



Des gens de soins et d'excellence

Centre hospitalier universitaire
de Sherbrooke

Unité d'évaluation des technologies et des modes d'intervention en santé



Évaluation d'un système de téléradiologie composé d'un numériseur de films





Des gens de soins et d'excellence

*CENTRE HOSPITALIER UNIVERSITAIRE DE SHERBROOKE
Hôpital Fleurimont, 3001, 12^e Avenue Nord, Sherbrooke (Québec) J1H 5N4
Hôtel-Dieu, 580, rue Bowen Sud, Sherbrooke (Québec) J1G 2E8
Téléphone : (819) 346-1110*

*Unité d'évaluation des technologies et de modes
d'intervention en santé*

ÉVALUATION D'UN SYSTÈME DE TÉLÉRADIOLOGIE COMPOSÉ D'UN NUMÉRISSEUR DE FILMS

10 février 2004

© *Unité ÉTMIS 2004*

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 Etienne-Le Bel du CHUS.

LES MEMBRES

D^r Maurice Roy, M.D., LL. M., M.A.P., directeur

Directeur général adjoint et directeur des Services professionnels du CHUS.

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

Physicien médical, Direction des services professionnels du CHUS

Christian Bellemarre, M. Sc., conseiller en évaluation des technologies

Mathématicien & informaticien, Unité ÉTMIS du CHUS.

AVANT-PROPOS

ÉVALUATION D'UN SYSTÈME DE TÉLÉRADIOLOGIE COMPOSÉ D'UN NUMÉRISSEUR DE FILMS

Le transfert à distance des images radiologiques numériques se voit comme un moyen d'améliorer le service de radiologie diagnostique à la clientèle des centres hospitaliers régionaux. Les grands centres hospitaliers à vocation tertiaire et quaternaire possèdent des départements de radiologie bien établis pouvant offrir une expertise unique au niveau diagnostique de première ligne. Cette expertise se reflète aussi au niveau de l'aide au diagnostic de deuxième ligne à d'autres cliniciens tels les urgentologues et ceux qui nécessitent une seconde opinion.

Le Centre de santé du Granit et l'Agence de développement des réseaux locaux de santé et des services sociaux (ADRLSSS) de l'Estrie ont demandé à l'Unité ÉTMIS du CHUS d'évaluer un système d'imageries numériques peu coûteux ayant la capacité d'envoyer des images radiographiques de qualité suffisante afin d'effectuer une lecture diagnostique à distance. Les demandeurs avaient formulé les demandes suivantes :

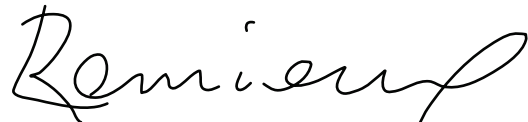
1. un coût suffisamment bas pour permettre l'acquisition de l'équipement par les plus petits centres de santé de la région;
2. utilisation des recommandations pour définir des critères d'acquisition de technologies similaires à l'intérieur d'un processus d'appel d'offres;
3. exportabilité de la technologie approuvée à l'ensemble du Réseau de la santé et des services sociaux du Québec.

Même si le PACS (*Picture Archiving and Communication System*) est de plus en plus utilisé en milieu hospitalier, de nombreux centres de santé régionaux et cliniques n'ont pas la capacité financière d'en acquérir un. Par contre, ces centres peuvent faire l'acquisition d'équipements de numérisation des images radiologiques dans le but de les transmettre par réseau informatique à des centres hospitaliers qui donnent des services de téléradiologie diagnostique. C'est dans ce contexte que cette recherche a été orientée.

En remettant ce rapport, l'Unité ÉTMIS du CHUS souhaite apporter aux demandeurs les éléments d'informations nécessaires leur permettant de conseiller adéquatement les centres de santé qui désirent acquérir une technologie efficace, mais peu coûteuse, pour développer ainsi un service de base de téléradiologie.



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



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

REMERCIEMENTS

Ce rapport a été préparé à la demande de l'Agence de développement des réseaux locaux de santé et de services sociaux de l'Estrie par Renald Lemieux, M. Ing., Ph. D., M.ETS., coordonnateur de l'Unité d'évaluation des technologies et des modes d'intervention en santé du Centre hospitalier universitaire de Sherbrooke.

L'Unité tient à remercier

1. L'équipe de recherche :

Christian Blais, M.D., radiologiste, CHUS.

Renald Lemieux, M. Ing., Ph. D., M.ETS., coordonnateur, Unité ÉTMIS, CHUS

Marc Verreault, M. Sc., informaticien, CHUS.

Christian Bellemare, M. Sc.

Informaticien et spécialiste en imagerie médicale, conseiller en évaluation des technologies au CHUS. Monsieur Bellemare a agit comme lecteur interne de ce document.

2. Les radiologistes qui ont participé à cette étude :

D^r Jean-Sébastien Billiard, D^{re} Julie Brien, D^{re} Brigitte Grandmont, D^r Yves Marmen, D^r Yves Patenaude, D^r Yves Fortier et D^{re} Mélissa Yipp.

3. Les lecteurs externes :

Guy Bisson, M.D., M. Sc.

Spécialiste en médecine nucléaire, directeur du Service des technologies de l'information et des communications, Faculté de médecine, Université de Sherbrooke.

Gérard Schmutz, M.D.

Radiologiste, chef du Département de radiologie au CHUS

Denis Derome, Ing., M. Sc. A.

Ingénieur biomédical, président de Derome Biomédical Inc., de Terrebonne, Québec.

4. Organismes de financement :

Centre de santé du Granit

Centre hospitalier universitaire de Sherbrooke

Faculté de médecine, Université de Sherbrooke

ABRÉVIATIONS ET DÉFINITIONS

ACR	<i>American College of Radiology</i>
Bit	<i>Diminutif de « Binary Unit ». Il est l'unité binaire d'information utilisé en informatique pour obtenir une réponse à une question qui ne se répond que par un OUI ou par un NON</i>
CAR	<i>Canadian Association of Radiologists</i>
CHUS	<i>Centre hospitalier universitaire de Sherbrooke</i>
CMQ	<i>Collège des Médecins du Québec</i>
DICOM	<i>Digital Imaging and Communication in Medicine</i>
D.O.	<i>Densité optique</i>
IC	<i>Intervalle de confiance</i>
JPEG	<i>Joint Photographic Experts Group</i>
Latitude	<i>Intervalle des niveaux de gris affichés</i>
Lux	<i>Unité de mesure de luminosité qui s'apparente au pied-lambert par la relation : 1 lux = 0,0929 pied-lambert</i>
MEDLINE	<i>Base de données bibliographiques de la National Library of Medicine des États-Unis</i>
PACS	<i>Picture Archiving and Communication System</i>
Pixel	<i>Diminutif de « Picture Element ». Il est l'unité d'information visuelle en imagerie</i>
SMPTE	<i>Society of Motion Picture and Television Engineers</i>
VPP	<i>Valeur prédictive positive</i>
VPN	<i>Valeur prédictive négative</i>

RÉSUMÉ

L'objectif principal de ce projet est d'évaluer l'efficacité pratique d'un système commercial de téléradiologie composé d'un numériseur de films.

L'étude d'efficacité comprend une évaluation technique avec des mires spécifiques et une étude descriptive où 7 radiologistes ont comparé 92 images numérisées de 61 cas cliniques de poumons, d'os et d'abdomen aux films originaux. Le remplacement du numériseur original par un numériseur plus performant et l'implantation d'un algorithme de linéarisation de la chaîne d'imagerie ont pallié aux faiblesses du système original.

Comparée à l'image originale sur film, l'image numérisée et corrigée est de qualité propice à l'établissement d'un diagnostic dans plus de 96% des cas. De ces 96%, 57,6% sont perçus avec une qualité équivalente ou supérieure au film. Les valeurs statistiques de sensibilité, spécificité, VPP, VPN, exactitude et Kappa sont légèrement, mais non significativement, plus élevées pour la lecture sur film que pour la lecture à l'écran. Toutes ces valeurs statistiques dépassent les 93% lorsque la lecture à l'écran est comparée à la lecture des films et ce dans chacun des trois systèmes anatomiques étudiés.

