centres hospitaliers. Des licences d'utilisations ont ainsi été acquises, mais n'ont pu être utilisées.

En 2008, l'UETMIS du CHUS a produit un rapport d'évaluation recommandant de n'utiliser la RV que dans les situations urgentes ou de pénuries de transcriptionnistes. La conduite d'une étude pilote avec la RV était également recommandée. Cette étude n'a cependant jamais pu être réalisée en raison d'un problème d'interface entre le système de RV et le dossier clinique informatisé (DCI) Ariane.

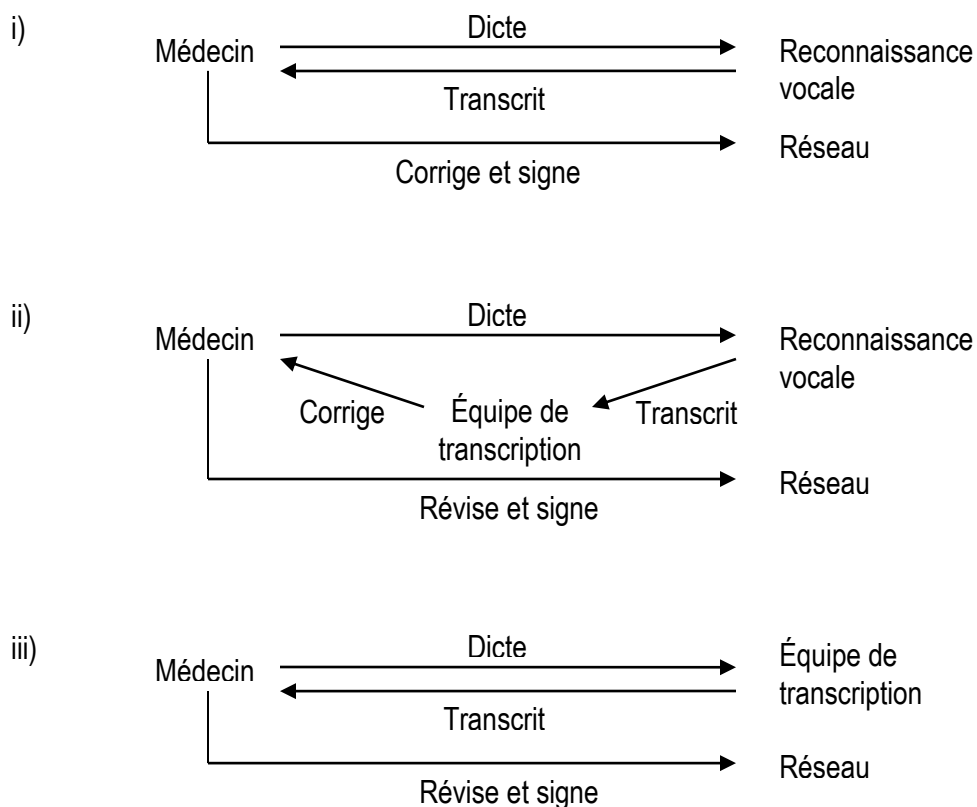


Figure 1. Principaux modes de transcription des rapports médicaux au sein d'un hôpital.

Notes: i) avant-plan (*i.e.* le médecin interagit avec un système de RV); ii) arrière-plan (*i.e.* le médecin n'interagit pas avec un système de RV); iii) système traditionnel sans système de RV.

Depuis avril 2015, l'ensemble des établissements de santé et de services sociaux de la région de l'Estrie, ainsi que ceux d'une partie de la Montérégie, ont été fusionnés. Certains de ces établissements possèdent un système de RV et la question de savoir si ce système doit être offert à l'ensemble des établissements du CIUSSS de l'Estrie – CHUS se pose. Par ailleurs, de nombreux médecins n'ont accès à aucun système de transcription et la mise à disposition d'un système de RV en avant-plan pourrait constituer une avenue intéressante pour combler ce besoin. En ce qui concerne les médecins ayant accès à un service de transcription, celui-ci est généralement assuré par le département des archives ou directement par des agentes administratives dans les départements des spécialités médicales concernées. Par contre, le CIUSSS de l'Estrie – CHUS éprouve de grandes

difficultés à recruter du personnel qualifié en transcription, avec un besoin d'embauche estimé à 90 employés pour l'année à venir. Pour faire face à cette pénurie, certains secteurs, comme l'imagerie à Cowansville, ont eu recours à de la main-d'œuvre indépendante, alors que d'autres, notamment au CHUS, ont eu recours à des heures supplémentaires. Une évaluation antérieure a également calculé un besoin supplémentaire de 72 secrétaires médicales si tous les médecins du CHUS dictaient leurs rapports de consultations externes. Actuellement, seules 2 à 3 secrétaires sont affectées à cette tâche. En outre, la planification de main-d'œuvre 1 à 5 ans évalue qu'il y aura plusieurs dizaines de secrétaires/transcriptionnistes qui partiront à la retraite, en congé de grossesse ou autres congés, dont une vingtaine uniquement au CHUS. Compte tenu de cette forte pénurie, les

délais de mise en disponibilité des rapports médicaux sont souvent bien trop longs et supérieurs à la durée souhaitée, notamment en imagerie. Finalement, l'offre de service de transcription est très inégalement répartie entre les spécialités médicales et les installations.

OBJECTIF DE L'ÉVALUATION

Procéder à une revue systématique des revues systématiques sur la performance des systèmes de RV afin de documenter s'il serait pertinent d'offrir ce service à un plus grand nombre d'installations au CIUSSS de l'Estrie – CHUS. Également identifier les barrières et facilitateurs à l'utilisation d'un tel système.

MÉTHODES

La méthodologie adoptée est celle d'une revue systématique de revues systématiques. Les moteurs de recherche Pubmed, Scopus, Cochrane Library et Center for Reviews and Dissemination (CRD) ont été consultés. Une recherche dans la littérature grise a été conduite sur Google. Les sites web consultés incluaient également les organismes d'évaluation de l'ACMTS et de l'INESSS, ainsi que les autres unités d'évaluation dans un CHU au Québec (*i.e.* McGill, CHUM, CHUQ-UL, Ste-Justine). La recension des écrits a été complétée par la lecture des références des articles inclus. La période de référence va jusqu'au 4 août 2017. Aucun critère de langue n'a été retenu.

Les mots clés et équations de recherche utilisés sont disponibles (annexe 1). Pour être incluses, les études devaient être des revues systématiques et comparer la RV avec une autre modalité de transcription de la dictée. La sélection des études et l'extraction des données ont été faites par deux évaluateurs. En cas de désaccord, un arbitrage par un troisième lecteur était prévu. L'évaluation de la qualité des études a été faite avec la grille AMSTAR (Shea et al. 2007).

Les variables d'intérêts retenues sont : spécialité médicale, technologie utilisée, utilisation en avant-plan ou en arrière-plan, taux de

reconnaissance (*i.e.* pourcentage de mots reconnus avec exactitude), taux d'erreurs et typologie, durée de la dictée et/ou de la correction, durée de la mise en disponibilité du rapport, capacité d'apprentissage du système de RV, taux de satisfaction des utilisateurs, barrières et facilitateurs à l'implantation.

Une enquête de « benchmark » a également été conduite afin d'identifier si la RV est utilisée ailleurs au Québec et de répertorier les facteurs de succès ou d'échec de son implantation. Pour ce faire, tous les responsables des unités en ETMI ont été contactés afin d'identifier des personnes ressources dans leurs établissements. Le questionnaire de l'enquête a aussi été envoyé aux responsables des services d'imagerie de ces établissements par Benoît Lebel du Service d'imagerie du CIUSSS de l'Estrie – CHUS.

RÉSULTATS ISSUS DE LA LITTÉRATURE

Les résultats de notre recherche dans les différentes bases de données ont permis de recenser 358 études sans doublon (*i.e.* incluant des revues de la littérature et des études primaires). De ce nombre, 17 revues de la littérature ont été sélectionnées pour lecture complète et 5 ont été retenues. Les résultats détaillés de notre stratégie de recherche documentaire sont fournis par le diagramme de flux PRISMA en annexe 2 et les études exclues en annexe 3.

À noter que nous avons trouvé deux rapports d'évaluation produits par les UETMIS du CHUM et du CHUS (UETMIS-CHUS 2008; DETMIS-CHUM 2012) ayant fait l'objet d'une publication sous la forme d'articles scientifiques (Poder et Bellemare 2009; Hammana et al. 2015). Ce sont ces deux articles scientifiques que nous avons évalués et dont les données ont été complétées par la lecture des rapports. De même, de nombreux écrits de la littérature grise ont été identifiés puis exclus sur la base de nos critères. Leurs contenus confirment cependant les conclusions des revues systématiques, notamment en ce qui concerne les barrières et facilitateurs à l'utilisation de la RV.

Considérant que les revues recensées indiquent une très grande hétérogénéité des résultats dans la littérature, ainsi que des définitions de variables parfois différentes (e.g. taux de reconnaissance, type d'erreurs, durée), des méta-analyses n'ont pu être conduites. Les revues recensées obtiennent des scores compris entre 3 et 8,5 sur 11 à la grille AMSTAR d'évaluation de la qualité méthodologique des revues systématiques. À noter cependant que 3 revues obtiennent un score de 8 ou plus (Poder et Bellemare 2009; Johnson et al. 2014; Hodgson et Coiera 2016). Les caractéristiques des études retenues sont fournies à l'annexe 4.

Des cinq revues recensées, celle de Hodgson et Coiera (2016) est celle dont la période de référence pour la recherche documentaire est la plus récente, avec une recherche portant jusqu'à décembre 2014. En raison de périodes de référence différentes, mais également de critères d'inclusion et d'exclusion différents (i.e. langue des études, type d'utilisation de la RV, spécialité médicale), le nombre d'études retenues dans ces revues varie de 14 (Johnson et al. 2014) à 52 (Poder et Bellemare 2009). De nombreux devis méthodologiques ont été recensés dans ces revues, le plus fréquent étant du type pré-post sans groupe de contrôle. Plusieurs études ont cependant comparé les résultats en utilisant les mêmes rapports. En général, les revues recensées qualifient la qualité de la preuve d'un niveau de faible à modéré.

Spécialités médicales

De nombreuses spécialités médicales ont été répertoriées dans la littérature comme utilisatrices de la RV. Les plus fréquemment mentionnées sont la radiologie, suivie de la pathologie, des urgences, de l'endocrinologie et de la santé mentale. Quelques études traitent également des soins dentaires, soins pédiatriques, anesthésie, consultation externe, soins infirmiers, réanimation, orthopédie et cardiologie. La radiologie est cependant la spécialité la plus citée avec près de la moitié des études qui lui sont consacrées.

Technologies utilisées

De nombreuses technologies de RV sont utilisées et le nom des modèles commercialisés varie avec le temps. Les plus cités sont le système Naturally Speaking de Dragon, ceux de Via Voice et MedSpeak d'IBM et le système Voice Express de L&H. D'autres produits des compagnies Phillips, Microsoft, LTI et Agfa sont aussi cités. Certaines études se sont attachées à comparer quelques-uns de ces systèmes. Ces comparaisons sont cependant trop anciennes pour être considérées d'actualité. Par ailleurs, il est à noter que les technologies de RV utilisées étaient le plus souvent en langue anglaise, bien que quelques études dans une autre langue aient été identifiées (i.e. allemand, chinois, danois, finnois, français, italien, japonais).

Type d'utilisation

En grande majorité, les études recensées font référence à l'utilisation de la RV en avant-plan. Les études en arrière-plan sont moins nombreuses et sont généralement de faible qualité, ce qui ne permet pas de conclure sur les avantages et inconvénients de la RV (Poder et Bellemare, 2009). À noter que seule la revue systématique de Hodgson et Coiera (2016) s'est attachée à n'inclure que des études utilisant la RV en avant-plan.

Taux de reconnaissance

Les taux de reconnaissances répertoriés dans la littérature sont très variables, allant de 84,5 % jusqu'à 99 % (Poder et Bellemare 2009; Ajami 2016). À noter cependant que la plupart des études rapportent des taux compris entre 92 % et 98 %. Comparativement à une transcription effectuée par des transcriptionnistes, le taux de reconnaissance est cependant plus bas, avec des taux de 97 % à 99,97 % pour les transcriptionnistes (Poder et Bellemare 2009; Johnson et al. 2014). On constate également que les taux de reconnaissances progressent peu dans le temps, avec un taux d'amélioration de seulement 0,03 % par an (Hodgson et Coiera 2016).