Le niveau de confiance des radiologistes à établir un diagnostic sur les images numérisées dépasse les 95%. Cependant, leur niveau de confiance à établir un aussi bon diagnostic qu'ils le feraient sur le film est nettement inférieur (59,5% - 82,7%), reflétant : 1) un malaise face à leur inexpérience à utiliser l'outil, plutôt que leur inhabilité à faire un diagnostic sur une image numérisée et; 2) un manque de capacité du système à afficher le degré perceptible de latitude que le film peut produire.

Deux actions sont proposées pour améliorer le niveau de confiance des lecteurs. Le premier est de former les utilisateurs et leur donner un temps d'apprentissage suffisant pour leur permettre de se familiariser avec l'outil. Le second est de remplacer le moniteur conventionnel de l'équipement par un moniteur de haute brillance afin d'améliorer suffisamment le degré de latitude perceptible des niveaux de gris affichés à l'écran pour augmenter de façon appréciable la confiance diagnostique des radiologistes.

Ces résultats permettent de conclure qu'un système composé d'un numériseur de films peut être utilisé en téléradiologie si les critères suivants sont respectés :

1. le numériseur de films fournit une résolution qui satisfait les critères de la CAR pour la plus petite dimension de films à numériser;
2. le moniteur fourni avec le système possède les caractéristiques d'un moniteur de haute brillance;
3. chaque composante du système est calibrée selon les spécifications des manufacturiers avant d'utiliser le système;
4. l'ensemble de la chaîne d'imagerie a un comportement linéaire.

TABLE DES MATIÈRES

1.	Introduction.....	1
2.	Besoins en téléradiologie	2
3.	Méthodologie.....	3
3.1	Recension de la littérature en téléradiologie	3
3.2	Système utilisé dans l'étude d'évaluation.....	3
3.3	Ajustement du système	4
3.4	Étude technique d'évaluation des images numérisées	5
3.5	Étude clinique d'évaluation des images numérisées.....	5
3.6	Paramètres statistiques	7
4.	Résultats obtenus.....	8
4.1	Littérature en téléradiologie	8
4.1.1	Critères reconnus d'évaluation de l'image numérique en radiologie	8
4.1.2	Le grand défi de la téléradiologie.....	9
4.1.3	Les bénéfices de la téléradiologie	10
4.2	Évaluation technique	10
4.2.1	Premiers essais de lecture des images numérisées	10
4.2.2	Correction de l'image radiologique	11
4.3	Évaluation clinique : lecture des images numérisées.....	14
4.3.1	Résultats généraux.....	14
4.3.2	Résultats obtenus par système anatomique.....	18
4.3.3	Niveau de confiance des radiologistes face à la lecture d'images numériques sur un écran d'ordinateur	18
5.	Discussion	21
5.1	Méthodologie utilisée.....	21
5.2	Chaîne d'imagerie améliorée.....	21
5.3	Qualité des images numériques	21
5.4	Niveau de confiance des radiologistes face à la lecture d'images numériques	20
5.5	Généralisation des résultats	20
6.	Conclusion.....	21

7. Recommandation	27
Annexe I Types de cas cliniques utilisés	27
Annexe II Formulaires.....	30
Annexe III Performance de chaque radiologiste	33
Annexe IV Niveau de confiance des radiologistes face à la lecture des images numériques pour chacun des systèmes anatomiques	35
Bibliographie_	37

LISTE DES TABLEAUX

Tableau 1 Distribution des publications concernant la téléradiologie de 1966 à 2004.....	3
Tableau 2 Valeurs de densités optiques mesurées sur le film «stepwedge» puis intégrées dans le numériseur de films.	4
Tableau 3 Détail des cas cliniques utilisés dans l'étude clinique.	6
Tableau 4 Dimensions des images exigées par l'ACR pour les modalités utilisées en téléradiologie.	9
Tableau 5 Niveaux de résolution atteints par les lecteurs de plaques au phosphore.	9
Tableau 6 Éléments caractérisant la résolution de l'image numérique et du film visualisés lors des premiers essais de lecture sur l'écran du système de numérisation des films.	11
Tableau 7 Nombre (pourcentage) d'images de qualité diagnostique et non diagnostique obtenu lors de l'étude comparative « Image numérisée vs Image originale analogique ».	14
Tableau 8 Analyse statistique de la détection de l'ensemble des anomalies (essentielle et accessoires) lors de la lecture des images numériques à l'écran et des images originales sur films.	15
Tableau 9 Comparaison directe de la lecture des images numériques à l'écran par rapport au standard établi, soit la lecture des images originales du film sur le négatoscope.	16
Tableau 10 Analyse statistique de la détection de l'ensemble des anomalies par la lecture des images numériques à l'écran et des images originales sur films.	17
Tableau 11 Comparaison directe de la détection de l'ensemble des anomalies par la lecture des images numériques à l'écran par rapport à la lecture des images originales sur film.	18
Tableau 12 Comparaison des images numériques par rapport aux images originales analogiques pour l'ensemble des évaluations effectuées (n = 427).	20
Tableau 13 Comparaison de la qualité des images numériques par rapport aux images originales analogiques pour les trois systèmes étudiés : abdominal, pulmonaire et osseux.	20

Tableau 14 Analyse statistique de la détection de l'ensemble des anomalies (essentielles et accessoire) de chaque radiologiste participant à l'étude, lors de la lecture des images numériques sur écran et des images originales sur films.....	33
Tableau 15 Comparaison directe de la lecture des images numériques à l'écran par rapport à la lecture des images originales du film sur le négatoscope, pour chacun des radiologistes participant à l'étude. La lecture de l'image originale est considérée comme la méthode standard.....	34
Tableau 16 Comparaison des images numériques par rapport aux images originales analogiques pour le système pulmonaire (n = 168).	35
Tableau 17 Comparaison des images numériques par rapport aux images originales analogiques pour le système osseux (n = 196).....	27
Tableau 18 Comparaison des images numériques par rapport aux images originales analogiques pour le système abdominal (n = 63).....	36

LISTE DES FIGURES

- Figure 1 : Linéarisation de la relation « valeurs des pixels » vs valeur de D.O. mesurées sur le film « stepwedge » avec le densitomètre X-rite, modèle 301. 4
- Figure 2 : Fonction de transfert du moniteur ($f_{\text{écran}}$) (ligne pointillée) corrigée pour rendre linéaire le système d'imagerie. La nouvelle fonction de transfert du moniteur ($f_{\text{écranC}}$) (ligne pleine mince) est comparée à la fonction de transfert du film (f_{film}) (ligne pleine grasse). 12
- Figure 3 : Valeurs de contraste calculées à partir des valeurs de densité optique mesurées sur les images « stepwedge » numériques, non corrigées (ligne pointillée) et corrigées (ligne pleine mince), et analogique (ligne pleine grasse). 13

CHAPITRE 1

1. INTRODUCTION

Ce rapport a été préparé par l'Unité d'évaluation des technologies et de modes d'intervention en santé (ETMIS) du CHUS suite aux besoins exprimés par un centre hospitalier de la région de l'Estrie concernant l'établissement d'un service de téléradiologie avec le CHUS.

Une étude d'efficacité pratique d'un équipement de téléradiologie composé d'un numériseur de films et d'un système de visualisation des images numériques, constitue l'essentiel de ce rapport. Les impacts éthiques et légaux de l'utilisation clinique de ce type de technologie en téléradiologie ne font pas partie de ce rapport. Ils sont traités dans le cadre de la rédaction d'un mémoire en évaluation des technologies de la santé [Lemieux, 2004]. La question de recherche à laquelle nous avons répondu est la suivante :

« Un système commercial composé d'un numériseur de films radiologiques peut-il permettre la visualisation sur un moniteur d'ordinateur d'un nombre suffisant d'anomalies permettant de procéder à l'établissement d'un diagnostic final avec la même capacité de détection et le même niveau de confiance que la lecture des films radiologiques sur négatoscope? »

La capacité des radiologistes d'effectuer un bon diagnostic sur les images visualisées ne fait pas l'objet de ce rapport. Ce dernier se concentre plutôt sur la capacité des radiologistes d'identifier les anomalies, un processus qui initie le processus de réflexion et qui mène à la décision diagnostique. La capacité d'établir un diagnostic est basée non seulement sur des critères objectifs que constitue le nombre d'anomalies trouvées sur les images, mais aussi sur des critères subjectifs qui dépendent des capacités physiques et mentales du radiologiste au moment de la visualisation des images et de son expérience clinique. L'interprétation de l'ensemble des anomalies trouvées sur les images en association avec l'histoire médicale du patient constituent le processus de prise de décision diagnostique. Nous concentrons notre étude sur le caractère objectif de la prise de décision diagnostique, soit l'identification des anomalies.

Ce projet d'évaluation comporte deux composantes. La première, d'ordre technique, consiste à mesurer la capacité du système à afficher l'image numérisée sans déformation perceptible par l'utilisateur. Cette étude est effectuée en laboratoire à l'aide de mires spécifiques. La seconde composante, d'ordre clinique, consiste à valider en milieu clinique la qualité diagnostique de l'image affichée sur un écran d'ordinateur par rapport au standard actuel, le film. Sept radiologistes ont participé à une étude comparative effectuée sur quatre-vingt-douze images provenant de soixante et un cas cliniques réels tirés de la salle d'urgence du CHUS.

Cette étude nous amène à recommander l'utilisation de ce type de technologie en téléradiologie à condition qu'elle soit bien calibrée et que la chaîne d'imagerie soit linéaire. L'implantation adéquate de cette technologie doit tenir en compte les considérations des professionnels de la santé directement impliqués et respecter l'environnement auquel la technologie est destinée.

CHAPITRE 2

2. BESOINS EN TÉLÉRADIOLOGIE

La téléradiologie et la visioconférence sont actuellement les activités les plus couramment pratiquées en télésanté au Québec. La téléradiologie se fait principalement dans les centres hospitaliers où un PACS (*Picture Archiving and Communication System*) est installé et fonctionnel. Par contre, les besoins en téléradiologie se retrouvent dans les plus petits centres hospitaliers situés en région éloignée. Ces centres régionaux sont dépourvus de radiologistes sur place. De plus, la présence sporadique d'un radiologiste ne suffit pas à combler la demande en examen de radiologie pour les patients de la région. Ils n'ont aucun support radiologique en dehors des heures régulières de travail, ce qui met en péril le continuum de soins dans les situations cliniques d'urgence ou de complications diagnostiques ou thérapeutiques. Ceci augmente le nombre des transferts de patients vers les grands centres hospitaliers.

L'avènement de la téléradiologie comblerait la carence locale d'expertise en lecture radiologique dont ces centres souffrent. Elle augmenterait la qualité des soins donnés aux patients [Wadley et coll., 2002], d'abord en augmentant la prise en charge complète des patients au centre régional, ce qui diminuerait les transferts vers les grands centres [Stranzinger et coll., 2003], et ensuite, en accélérant la prise en

charge thérapeutique du patient par la confirmation rapide d'un diagnostic [Ricci & Borelli, 2002].

Malheureusement, ces centres hospitaliers régionaux n'ont pas la capacité financière d'acquérir un PACS dont le prix varie entre 0,5 et 13 millions de dollars, selon la grosseur du centre hospitalier où le système est installé. Par contre, tout en conservant la gestion papier des examens radiologiques, trois options de numérisation des images radiologiques s'offrent à eux :

1. l'acquisition d'un appareil de radiographie numérique qui coûte au-delà de 500 000 \$;
2. l'utilisation d'un lecteur de plaques au phosphore dont le prix est de l'ordre de 150 000 \$ incluant un système d'archivage et une imprimante;
3. l'utilisation d'un numériseur de films, dont le coût d'acquisition s'élève aux environs des 60 000 \$.

Cette dernière option semble la plus intéressante pour les petits centres hospitaliers qui font peu d'examens radiologiques et qui n'ont pas la capacité financière d'acquérir une des deux autres technologies déjà éprouvées et disponibles commercialement

CHAPITRE 3

3. MÉTHODOLOGIE

3.1 Recension de la littérature en téléradiologie

Une recherche sur l'engin MEDLINE a permis de répertorier sept cent treize articles sur la téléradiologie selon la répartition présentée au tableau 1.

Tableau 1
Distribution des publications concernant la téléradiologie de 1966 à 2004.

#	Historique de la recherche sur la téléradiologie	Résultats
1	Nombre total de publications	713
2	Nombre de publications de 1966 à 1993	5
3	Nombre de publications de 1994 à 1998	386
4	Nombre de publications de 1999 à 2004	322

Les mots-clés utilisés dans cette recherche sont : classification, organisation et administration, économie, standards, historique, statistiques et données numériques, instrumentation, tendances technologiques, législation et jurisprudence, utilisation et méthodes. Ils ont permis d'effectuer un lien entre numérisation d'images radiologiques et téléradiologie.

3.2 Système utilisé dans l'étude d'évaluation

Le système testé est le modèle SDX-125 de la compagnie Cifra Médical Inc. de Québec. Il est composé d'un numériseur de films Lumiscan 20 de la compagnie Lumisys, d'un ordinateur de type Pentium III et d'un logiciel de visualisation et de gestion des images appelé « Cifravisión ».

Une étude technique du numériseur Lumiscan 20 a démontré que ce dernier ne remplissait pas les critères de l'Association canadienne de radiologie. Il a donc été remplacé par un appareil plus performant, soit le Lumiscan 75 dont la résolution spatiale atteint 102,4 pixels par centimètre, soit 3 584 pixels sur une largeur de 35 centimètres. Sa profondeur est 8 ou 12 bits au choix de l'utilisateur. L'écran de l'ordinateur possède une matrice d'affichage qui peut atteindre 1600 x 1200 pixels, avec une résolution en profondeur de 32 bits « *true color* ». Le logiciel de visualisation et de gestion permet de modifier les images visualisées à l'écran à l'aide des fonctions suivantes : agrandissement, contraste, luminosité, région d'intérêt, règle pour mesurer la dimension des structures anatomiques, inversion des teintes de gris et annotations.

3.3 Ajustement du système

Le numériseur de films a d'abord été calibré selon la méthode fournie par la compagnie Lumisys [Lim, A.J., 1996]. Un film «stepwedge» de 32 pas, dont la densité optique (D.O.) varie entre 0,07 et 3,76, a été utilisé. Une correspondance linéaire entre les

valeurs de D.O. lues durant la numérisation du film et celles enregistrées dans le numériseur de films a été établie par minimisation des moindres carrés (Figure 1). Les valeurs de D.O. intégrées dans le numériseur de films avaient préalablement été mesurées avec le densitomètre modèle 301 de la compagnie X-rite (Tableau 2).

Tableau 2
Valeurs de densité optique mesurées sur le film «stepwedge» puis intégrées dans le numériseur de films.