Taux d'erreurs et leur typologie

Selon la revue la plus récente de Hodgson et Coiera (2016), le taux d'erreur avec la RV est de 0,05 à 6,66 par rapport, alors qu'avec un transcriptionniste, ce taux est de 0,02 à 0,4. Ces auteurs indiquent également que le nombre d'erreurs supplémentaires avec la RV varie de 0,03 à 19,53 par rapport. Selon Hammana et al. (2015), le pourcentage de rapport présentant au moins une erreur est quant à lui de 4,8 % à 89 % avec la RV contre 2,1 % à 22 % avec les transcriptionnistes. Johnson et al. (2014) rapportent pour leur part que la RV produit entre 2 à 16 fois plus d'erreurs que le système de transcription traditionnel. Par contre, les erreurs grammaticales et orthographiques sont réduites avec la RV, jusqu'à 3 fois moins selon Poder et Bellemare (2009). Avec la RV, les erreurs les plus fréquentes sont associées à des substitutions et omissions de mots, ainsi qu'à des insertions inadéquates. Dans certaines études, on note également des phrases sans sens et des problèmes de ponctuation (Hodgson et Coiera 2016). De ce fait, il est considéré dans la grande majorité des études répertoriées que les erreurs majeures sont plus fréquentes avec la RV qu'avec un système de transcription traditionnel. Dans la revue de Hodgson et Coiera (2016), il apparaît que les erreurs majeures seraient environ 3 fois plus nombreuses avec la RV. Johnson et al. (2014) rapportent que ces erreurs graves avec la RV pourraient concerner jusqu'à 7,8 % des rapports.

Durée de la dictée et/ou de la correction

La plupart des études rapportent un allongement de la durée de la dictée avec la RV (Hammana et al. 2015). Il existe cependant une grande variabilité des résultats, allant de -61 % à +190 % (Poder et Bellemare 2009). Le temps de correction en avant-plan est encore plus long avec une variation allant de -16,5 % à +1876 % pour le médecin (Hodgson et Coiera 2016). Lorsque réalisée en arrière-plan, la variabilité est également grande, allant de -44 % à + 28 % (Poder et Bellemare 2009). Si Ajami (2016) et Johnson et al. (2014) ne rapportent pas de données chiffrées, ils indiquent également que la

RV augmente souvent la durée de dictée et de correction, surtout lorsque ce sont les médecins qui effectuent la correction du rapport généré.

Durée de mise en disponibilité

À la différence de la durée de dictée et/ou de correction, la durée de mise en disponibilité est considérablement réduite avec la RV. Lorsque utilisée en avant-plan, on parle d'une réduction de 16 % à 99,5 % (Poder et Bellemare 2009; Johnson et al. 2014; Hammana et al. 2015; Hodgson et Coiera 2016). En arrière-plan, les résultats sont également positifs avec une réduction de 36 % des délais. À noter que dans certaines études, le passage à la RV en avant-plan a permis de réduire cette durée de plusieurs jours à quelques minutes seulement, ce qui semble représenter le principal avantage de la RV.

Capacité d'apprentissage

Seule la revue systématique de Poder et Bellemare (2009) fait référence à la capacité d'apprentissage du système de RV. Ces auteurs ont pu constater que la capacité d'apprentissage du système est assez forte, avec une amélioration du taux de reconnaissance pouvant atteindre 11 à 13 points de pourcentage supplémentaires après un maximum de 4 adaptations, avant de se stabiliser. Une adaptation est définie comme un processus où le médecin doit lire un certain nombre de mots à voix haute. À noter toutefois qu'avec certains nouveaux systèmes de RV, les utilisateurs n'ont plus à lire une liste de mots. Cette liste a été remplacée par le téléchargement de rapports médicaux existants. Le système de RV s'ajuste ensuite au fur et à mesure des corrections effectuées par l'utilisateur ou le transcriptionniste.

Concernant la durée d'apprentissage du système de RV par les utilisateurs, Poder et Bellemare (2009) et Johnson et al. (2014) mentionnent que cela peut aller de quelques minutes jusqu'à 6 heures pour la 1^{ère} formation. Des ajustements sont cependant nécessaires pendant plusieurs semaines afin que l'utilisateur se familiarise avec la RV et que celle-ci s'adapte aux nouveaux mots

utilisés. Selon ces mêmes auteurs, avant que les utilisateurs ne soient familiers avec la RV et obtiennent des résultats satisfaisants, cela peut prendre de 2 à 6 mois. Cette durée varie cependant avec l'habileté des utilisateurs et le type de matériel utilisé (e.g. système de RV, micro, carte son) (Poder et Bellemare 2009).

Taux de satisfaction

La revue systématique de Johnson et al. (2014) indique des taux de satisfaction positif et négatif avec la RV très similaires (i.e. 44 % vs. 46 %). Selon Hammana et al. (2015), le taux de satisfaction avec la RV dépend d'un arbitrage à faire entre le gain pour le département (i.e. mise en disponibilité plus rapide du rapport) et la perte pour le médecin (i.e. temps de dictée et de correction plus long).

Barrières

De nombreuses barrières à l'utilisation de la RV sont relevées. Certaines peuvent cependant être évitées, notamment celles liées à la technologie utilisée. Par exemple, les bruits de fond peuvent être évités en utilisant un local calme ainsi qu'un micro de bonne qualité et bien placé. De même, si les erreurs ne sont pas immédiatement corrigées, le logiciel de RV considérera alors les termes utilisés comme exacts pour les prochaines utilisations. Par contre, certaines barrières seront plus difficiles à lever, telles que l'accent de l'utilisateur, la faible expérience des médecins (i.e. résidents), leur intérêt et capacité à utiliser la RV (selon Kopach et al. (2005), seuls 50 % des médecins seraient habiles à utiliser la RV), la surcharge cognitive demandée à l'utilisateur et la surcharge de travail si utilisée en avant-plan, le vocabulaire et la complexité grammaticale utilisée, les interruptions fréquentes dans la dictée, ainsi que le temps requis avant d'arriver à une utilisation satisfaisante de la RV. À noter qu'en arrière-plan, un obstacle supplémentaire est lié au fait que pour les transcriptionnistes, il ne s'agit plus du même type de travail et donc qu'un temps d'adaptation peut être nécessaire (David et al. 2009).