Pas	D.O. mesurées	Pas	D.O. mesurées	Pas	D.O. mesurée	Pas	D.O. mesurée
1	0,07 ± 0,01	9	0,95 ± 0,01	17	1,90 ± 0,01	25	2,92 ± 0,01
2	0,17 ± 0,01	10	1,08 ± 0,01	18	2,03 ± 0,01	26	3,03 ± 0,01
3	0,27 ± 0,01	11	1,19 ± 0,01	19	2,14 ± 0,01	27	3,17 ± 0,01
4	0,38 ± 0,01	12	1,30 ± 0,01	20	2,26 ± 0,01	28	3,29 ± 0,01
5	0,50 ± 0,01	13	1,41 ± 0,01	21	2,38 ± 0,01	29	3,41 ± 0,01
6	0,62 ± 0,01	14	1,54 ± 0,01	22	2,52 ± 0,01	30	3,56 ± 0,01
7	0,75 ± 0,01	15	1,65 ± 0,01	23	2,66 ± 0,01	31	3,71 ± 0,01
8	0,86 ± 0,01	16	1,77 ± 0,01	24	2,78 ± 0,01	32	3,87 ± 0,01

Calibrage du numériseur Lumiscan 75

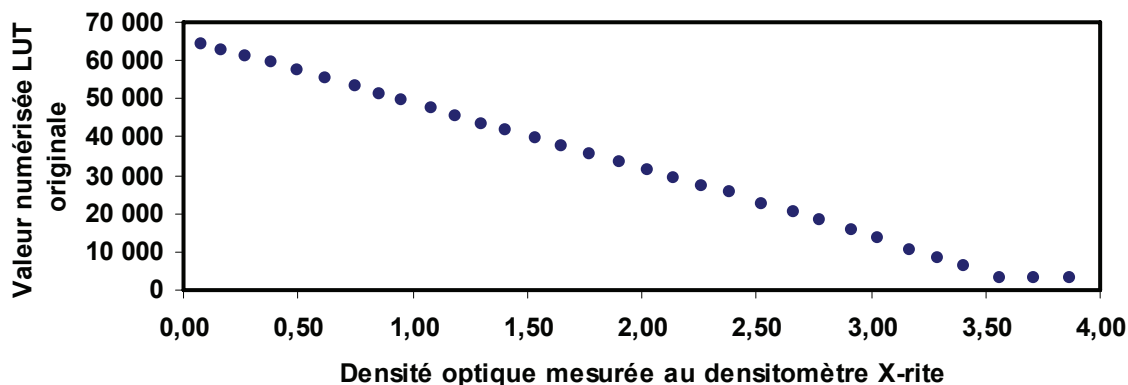


Figure 1 : Linéarisation de la relation «valeurs des pixels» vs valeur de D.O. mesurées sur le film «stepwedge» avec le densitomètre X-rite, modèle 301.

Le processus de calibrage élimine la détection des densités optiques supérieures à 3,56 (30^{ième} ligne du film « stepwedge »). Ceci correspond à la portion horizontale de la courbe de calibrage montrée à la figure 1. Cette limite se retrouve à l'extérieur de la bande de niveaux de gris utiles en radiologie. Elle est donc acceptable pour la numérisation de films radiologiques [Bushberg et coll, 1994].

3.4 Étude technique d'évaluation des images numérisées

Cinq mires et un film « stepwedge » ont été utilisés pour comparer la résolution de l'image numérique affichée à l'écran d'ordinateur par rapport à l'image originale sur film.

D'abord, la résolution à bas contraste a été évaluée en utilisant le film « stepwedge », la mire Philips avec une plaque de cuivre de 0,8 mm d'épaisseur et la mire SMPTE « Standard Society of Motion Picture and Television Engineers ». La mire Philips permet de détecter 6 niveaux de gris différents à partir d'objets dont la dimension minimale est de 0,7 millimètre. Ces niveaux de gris peuvent être distingués du niveau de gris produit par le matériau composant la mire même. Similairement, la mire SMPTE permet la détection de 10 niveaux de gris qui peuvent être distingués à un onzième niveau qui constitue la base même de la mire. Le film « stepwedge » a permis d'identifier la densité maximale pouvant être visualisée à partir des 32 niveaux de gris offerts.

Ensuite, la résolution à haut contraste a été évaluée avec la mire SMPTE et la mire Nuclear Associates No. 07-527 dans laquelle le nombre de paires de lignes par millimètre varie entre 0,6 et 5,0. La résolution à haut contraste de l'image numérique a été compa-

rée à la résolution de l'image originale analogique sur le film.

Finalement, l'uniformité de l'image numérique a été évaluée avec une mire constituée d'une plaque de cuivre de 0,8 mm d'épaisseur parfaitement uniforme. Cette mire permettait de visualiser les imperfections pouvant apparaître lors du développement du film et lors de la numérisation de ce dernier.

3.5 Étude clinique d'évaluation des images numérisées

Sept radiologistes généraux, dont cinq sont certifiés par le Collège Royal du Canada et deux sont des résidents seniors, ont été sélectionnés pour les fins de cette étude. Deux radiologistes avaient plus de 25 ans d'expérience et n'avaient jamais utilisé l'image radiographique numérique dans le cadre de leur pratique. Les deux résidents avaient déjà utilisé un PACS dans le cadre de leur formation. Les trois autres radiologistes, d'expérience variable, n'avaient jamais lu d'images radiographiques numériques dans le cadre de leur pratique. Tous pratiquent la radiologie générale en salle d'urgence et lisent des clichés de tout genre.

Quatre vingt douze (92) images provenant de soixante et un cas cliniques ont été évaluées par chacun des sept radiologistes participant à cette étude (Tableau 3). Les cas choisis ont été sélectionnés judicieusement de manière à varier la complexité de lecture pour qu'elle soit représentative, selon le radiologiste expert du groupe de recherche, d'une journée habituelle à l'urgence de n'importe quel centre hospitalier régional. Les systèmes anatomiques choisis sont ceux qui sont les plus fréquemment radiographiés à l'urgence. L'éventail des cas choisis com-

portait en moyenne 15% de cas normaux. Un cas est dit normal lorsqu'il ne présente au-

cune pathologie.

Tableau 3
Détail des cas cliniques utilisés dans l'étude clinique.

	# total de cas	# cas normaux (%)	# total d'images
Poumons	24	4 (16%)	36
Ostéo-articulaires	28	4 (14%)	41
Abdomen	9	1 (11%)	15
TOTAL	61	9 (15%)	92

Toutes les images ont été numérisées à une profondeur de 12 bits et avec la résolution spatiale du numériseur, soit 5,1 paires de lignes par millimètre.

Les images numériques affichées sur écran ont été comparées aux images originales analogiques visualisées sur un négatoscope. Aucune information clinique n'était disponible pour faciliter la tâche des radiologistes. Ces derniers ne pouvaient se fier que sur leur capacité d'observer les images. Les valeurs de sensibilité, de spécificité et les valeurs prédictives ont été établies pour chaque mode de lecture, séparément et en comparaison de l'un avec l'autre.

Dans le but de vérifier si la qualité de l'image numérique était suffisante pour effectuer un diagnostic, des éléments précis de l'image devaient être reconnus par les radiologistes. Nous avons qualifié ces éléments d'anomalies et ces dernières ont été divisées en deux catégories : les anomalies essentielles et les anomalies accessoires. Une anomalie est jugée essentielle lorsqu'elle permet l'établissement d'un diagnostic ou lorsqu'elle supporte ou modifie un diagnostic préétabli par l'examen d'autres données cliniques. Par contre, une anomalie est jugée accessoire lorsqu'elle

complète la lecture sans qu'elle soit liée nécessairement au diagnostic. Les fiches utilisées lors des sessions de lecture sont disponibles à l'annexe 2.

Puisque notre hypothèse de départ était que la capacité de détection des anomalies lors de la lecture des images numériques serait moindre que la capacité de détection lors de la lecture des films, l'image numérique a été analysée en premier. Notre session de visualisation des images comportait trois étapes :

1. l'image numérique a d'abord été lue et si des anomalies étaient détectées, elles étaient notées sur une fiche par un assistant non initié à la lecture d'images radiologiques;
2. l'image originale sur film était ensuite affichée sur un négatoscope puis lue par le même radiologiste, afin de servir de comparaison. Les anomalies détectées étaient alors notées sur une fiche par le même assistant;
3. finalement, les deux images, numériques et analogiques, étaient affichées en même temps afin que le radiologiste qualifie l'image

numérique par rapport à l'image originale analogique.

Chaque image a été préalablement analysée attentivement par le radiologiste membre de l'équipe de recherche pour poursuivre deux objectifs. D'abord, identifier le plus d'anomalies possibles et ensuite, identifier les fausses anomalies dues à des artéfacts dans l'image ou dues à une mauvaise interprétation des lecteurs. Puisqu'il s'agissait de cas cliniques réels, le diagnostic était connu par le radiologiste-chercheur, mais non disponible aux radiologistes participant à cette étude.