Facilitateurs

D'un point de vue technologique, la performance du système de RV peut être améliorée grâce à une bonne carte son et à une bonne carte mémoire. Télécharger des rapports médicaux déjà existants dans le logiciel de RV devrait permettre un meilleur taux de reconnaissance dès le départ. Pour l'utilisateur, une diction claire et une vitesse de prononciation adéquate, ainsi que des phrases synthétiques et bien structurées, sont des éléments favorables à une bonne utilisation de la RV. Un vocabulaire répétitif est également un facilitant, ce qui peut expliquer pourquoi certaines spécialités comme la radiologie utilisent la RV avec succès alors les résultats dans d'autres spécialités médicales sont moins positifs (e.g. urgence, psychiatrie). Une bonne formation à la RV est aussi considérée comme facilitant, de même qu'une utilisation sélective selon le type de rapport (e.g. complexité) ou le moment de la journée (Hammana et al. 2015; Johnson et al. 2014).

À noter que plusieurs avantages de la RV constituent de par leur nature des facilitateurs à son adoption. En particulier le fait que le rapport est immédiatement disponible pour être téléversé dans le dossier électronique du patient, ce qui représente un gain à la fois pour le médecin utilisateur, mais aussi pour ses collègues et le patient traité. De plus, avec la RV en avant-plan, il n'y a plus besoin de se « replonger » dans le rapport médical plusieurs jours après qu'il ait été transcrit par les transcriptionnistes, ce qui parfois exigeait du médecin qu'il réexamine entièrement le dossier du patient. La RV peut aussi être très adaptée quand les yeux et les mains sont occupés à d'autres tâches. De même, elle peut être utilisée pour accélérer le processus de facturation et de paiements des honoraires. En outre, si les transcriptionnistes sont considérés comme peu rapides, l'utilisation de la RV pourrait leur permettre de gagner en efficacité (Johnson et al. 2014).

Finalement, considérant que la RV augmente le temps de dictée et de correction pour les médecins, ces derniers pourraient être tentés de

produire des rapports plus concis. Si aucune information importante n'est perdue, le fait de forcer à produire des rapports plus concis et parfois plus clairs pourrait être un facilitant à l'adoption de la RV (Poder et Bellemare 2009). De même, l'utilisation de gabarit (i.e. macro/template) est rapportée dans la littérature comme un élément pouvant permettre de produire des rapports médicaux plus rapidement (Poder et Bellemare 2009; Johnson et al. 2014; Ajami 2016).

RÉSULTATS ISSUS DU BENCHMARK

Les résultats du benchmark effectués à l'extérieur du CIUSSS de l'Estrie – CHUS indiquent que la grande majorité des établissements de santé au Québec n'utilisent pas la RV pour la dictée et la transcription des rapports médicaux. Sur un total de 56 établissements ou installations pour lesquels nous avons pu collecter de l'information, seuls 7 disposaient de la RV. Ces 7 établissements ou installations sont le Centre hospitalier universitaire de Montréal (CHUM), le Centre hospitalier universitaire de Québec – Université Laval (CHUQ-UL), l'Hôpital général juif (HGJ), le Centre universitaire de santé McGill (CUSM), l'Hôpital Anna Laberge (HAL) à Chateauguay, l'Institut universitaire de cardiologie et de pneumologie de Québec (IUCPQ) et l'Hôpital de Rivière-du-Loup (HRL). La plupart de ces établissements utilisent la RV en arrière-plan et en grande majorité pour l'imagerie médicale. Ceci dit, d'autres spécialités utilisent aussi la RV (i.e. urgence, consultation en orthopédie, chirurgie générale, neurochirurgie, urologie, gastro-entérologie, médecine nucléaire). Il existe également une utilisation plus restreinte de la RV en avant-plan au CHUM et à HAL. Les systèmes de RV utilisés sont principalement ceux de Crescendo (Medirad, MedRite, Centro) et de Comdic (Fusion, Post-script). Précisons ici que HRL utilise aussi Speechmagic et Nuance dragon en plus de Crescendo. Les personnes que nous avons contactées dans ces établissements indiquent toutes que la RV en arrière-plan permet d'augmenter l'offre de service en transcription grâce à un gain de productivité

chez les transcriptionnistes. Selon les établissements, ce gain varie de 10 à 30 %. En avant-plan, le gain se situe au niveau de la mise en disponibilité immédiate du rapport médical, mais nécessite de nombreuses corrections de la part de l'utilisateur. À noter que 2 établissements sur les 56 ont indiqué avoir connu une expérience négative avec la RV (Centro de Crescendo), ce qui les a conduit à l'abandonner. Selon les informations disponibles, les raisons invoquées étaient l'absence d'un des modules spécifiques permettant au système de fonctionner adéquatement, ainsi qu'un trop grand nombre d'erreurs de transcription. Un établissement rapporte aussi une faible autonomie pour apporter des modifications au lexique, ce qui en limite l'usage.

Concernant les conditions gagnantes permettant une utilisation réussie de la RV, plusieurs ont été fournies (e.g. soutien massif et continu de la compagnie de RV et du service informatique de l'établissement, désigner un chargé de projet, dispenser une formation de qualité, être proactif avec les utilisateurs, donner des consignes claires et faire des retours sur ce qui fonctionne ou non, cibler les médecins les plus aptes à la RV ainsi que les plus gros utilisateurs, faire une implantation graduelle, favoriser les rapports courts, rehausser et uniformiser le matériel informatique). Spécifiquement à l'utilisation de gabarits ou rapports synoptiques, seul le CHUQ-UL en utilise dans certains cas, alors que HRL indique avoir essayé sans succès.

DISCUSSION

La réalisation de rapports médicaux connaît depuis plusieurs décennies une nette tendance à la hausse (Van Therheyden, 2005). Cette forte demande est alimentée par l'accroissement d'une population vieillissante nécessitant un grand nombre de tests et de traitements médicaux, et par le besoin continu des différents acteurs du système de santé de partager ces informations médicales (e.g. médecins référents, organismes de remboursements). L'offre de transcription humaine pourrait bientôt ne plus être suffisante pour couvrir ce flux, ce qui

démontre le besoin de faire appel à des technologies de substitution.