Une révision des cas cliniques a été effectuée par le radiologiste membre de l'équipe de recherche peu de temps après l'étude comparative. Toutes les anomalies préalablement manquées lors de l'évaluation des images par le radiologiste-chercheur, durant la phase d'élaboration de la liste des anomalies à détecter, mais qui ont été identifiées durant l'étude par au moins un des sept radiologistes de l'équipe, ont été validées en réévaluant les images concernées numériques et analogiques. S'il s'avérait que les nouvelles anomalies identifiées sur les images numériques étaient aussi présentes sur les films, ou vice-versa, elles étaient ajoutées à notre liste d'anomalies. Dans le cas contraire, elles entraient dans la catégorie des « fausses anomalies ». Ceci a permis de compléter la liste d'anomalies prévues à être identifiées durant l'étude clinique.

3.6 Paramètres statistiques

Les paramètres statistiques quantitatifs utilisés dans cette étude sont :

- la sensibilité et la spécificité;
- les valeurs prédictives positives et négatives;
- l'exactitude d'un test diagnostique;
- le coefficient de concordance Kappa.

Les paramètres statistiques quantitatifs les plus importants sont l'exactitude du test diagnostique et le coefficient de concordance Kappa. Le premier indique le niveau d'insensibilité du test aux erreurs systématiques et biais de toutes sortes qui pourraient se retrouver soit dans l'équipe de radiologistes participant à l'étude ou dans la manière de performer le test. Le second est une mesure quantitative de la reproductibilité d'une mesure chez un groupe d'individus. Il quantifie la concordance entre deux mesures effectuées à des temps différents, chez un même groupe d'individus, dans les mêmes conditions expérimentales (Bouyer et coll., 1995).

L'appréciation subjective des radiologistes concernant 1) leur appréciation de la qualité de l'image numérique et 2) leur niveau de confiance d'effectuer le bon diagnostic sur l'image numérique, est exprimé en terme de pourcentage.

CHAPITRE 4

4. RÉSULTATS OBTENUS

4.1 Littérature en téléradiologie

4.1.1 Critères reconnus d'évaluation de l'image numérique en radiologie

Plusieurs critères objectifs de qualité de l'image radiologique imprimée sur film sont émis par la «Canadian Association of Radiologists» (CAR) et «American College of Radiology» (ACR) [CAR, 1999; ACR Guidelines, 2002]. Les critères principaux sont la résolution, le contraste et la précision. Ils sont déterminés par les médecins à partir de tests précis effectués sur des images montrant des mires spécifiques [Siebert et coll., 1994].

Les critères de la CAR sont publiés de façon succincte. Ils se résument comme suit : la numérisation d'une image radiologique doit se faire avec une résolution spatiale minimale de 2,5 paires de lignes par mm et une résolution en contraste d'au moins 12 bits. Par contre, la visualisation d'une image radiologique doit se faire sur un écran dont la profondeur minimale est de 8 bits, mais de

résolution spatiale de 2000 x 2000 pixels. La luminosité de l'écran devrait dépasser les 4000 lux (360 pieds-lambert).

Par contre, les critères de l'ACR sont plus explicites. Ils sont indiqués comme suit :

1. l'image provenant d'une modalité dont la matrice d'images est de faible dimension (tomodensitométrie axiale, IRM, échographie, médecine nucléaire, fluorographie, angiographie) doit être acquise à une dimension minimale de 512 x 512 x 8 bits et doit être affichée sur un moniteur d'au moins 50 pieds-lambert de luminosité;
2. l'image provenant d'une modalité dont la matrice d'images est de dimension élevée (radiographie numérique, numériseur de films radiologiques) doit être acquise avec une résolution spatiale minimale de 2,5 paires de lignes par mm et une résolution en contraste de 10 bits, et doit être affichée sur un moniteur d'au moins 50 pieds-lambert de luminosité.

Tableau 4
Dimensions des images exigées par l'ACR pour les modalités utilisées en
téléradiologie.

Modalité	Dimension de l'image produite (pixel ²)
Mammographie	non déterminée à ce jour
Radiographie	résolution spatiale de 2,5 paires de lignes par mm
Tomodensitométrie, IRM, échographie	512 x 512

Les dimensions minimales des images exprimées par l'ACR pour chaque modalité, dans un contexte d'utilisation en téléradiologie, sont résumées au tableau 4.

4.1.2 Le grand défi de la téléradiologie

La difficulté de la téléradiologie réside dans la réception d'images de qualité diagnostique, transmise à distance par un lien réseautique, que ce dernier soit sans fil ou constitué de fibres optiques ou autres, et ce, de façon sécurisée.

Aujourd'hui, toutes les modalités, peu importe le fournisseur, produisent des images de dimensions égales ou supérieures à celles exigées par l'ACR. À titre d'exemple, l'image de radiographie numérique produite possède une résolution spatiale d'au moins 5 pixels par mm, ce qui équivaut à 2,5 paires de lignes par mm. Le tableau 5 montre les niveaux de résolution que peut atteindre un lecteur de plaques au phosphore couramment utilisé dans les systèmes PACS [Kato, 1987].

Tableau 5
Niveaux de résolution atteints par les lecteurs de plaques au phosphore.

Dimension de la plaque (cm)	Résolution spatiale (pixels/mm)	Résolution en densité (bits/pixel)
43 x 35	5	8
35 x 35	5	8
30 x 25	6,7	8
20 x 25	10	8

Considérant les caractéristiques des images radiologiques, on peut s'attendre à ce que les images des modalités de basse

résolution (tomodensitométrie, résonance magnétique, échographie) soient plus facilement utilisables en téléradiologie. Leur

plus faible résolution permet leur transport sur un réseau de faible puissance [Glowacki et coll., 2002] en format JPEG (*Joint Photographic Expert Group*) sans détérioration perceptible de leur qualité à la lecture [Kondo, 2002]. Ce format est communément utilisé aujourd'hui dans diverses applications d'imageries, incluant les caméras numériques disponibles sur le marché.

4.1.3 Les bénéfices de la téléradiologie

Deux bénéfices ont été rapportés dans la littérature. Le premier est l'amélioration des soins aux patients en région et le second est d'ordre économique. Les bienfaits de la téléradiologie se retrouvent dans les régions éloignées où l'absence de radiologiste rend difficile le maintien d'un service de radiologie et dans les petits hôpitaux qui n'ont pas de radiologiste à demeure et qui dépendent de la visite sporadique de celui-ci pour offrir un service de radiologie à la population. Dans ces régions, la téléradiologie améliore la qualité des soins aux patients [Crowe et coll., 1996; Bodemeyer et coll., 2002; Wadley et coll., 2002; Johansen, 2000], d'abord, en accélérant le diagnostic final en vue de débiter le traitement plus rapidement [Kiuru et coll., 2002; Ricci & Borelli, 2002] et ensuite, en diminuant le nombre de transferts inutiles vers les grands centres spécialisés [Maass, 2000; Vuletic, 2001; Stranzinger et coll., 2003; Jithoo et coll., 2003]. Ceci favorise la prise en charge complète des patients dans leur milieu de vie.

Les économies réalisées par l'utilisation de la téléradiologie peuvent être substantielles, comme dans les cas d'urgence [Darkins et coll., 1996]. À titre d'exemple, Brumage et collègues [2001] rapportent que l'utilisation de la téléradiologie dans les camps d'entraînement de l'armée américaine

a généré des économies estimées à plus de 176 500 \$ annuellement. Sur une période de 51 jours, 29 transferts de soldats vers un hôpital civil situé à 55 kilomètres plus loin ont pu être évités. Cette période seulement a généré une économie de 36 600 \$. Les radiographies étaient prises directement au camp d'entraînement et les images étaient interprétées à distance par des radiologistes localisés à l'hôpital civil.