Cette revue systématique des revues systématiques rapporte que la RV pourrait permettre de contribuer au processus de documentation médicale et de supporter les établissements de santé et de services sociaux dans leurs obligations légales. Si les données scientifiques apparaissent comme étant de faibles qualités, elles semblent cependant converger en soulignant que la RV n'est pas une solution miracle, que celle-ci produit davantage d'erreurs que la transcription humaine, qu'elle demande un effort supplémentaire de la part des utilisateurs, à la fois pour la dictée et la correction, mais qu'elle permet un résultat immédiat lorsqu'utilisée en avant-plan, c'est-à-dire un rapport médical immédiatement disponible. C'est donc dans cette rapidité de mise en disponibilité que réside l'avantage principal de la RV. Par ailleurs, lorsque les utilisateurs n'avaient pas auparavant accès à un service de transcription, la RV pourrait éventuellement leur permettre d'avoir accès à ce service, pour autant qu'ils soient prêts à corriger les erreurs générées.

Avec un système de reconnaissance vocale, arriver à un résultat sans erreur de transcription peut donc s'avérer être long et coûteux (Pezzullo et al, 2008). Cependant, ne procéder à aucune correction ou à une correction minimale avec la RV 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 (*i.e.* patients, tierces médecins, gestionnaires) et que cela pourrait être interprété comme un signe de travail mal fait ou d'un manque d'intérêt de la part du médecin. De plus, d'un point de vue médico-légal, chaque rapport médical doit être vérifié en détail par son auteur ou un correcteur médical humain.

Les nouvelles études publiées depuis décembre 2014 et la dernière recension des écrits de Hodgson et Coiera (2016) ne semblent pas indiquer une amélioration des performances de la RV. Par exemple, les études de du Toit et al. (2015) et de Motyer et al. (2016) indiquent que la RV continue à générer des erreurs majeures dont l'interprétation pourrait être potentiellement dommageable pour les patients. L'étude de Prevedello et al. (2014) indique pour sa part que les durées de mise en disponibilité des rapports médicaux sont considérablement réduites avec la RV. Ces résultats confirment ainsi la conclusion de Hodgson et Coiera (2016) selon laquelle il existe depuis le début des années 2000 autant d'études indiquant des résultats positifs que négatifs. Un arbitrage est donc nécessaire entre ses bénéfices et ses inconvénients avant de statuer sur son éventuelle implantation et ce d'autant plus qu'il est difficile, voire impossible, de déterminer dans quel cadre clinique et pour quelles tâches la RV serait la plus adaptée et la plus bénéfique.

Il est d'autant plus difficile d'apporter une conclusion claire sur la pertinence d'utiliser la RV que les nombreuses études répertoriées présentent un fort niveau d'hétérogénéité, principalement en raison des types de rapports produits et des différentes définitions utilisées pour mesurer la productivité ou les types d'erreurs. À cet égard, la revue systématique de Hammana et al. (2015), bien que centrée uniquement sur la radiologie, indique que cette hétérogénéité reste présente et qu'elle est dans ce cas-ci principalement due aux modalités d'imagerie évaluées et à la façon de mesurer les issues.

Combinée avec d'autres technologies (*e.g.* Picture Archiving and Communication System (PACS), dictionnaire terminologique), la RV a également le potentiel de générer automatiquement des enregistrements cliniques standardisés et d'interagir dynamiquement avec les systèmes d'information clinique pour améliorer la prise de décision clinique et le temps de diagnostic. Elle pourrait également accroître le

flux de patients et réduire les délais d'attente de ces derniers. Cependant, cela reste à démontrer.

En ce qui concerne le point particulier de la langue utilisée par le système de RV, aucune étude ne s'est attachée à identifier si la performance de la RV pouvait être amoindrie dans une langue autre que l'anglais. Par exemple, seules deux études ont été publiées avec un système de RV utilisant le français (Capel et al. 2004; Happe et al. 2003). Les conclusions de ces études semblent cependant indiquer des résultats semblables à ceux répertoriés dans notre revue systématique des revues systématiques. Par contre, les expériences issues de notre benchmark au Québec (*i.e.* ICM, Drummondville) pourraient indiquer l'existence de certaines spécificités au Français (*i.e.* spectre sonore) qui viendrait limiter la pertinence d'utiliser la RV.

CONCLUSION

La RV présente plusieurs avantages, mais également quelques inconvénients. Son principal avantage, lorsque utilisée en avant-plan, réside dans la réduction considérable de la durée de mise en disponibilité des rapports médicaux dans le DCI, ainsi que des économies de coûts potentielles, puisque le service de transcription n'aurait plus à traiter ces rapports. Par contre, ce gain se fait en contrepartie d'une baisse de la productivité des utilisateurs dans la mesure où le temps de dictée et de correction est augmenté. Il est également à noter qu'un grand nombre de conditions doivent être réunies pour parvenir à une utilisation satisfaisante de la RV, que cela soit en avant ou en arrière-plan. Sous cet angle, la RV pourrait être une avenue intéressante pour le CIUSSS de l'Estrie – CHUS si un plan adéquat d'implantation est établi (*e.g.* formation, qualité du matériel) et s'il démontre suffisamment de flexibilité pour s'assurer l'adhésion de ses utilisateurs (*e.g.* utilisation libre, possibilité de basculer en arrière-plan). Finalement, un résumé des principaux avantages et inconvénients de la RV est présenté sous l'angle des 6 axes de la performance du CIUSSS de l'Estrie – CHUS (voir Figure 2).

REMERCIEMENTS

Nous remercions l'ensemble de nos partenaires des UETMI de la province du Québec pour l'aide apportée durant le benchmark, Dre Colette Bellavance, Normand Bilodeau, Mélanie Boisvert, Dre Édith Grégoire, Maryse Lachance, Benoît Lebel et Line Ménard pour leur collaboration, Dre Véronique Déry pour son travail de relecture et sa contribution à la préparation de cette note de synthèse, ainsi que Mariève Desrochers-Rancourt et Mathieu Roy pour la mise en page et correction du texte.

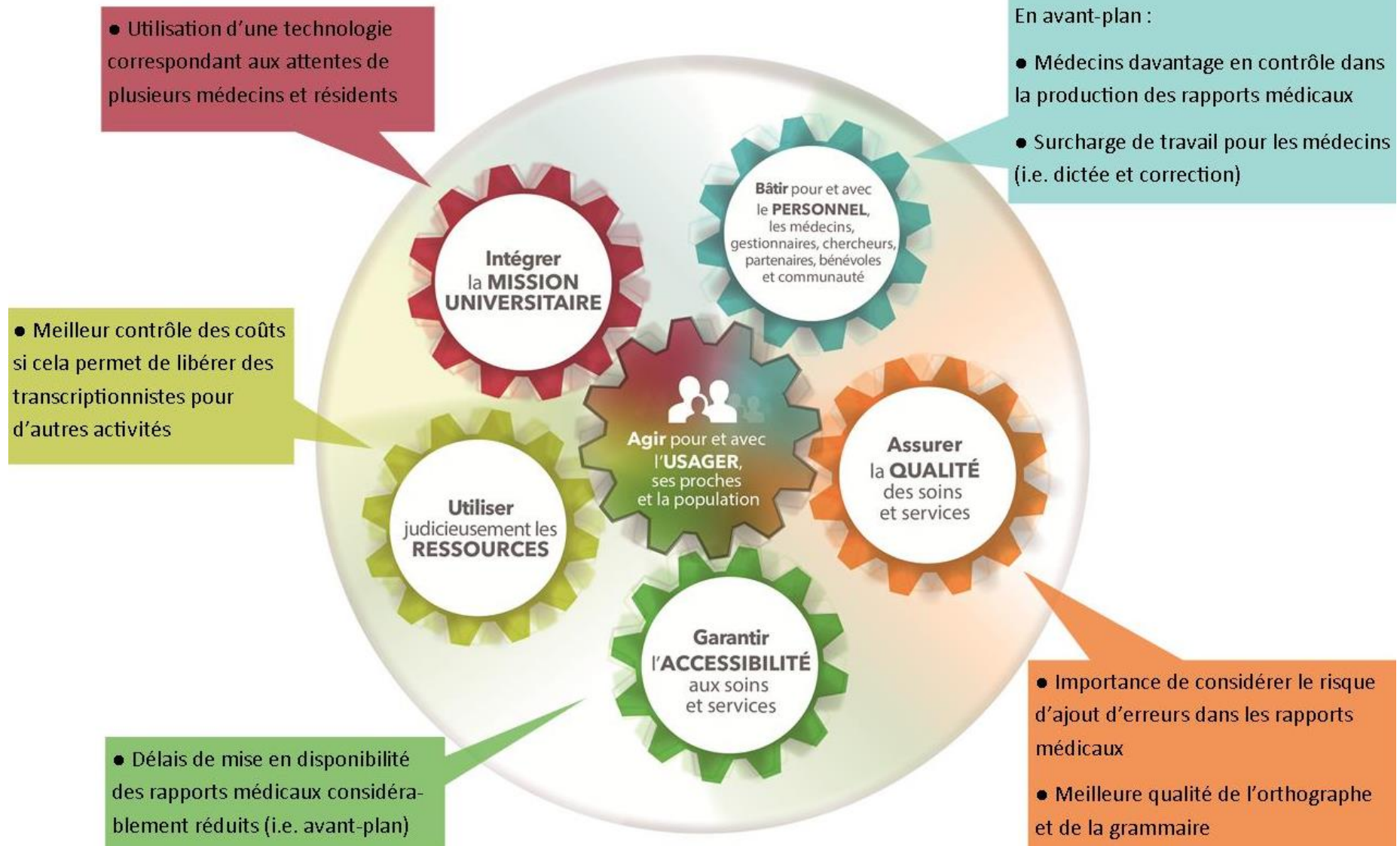


Figure 2. Impact de l'utilisation de la RV sur les 6 axes de la performance du CIUSSS de l'Estrie – CHUS (adapté du document interne sur « L'identité du CIUSSS de l'Estrie – CHUS », octobre 2016)

RÉFÉRENCES

- Ahlgren, C., Maenner, O., Baumstark, M.W. 2016. Introduction of digital speech recognition in a specialised outpatient department: a case study. *BMC Medical Informatics and Decision Making*, 16:132.
- Ajami, S. 2016. Use of speech-to-text technology for documentation by healthcare providers. *The National Medical Journal of India*, 29(3):148-152.
- Alapetite, A., Andersen, H.B., Hertzum, M. 2009. Acceptance of speech recognition by physicians: A survey of expectations, experiences and social influence. *International Journal of Human-Computer Studies*, 67:36–49.
- 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*, 23:344-348.
- David, G.C., Garcia, A.C., Warfield Rawls, A., Chand, D. 2009. Listening to what is said – transcribing what is heard: the impact of speech recognition technology (SRT) on the practice of medical transcription (MT). *Sociology of Health & Illness*, 31(6):924–938.
- DETMIS-CHUM. 2012. L'utilisation de la reconnaissance vocale au CHUM. Préparé par Imane Hammana et Luigi Lepanto, Direction de l'évaluation des technologies et des modes d'intervention en santé (DETMIS), Centre hospitalier de l'Université de Montréal, QC, Canada.
- du Toit, J., Hattingh, R., Pitcher, R. 2015. The accuracy of radiology speech recognition reports in a multilingual South African teaching hospital. *BMC Medical Imaging*, 15:8.
- Hodgson, T., Coiera, E. 2016. Risks and benefits of speech recognition for clinical documentation: a systematic review. *Journal of the American Medical Informatics Association*, 23(e1):e169-79.
- Hammana, I., Lepanto, L., Poder, T., Bellemare, C., Ly. M.-S. 2015. Speech recognition in the radiology department: a systematic review. *Health Information Management Journal*, 44(2):4-10.
- Happe, A., Pouliquen, B., Burgun, A., Cuggia, M., Le Beux, P. 2003. Automatic concept extraction from spoken medical reports. *International Journal of Medical Informatics*, 70:255-263.
- Johnson, M., Lapkin, S., Long, V., Sanchez, P., Suominen, H., Basilakis, J., Dawson, L. 2014. A systematic review of speech recognition technology in health care. *BMC Medical Informatics and Decision Making*, 14:94.
- Kopach, R., Sadat, S., Gallaway, I.D., Geiger, G., Ungar, W.J., Coyte, P.C. 2005. Cost-effectiveness analysis of medical documentation alternatives. *International journal of technology Assessment in Health care*, 21(1):126-131.
- Leeming, B.W., Porter, D., Jackson, J.D., Bleich, H.L., Simon, M. 1981. Computerized radiologic reporting with voice data-entry. *Radiology*, 138(3):585–588.
- Mehta, A., McLoud, T.C. 2003. Voice recognition. *J Thorac Imaging*, 18(3):178–182.
- Motyer, R.E., Liddy, S., Torreggiani, W.C., Buckley, O. 2016. Frequency and analysis of non-clinical errors made in radiology reports using the National Integrated Medical Imaging System voice recognition dictation software. *Irish Journal of Medical Science*, 185:921-927.
- Pezzullo, J.A., Tung, G.A., Rogg, J.M., Davis, L.M., Brody, J.M., Mayo-Smith, W.W. 2008. Voice recognition dictation: radiologist as transcriptionist. *J Digit Imaging*, 21(4):384–389.
- Poder, T., Bellemare, C. 2009. Émergence d'une nouvelle technologie au service de la gestion des hôpitaux : les systèmes de reconnaissance vocale. *Journal d'Économie Médicale*, 27(5):329-339.
- Prevedello, L.M., Ledbetter, S., Farkasa, C., Khorasani, R. 2014. Implementation of Speech Recognition in a Community-based Radiology Practice: Effect on Report Turnaround Times. *J Am Coll Radiol*, 11:402-406.
- Robbins, A.H., Horowitz, D.M., Srinivasan, M.K., Vincent, M.E., Shaffer, K., Sadowsky, N.L., Sonnenfeld, M. 1987. Speech-controlled generation of radiology reports. *Radiology*, 164(2):569–573.
- Shea, B.J., Grimshaw, J.M., Wells, G.A., Boers, M., Andersson, N., Hamel, C., Porter, A.C., Tugwell, P., Moher, D., Bouter, L.M., 2007. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med. Res. Methodol.* 7:10.
- UETMIS-CHUS. 2009. Efficience de la reconnaissance vocale au CHUS. Préparé par Thomas Poder et Christian Bellemare, Unité d'évaluation des technologies et des modes d'intervention, Centre hospitalier de l'Université de Sherbrooke, QC, Canada.
- Van Terheyden, N. 2005. Is speech recognition the Holy Grail? *Health Management Technology*, 26(2):42-45.