4.2 Évaluation technique

4.2.1 Premiers essais de lecture des images numérisées

Même si toutes les caractéristiques de résolution à bas et à haut contraste, ainsi que l'uniformité de l'image, ont été visualisées tant à l'écran que sur le film (Tableau 6), tous les premiers essais de numérisation des films se sont avérés infructueux. Trois éléments sont ressortis lors de la lecture des images par le radiologiste membre de l'équipe de recherche :

1. la résolution des images numérisées était insuffisante pour porter un diagnostic, particulièrement pour les images ostéo-articulaires;
2. les niveaux de contraste des images de poumons et d'abdomen étaient insuffisants et ne pouvaient être corrigés adéquatement par les possibilités qu'offrait le logiciel offert avec le système;
3. les niveaux de gris affichés à l'écran étaient trop pâles pour l'ensemble des images et ne pouvaient être corrigés ni par les fonctions offertes par le logiciel de visualisation, ni par les fonctions offertes par le moniteur utilisé. Le logiciel ne comportait pas

de fonctions avancées de traitement d'images.

linéaire de la chaîne d'imagerie, particulièrement du moniteur puisque le numériseur a révélé un comportement linéaire lors de son calibrage. Un filtre qui inverse la fonction de transfert non linéaire du moniteur a été développé puis appliqué à l'image juste avant sa visualisation à l'écran.

4.2.2 Correction de l'image radiologique

L'évaluation technique de l'équipement a permis de mettre en évidence l'aspect non

Tableau 6
Éléments caractérisant la résolution de l'image numérique et du film visualisés lors des premiers essais de lecture sur l'écran du système de numérisation des films.

Éléments à visualiser	Observations sur l'image numérique	Observations sur le film
Mire Philips		
6 cercles de 9,0 mm de diamètre, de densité croissante du centre vers la périphérie.	6 cercles identifiés clairement	6 cercles identifiés clairement
1 cercle de 55,0 mm de diamètre	1 cercle identifié clairement	1 cercle identifié clairement
1 cercle de 40,0 mm de diamètre	1 cercle identifié clairement	1 cercle identifié clairement
2 cercles de 20,0 mm de diamètre.	2 cercles identifiés clairement	2 cercles identifiés clairement
1 anneau de 7,5 mm de largeur	1 anneau identifié clairement	1 anneau identifié clairement
12 cercles de 1,0 mm de diamètre	12 cercles identifiés clairement	12 cercles identifiés clairement
16 cercles de 0,7 mm de diamètre	16 cercles identifiés clairement	16 cercles identifiés clairement
Mire NA #07-527		
# de paires de lignes par mm identifiées clairement	2,5	2,5
Film « stepwedge »		
Densité optique identifiée	3,56	3,87
Mire SMPTE		
# Niveaux de gris identifiés	11 sur 11	11 sur 11
Contraste 0%/5%	Identifié clairement	Identifié clairement
Contraste 95%/100%	Peu visible	Identifié clairement
Patron haute résolution 3:3	Identifié clairement	Identifié clairement
Patron haute résolution 2:2	Identifié clairement	Identifié clairement
Patron haute résolution 1:1	Non identifié	Identifié clairement
Mire Plaque simple de cuivre 0,8 mm d'épaisseur		
Vérification de l'uniformité de l'image	Image uniforme	Image uniforme

Basée sur la mesure de l'intensité lumineuse traversant le film «stepwedge» (transmittance) placé sur un négatoscope, de luminance de 5 918 lux, et de l'intensité lumineuse affichée sur l'écran, la méthode utilisée a permis d'établir la fonction de transfert du moniteur ($f_{\text{écran}}$). La modélisation de cette fonction a été effectuée par une minimisation au sens des moindres carrés entre un polynôme d'ordre 3 et les valeurs d'intensités mesurées. Le polynôme suivant a été obtenu.

$$f_{\text{écran}} = 12,5285 - 0,18586x + 0,00937018x^2 - 0,00046452x^3$$

($rms = 0,164$)

Équation 1

L'intensité lumineuse de chaque pas du film a été mesurée avec un luxmètre modèle Lunasix 3 de la compagnie Gossen (Nürnberg, Allemagne).

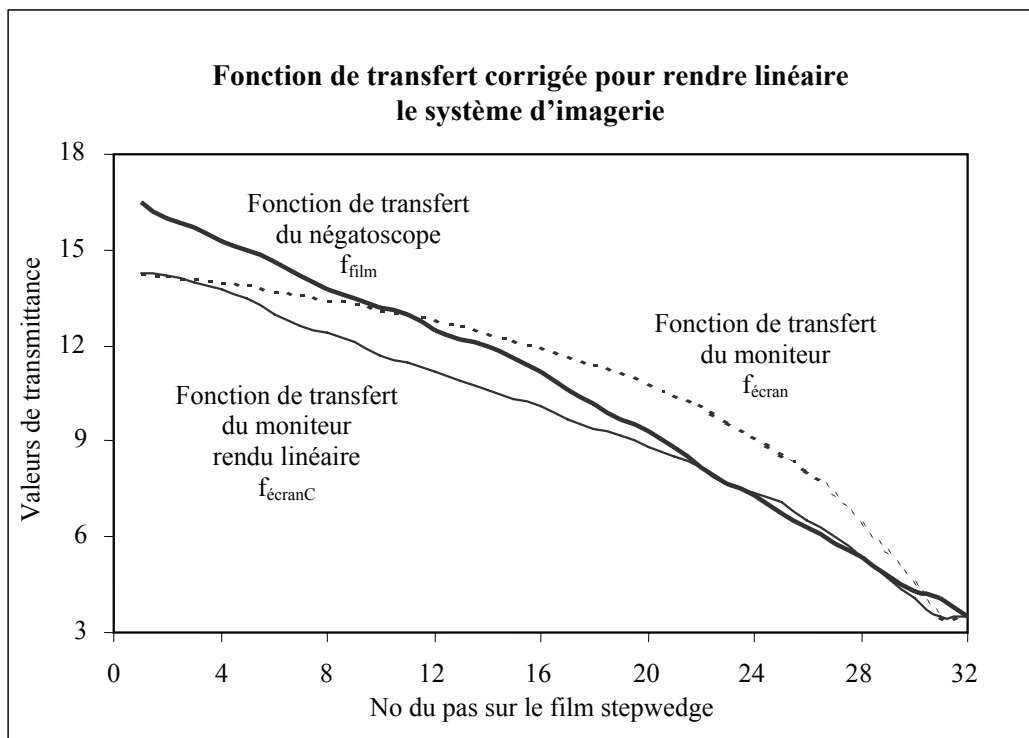


Figure 2 : Fonction de transfert du moniteur ($f_{\text{écran}}$) (ligne pointillée) corrigée pour rendre linéaire le système d'imagerie. La nouvelle fonction de transfert du moniteur ($f_{\text{écranC}}$) (ligne pleine mince) est comparée à la fonction de transfert du film (f_{film}) (ligne pleine grasse).

La fonction de transfert du moniteur ($f_{\text{écran}}$) a été inversée pour obtenir un polynôme ($f_{\text{polynôme}}$) dont les paramètres ont été déterminés de la façon suivante. Le produit des fonctions $f_{\text{écran}}$ et $f_{\text{polynôme}}$, a subi un processus d'optimisation par une minimisation de l'erreur quadratique lorsque le produit de ces fonctions a été comparé à la droite caractérisant la correspondance

parfaitement linéaire entre l'image numérisée et l'image affichée à l'écran. Les limites physiques du moniteur utilisé ont été prises en considération durant ce processus. La figure 2 montre l'allure des fonctions de transfert originales de chacun des médiums, ainsi que la nouvelle fonction de transfert du moniteur ($f_{\text{écranC}}$) une fois le processus d'optimisation terminée. La fonction $f_{\text{polynôme}}$

a été appliquée aux images numérisées juste avant leur visualisation à l'écran.