L'Unité d'évaluation des technologies et des modes d'intervention en santé et services sociaux (UETMISSS) du CIUSSS de l'Estrie – CHUS a pour mission de soutenir et conseiller les décideurs du CIUSSS de l'Estrie – CHUS et du RUIS de l'Université de Sherbrooke dans la prise de décision par la synthèse et la production de connaissances pour l'utilisation efficiente des ressources et l'amélioration de la qualité des soins et des services aux patients et de contribuer à la mission universitaire du CIUSSS de l'Estrie – CHUS par ses pointes d'excellence, ses activités de transfert de connaissances et ses partenariats.

<http://www.chus.qc.ca/professionnels/evaluation-des-technologies/>

Annexe 1. Mots clés et stratégies de recherche documentaire

PubMed :

(speech AND recognition AND medical AND (report OR record OR system OR dictation OR dictating)) OR (voice AND recognition AND medical AND (report OR record OR system OR dictation OR dictating)) OR (vocal AND recognition AND medical AND (report OR record OR system OR dictation OR dictating)) OR (reconnaissance AND vocale) AND (error OR workflow OR turnaround OR productivity)

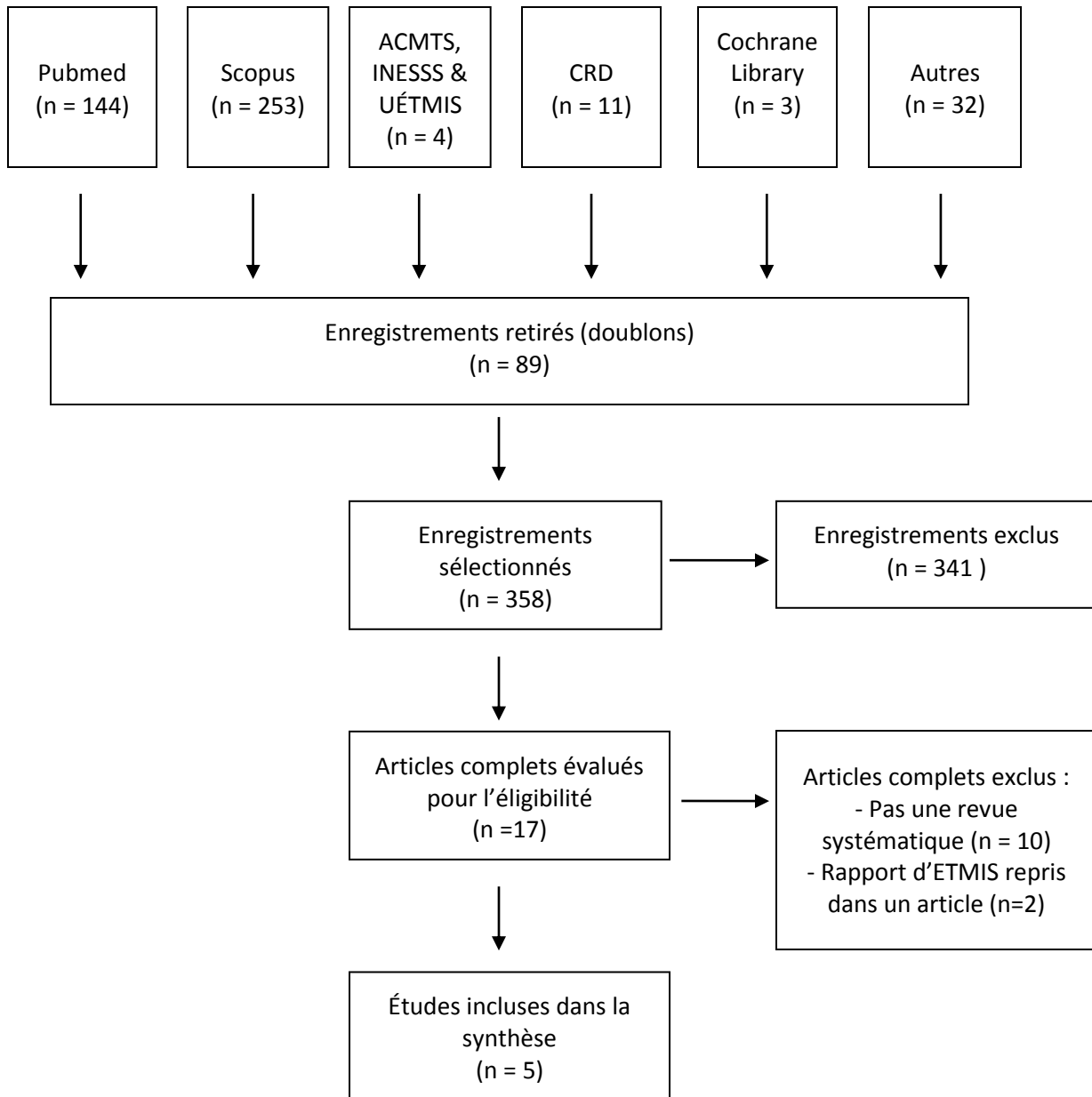
Scopus :

(TITLE-ABS-KEY(speech) AND TITLE-ABS-KEY(recognition) AND TITLE-ABS-KEY(medical) AND (TITLE-ABS-KEY(report) OR TITLE-ABS-KEY(record) OR TITLE-ABS-KEY(system) OR TITLE-ABS-KEY(dictation) OR TITLE-ABS-KEY(dictating))) OR (TITLE-ABS-KEY(voice) AND TITLE-ABS-KEY(recognition) AND TITLE-ABS-KEY(medical) AND (TITLE-ABS-KEY(report) OR TITLE-ABS-KEY(record) OR TITLE-ABS-KEY(system) OR TITLE-ABS-KEY(dictation) OR TITLE-ABS-KEY(dictating))) OR (TITLE-ABS-KEY(vocal) AND TITLE-ABS-KEY(recognition) AND TITLE-ABS-KEY(medical) AND (TITLE-ABS-KEY(report) OR TITLE-ABS-KEY(record) OR TITLE-ABS-KEY(system) OR TITLE-ABS-KEY(dictation) OR TITLE-ABS-KEY(dictating))) OR (TITLE-ABS-KEY(reconnaissance) AND TITLE-ABS-KEY(vocale)) AND (TITLE-ABS-KEY(error) OR TITLE-ABS-KEY(workflow) OR TITLE-ABS-KEY(turnaround) OR TITLE-ABS-KEY(productivity))

CRD & Cochrane Library & autres sites web :

Speech recognition OR Voice recognition OR Reconnaissance vocale

Annexe 2. Diagramme de flux PRISMA au 4 août 2017



Source: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097 [Traduction libre]

Annexe 3. Études exclues

- Bajaj, S. 2004. Expanding the Reach of Speech Recognition. Extending speech technology benefits to transcription workflow can boost productivity. *Health Management Technology* 25(3):16,20-2.
- Branstetter, B.F., Kim, W., Prevedello, L.M., Weiss, D.L. 2013. Reporting and Communication. In: *IT Reference Guide for the Practicing Radiologist*. American College of Radiology, pp. 1-21. Disponible au : <https://www.acr.org/~/media/ACR/Documents/PDF/Advocacy/IT%20Reference%20Guide/IT%20Ref%20Guide%20Reporting%20Communication.pdf>
- CADTH 2011. The Use of Voice Recognition Tools for Diagnostic Imaging Management: Clinical Evidence and Cost-Effectiveness. Rapid Response Report: Summary of Abstracts. Canadian Agency for Drugs and Technologies in Health, Ottawa, Canada.
- CADTH 2012. Voice Recognition Documentation Systems: Clinical Effectiveness. Rapid Response Report: Summary of Abstracts. Canadian Agency for Drugs and Technologies in Health, Ottawa, Canada.
- DETMIS-CHUM. 2012. L'utilisation de la reconnaissance vocale au CHUM. Préparé par Imane Hammana et Luigi Lepanto, Direction de l'évaluation des technologies et des modes d'intervention en santé (DETMIS), Centre hospitalier de l'Université de Montréal, QC, Canada.
- Fox, M.A., Aschkenasi, C.J., Kalyanpur, A. 2013. Voice recognition is here comma like it or not period. *Indian Journal of Radiology and Imaging* 23(3):191-4.
- Herman, S.J. 2004. Speech recognition and the creation of radiology reports. *Applied Radiology* 5;23-28.
- Lackey, A.E., Pandey, T., Moshiri, M., Lalwani, N., Lall, C., Bhargava, P. 2014. Productivity, Part 2: Cloud Storage, Remote Meeting Tools, Screencasting, Speech Recognition Software, Password Managers, and Online Data Backup. *J Am Coll Radiol* 11:580-588.
- Shagoury, J. 2010. Dr. "Multi-Task": Using Speech to Build Up Electronic Medical Records While Caring for Patients. In: A. Neustein (ed.), *Advances in Speech recognition: Mobile Environments, Call Centers and Clinics*. Chapter 11, pp. 247-73. Springer, Boston, MA.
- 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);54-55.
- UETMIS-CHUS. 2009. Efficience de la reconnaissance vocale au CHUS. Préparé par Thomas Poder et Christian Bellemare, Unité d'évaluation des technologies et des modes d'intervention, Centre hospitalier de l'Université de Sherbrooke, QC, Canada.
- Weiss, D.L., Kim, W., Branstetter, B.F., Prevedello, L.M. 2014. Radiology Reporting: A Closed-Loop Cycle from Order Entry to Results Communication. *J Am Coll Radiol* 11:1226-1237.

Annexe 4. Caractéristiques des études incluses

Auteurs	Période de recension	Spécialités médicales	Utilisation en avant ou arrière-plan	Langue(s) utilisée(s)	Nombre d'études incluses	Score AMSTAR
Poder et Bellemare (2009)	01-1992 à 03-2008	Multiples	Les deux	All, Ang, Chin, Fra, Ita, Jap.	52	8
Johnson et al. (2014)	2000 à 2012	Multiples	Les deux	Ang, Fin, Dan	14	8
Hammana et al. (2015)	01-1992 à 10-2013	Radiologie	Les deux	Non précisé	20	6,5
Ajami (2016)	01-1990 à 07-2014	Multiples	Non précisé	All, Ang, autres	42	3
Hodgson et Coiera (2016)	Jusqu'à 12-2014	Multiples	Avant-plan	Non précisé	23	8,5

Notes : All : allemand; Ang : anglais; Chin : chinois; Fra : français; Ita : italien ; Jap : japonais; Fin : finnois; Dan : danois.