Les valeurs de contraste entre chaque pas de l'image « stepwedge » ont été calculées pour les trois images suivantes : l'image numérique non corrigée, l'image numérique corrigée et l'image analogique sur négatoscope. La figure 3 montre la proche correspondance entre la courbe de l'image numérique corrigée (ligne pleine mince) et

celle de l'image originale analogique (ligne pleine grosse), comparativement à la courbe de l'image numérique non corrigée (ligne pointillée). Une diminution de contraste dans une image radiographique limite la distinction de tissus voisins dont les densités sont proches l'une de l'autre, surtout si l'un de ces tissus est pathologique. C'est le cas de la courbe de l'image numérique non corrigée (ligne pointillée) pour les 15 premiers pas

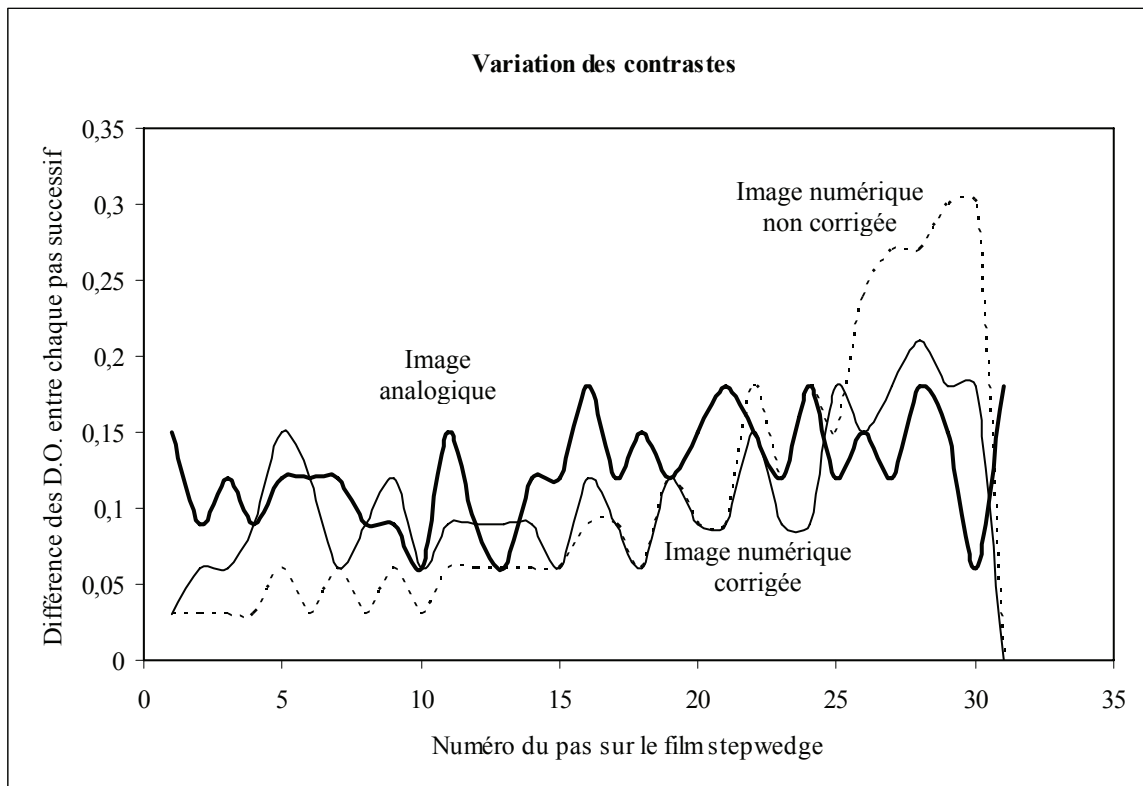


Figure 3 : Valeurs de contraste calculées à partir des valeurs de densité optique mesurées sur les images « stepwedge » numériques, non corrigées (ligne pointillée) et corrigées (ligne pleine mince), et analogique (ligne pleine grosse).

L'application de notre algorithme ($f_{\text{polynôme}}$) sur l'image numérisée de la mire SMPTE a augmenté la densité de tous les niveaux de gris de 15%, à la satisfaction du radiologue membre de l'équipe de recherche. Par contre, elle a rendu similaire les deux niveaux de gris les plus foncés de manière à faire disparaître le contraste 0/5%. Les valeurs de densité optique de ces deux niveaux de gris sont à l'extérieur de l'intervalle des densités optiques requises

pour la lecture des images radiographiques. La disparition du contraste 0/5% n'a pas de conséquence négative sur la qualité des images affichées sur le moniteur.

Cette correction de l'image nous a permis de passer à l'étape suivante qui consistait à procéder à la comparaison du processus de lecture des images numériques à l'écran au processus de lecture des images originales sur films.

4.3 Évaluation clinique : lecture des images numérisées

les radiologistes de l'équipe comme étant de qualité propice à l'établissement d'un diagnostic dans 96% des cas. De ces 96%, 57,6% des images numériques sont perçues comme étant de qualité équivalente ou supérieure à l'image analogique sur film.

4.3.1 Résultats généraux

Le tableau 7 montre que l'image numérisée et corrigée a été déterminée par

Tableau 7
Nombre (pourcentage) d'images de qualité diagnostique et non diagnostique obtenu lors de l'étude comparative « Image numérisée vs Image originale analogique ».

	Nombres (pourcentages)			% moyen ± D.S
	Images de poumons	Images d'os	Images d'abdomen	
Qualité diagnostique	34 (94%)	40 (98%)	14 (96%)	96,0% ± 1,6%
Numérique < analogique	20 (54%)	11 (28%)	5 (33%)	38,3% ± 11,3%
numérique = analogique	14 (40%)	28 (69%)	9 (63%)	57,3% ± 12,5%
numérique > analogique	0 (0%)	1 (1%)	0 (0%)	0,3% ± 0,5%
Qualité non diagnostique	2 (6%)	1 (2%)	1 (4%)	4,0% ± 1,6%
Nombre d'images	36 (100%)	41 (100%)	15 (100%)	
Nombre de cas cliniques	24	28	9	

Les valeurs statistiques observées au tableau 8 se ressemblent pour les deux modes de lecture. Bien que les valeurs reliées à la lecture sur films soient légèrement supérieures à celles reliées à la lecture sur l'écran, sauf pour la spécificité et la valeur prédictive positive, les différences observées ne sont pas significatives. Les faibles valeurs Kappa (0,24 pour l'écran et 0,26 pour le film) montrent une forte influence du hasard dans le processus de détection des anomalies, tant à l'écran que sur le film. Les résultats montrent que la détection d'anomalies sur des images est un processus peu sensible, peu exacte, mais très spécifique.

La comparaison entre les deux processus de lecture montre une forte correspondance entre capacité de détecter des anomalies lors de la lecture à l'écran et la capacité de détecter des anomalies lors de la lecture sur films. Une telle correspondance est révélée par des valeurs statistiques très élevées : Kappa (κ)= 0,94, sensibilité = 94,4% et spécificité = 99,9%. La lecture à distance d'images numériques pourrait donc s'avérer aussi efficace que la lecture de films pour l'ensemble des cas de salle d'urgence, soit les poumons, l'ostéo-articulaire et l'abdomen. La valeur élevée à 99,9% pour la spécificité est due au fait qu'un seul radiologiste a identifié une seule fausse anomalie sur l'écran d'ordinateur, mais pas

sur le film (spécificité = 99,2%) (voir annexe 3). Il s'agit d'une anomalie essentielle d'un cas de poumons. Trois autres fausses anomalies ont été identifiées sur le film,

mais pas sur l'écran. Il s'agit d'un cas d'abdomen et de deux cas d'ostéo-articulaire.

Tableau 8
Analyse statistique de la détection de l'ensemble des anomalies (essentielles et accessoires) lors de la lecture des images numériques à l'écran et des images originales sur films.

	Ensemble des anomalies		Total
	Présentes sur les images	Absentes des images, mais créées suite à une erreur d'interprétation par au moins 1 des 7 radiologistes	
Détectées à l'écran	861	30	891
Non détectées à l'écran	679	257	936
Détectées sur le film	911	33	944
Non détectées sur le film	629	254	883
Total	1540	287	1827

	Écran	Film
Sensibilité (%)	55,9%	59,2%
Spécificité (%)	89,5%	88,5%
Valeur prédictive positive, VPP (%)	96,6%	96,5%
Valeur prédictive négative, VPN (%)	27,5%	28,8%
Exactitude (%)	61,1%	63,8%
Kappa	0,24	0,26

Tableau 9
Comparaison directe de la lecture des images numériques à l'écran par rapport au standard établi, soit la lecture des images originales du film sur le négatoscope.

L'ensemble des anomalies

Anomalies à l'écran	Anomalies sur film		Total
	détectées	non détectées	
détectées	890	1	891
non détectées	54	882	936
Total	944	883	1827
Sensibilité (%)	94,3%	VPP (%)	99,9%
Spécificité (%)	99,9%	VPN (%)	94,2%
Exactitude (%)	97,0%	Kappa	0,94

Le même phénomène se reflète pour chacun des radiologistes participant à l'étude, comme le montre l'annexe 3. Les paramètres statistiques sont variables d'un individu à l'autre. Ici encore, ils sont légèrement plus élevés pour la lecture sur films que pour la lecture à l'écran. Par contre, elles sont très élevées pour chacun des radiologistes lorsque la lecture à l'écran est comparée à la lecture sur films.

La comparaison inter-radiologistes que montre l'annexe 3 ne permet pas de se prononcer sur la valeur distinctive entre la lecture à l'écran et la lecture sur film. Elle montre la variabilité naturelle entre les lecteurs. Cette variabilité est basée sur des critères subjectifs comme l'expérience, les connaissances, la méthodologie de lecture et l'état mental des lecteurs durant l'étude. Une telle variabilité demeure quelque soit le mode de lecture.

Tableau 10
Analyse statistique de la détection de l'ensemble des anomalies par la lecture des images numériques à l'écran et des images originales sur films.

Ensemble des anomalies identifiées lors de la lecture des images

Mode de lecture	Présentes sur les images			Absentes des images, mais créées suite à une erreur d'interprétation par au moins 1 des 7 radiologistes			Total		
	Abdo	Poumons	Os	Abdo	Poumons	Os	Abdo	Poumons	Os
Écran +	161	361	339	0	12	18	161	372	357
Écran -	196	241	242	28	93	136	224	334	378
Film +	171	385	355	2	11	20	173	396	375
Film -	186	217	226	26	94	134	212	311	360
Total	357	602	581	28	105	154	385	707	735

Statistiques	Abdomen	Poumons	Os
Sensibilité écran (%)	45,1%*	60,0%	58,3%
Spécificité écran (%)	100,0%*	88,6%	88,3%
VPP écran (%)	100,0%*	96,8%	95,0%
VPN écran (%)	12,5%*	27,8%	36,0%
Exactitude écran (%)	49,1%*	64,2%	64,6%
Kappa écran	0,11	0,26	0,30
Sensibilité film (%)	47,9%*	64,0%	61,1%
Spécificité film (%)	92,9%	89,5%	88,2%*
VPP film (%)	98,8%	97,2%	94,7%*
VPN film (%)	12,3%*	30,2%	37,2%
Exactitude film (%)	51,2%*	67,8%	66,5%
Kappa film	0,14	0,33	0,34

* Différence significative entre les valeurs des systèmes (95% IC)

4.3.2 Résultats obtenus par système anatomique

Une analyse séparée de chacun des systèmes anatomiques (abdominal, pulmonaire et osseux) a été réalisée. Nous observons une différence significative dans la performance des radiologistes entre les deux modes de lecture pour le système abdominal. Pour les deux autres systèmes anatomiques, la différence n'est pas significative (voir tableau 10). Le tableau 10 montre que, pour chacun des systèmes anatomiques étudiés, les valeurs de spécificité et VPP sont très élevées alors que les valeurs de VPN sont basses. La valeur de 100% pour la lecture à l'écran des cas

d'abdomen est due au fait qu'il n'y a eu aucune identification de fausses anomalies lors de la lecture à l'écran. Par contre, deux fausses anomalies, une essentielle et un accessoire, ont été détectées lors de la lecture sur film. Finalement, l'ensemble des valeurs statistiques sont légèrement, mais non significativement, supérieures pour le mode de lecture sur films par rapport à la lecture à l'écran.

Ces résultats nous amènent à la prudence quant aux conclusions que nous pouvons en tirer puisque l'utilisation de seulement neuf cas d'abdomen pourrait ne pas permettre la généralisation des capacités de lecture de notre équipe de radiologistes en conditions cliniques normales.

Tableau 11
Comparaison directe de la détection de l'ensemble des anomalies par la lecture des images numériques à l'écran par rapport à la lecture des images originales sur film.

Ensemble des anomalies identifiées à l'écran	Ensemble des anomalies identifiées sur le film						Total		
	détectées			non détectées			Ado	Poumons	Os
	Abdo	Poumons	Os	Abdo	Poumons	Os			
détectées	161	371	355	0	1	0	161	372	355
non détectées	12	24	18	205	310	360	217	334	378
Total	173	395	373	205	311	360	378	706	733

Statistiques	Abdomen	Poumons	Os
Sensibilité écran (%)	93,1%	93,9%	95,2%
Spécificité écran (%)	100,0%	99,7%	100,0%
VPP écran (%)	96,8%	96,5%	97,5%
VPN écran (%)	100,0%	99,8%	100,0%
Exactitude écran (%)	94,5%	92,8%	95,2%
Kappa	0,94	0,93	0,95

La comparaison de la lecture à l'écran par rapport à la lecture sur film montre une forte correspondance dans la capacité des radiologistes de détecter des anomalies sur les images. Toutes les valeurs statistiques incluant les coefficients Kappa du tableau 11 en témoignent par des valeurs dépassant les 92%. Les valeurs de spécificité de 100% pour les systèmes abdominaux et osseux indiquent que les fausses anomalies qui ont été détectées à l'écran l'ont aussi été sur le film, d'où les valeurs zéro du tableau 11.

Les résultats statistiques présentés au tableau 9 et au tableau 11 indiquent une même capacité des radiologistes de détecter des anomalies sur des images radiologiques, que ces dernières soient numériques et affichées sur un moniteur, ou analogiques et illuminées par un négatoscope. Toutes les valeurs statistiques dépassent les 92% lorsque le mode de lecture à l'écran est comparé au mode de lecture sur film.

Cependant, ces données ne tiennent pas compte du niveau de confiance que les radiologistes de l'équipe de recherche avaient par rapport à leur propre capacité de lire des images numériques sur un écran d'ordinateur, alors qu'ils étaient très familiers avec la lecture de film sur négatoscope. La prochaine section établit le niveau de confiance des radiologistes de l'équipe face à la lecture d'images numériques sur un écran d'ordinateur.

4.3.3 Niveau de confiance des radiologistes face à la lecture d'images numériques sur un écran d'ordinateur

Nous avons demandé à l'équipe des 7 radiologistes participant à cette étude, de qualifier les images numériques affichées sur le moniteur et de les comparer aux images originales sur films. Dans cette étape de l'étude, la lecture sur film a été utilisée comme étant le standard. Les images étaient

donc considérées comme parfaites, même si en réalité elles ne l'étaient pas. L'annexe 4 détaille les résultats pour chacun des trois systèmes anatomiques choisis.

Cinq catégories de qualité des images ont été utilisées pour comparer la qualité des images numériques par rapport aux films : nettement moindre, légèrement moindre, égale, légèrement supérieure et nettement supérieure. Il s'avère que seule la catégorie « nettement moindre » était considérée comme inacceptable à la lecture diagnostique par l'équipe de radiologistes participant à cette étude. Les quatre autres catégories se sont donc retrouvées à la classe des images de qualité diagnostique. Le tableau 12 et les tableaux de l'annexe 4 montrent ce classement.

L'ensemble des 61 cas cliniques revus par les 7 radiologistes de l'équipe représente un total de 427 évaluations. Treize évaluations (3,0%) présentaient des images numériques jugées de qualité non diagnostique comme le montre le tableau 13. Par contre, cinq évaluations (1,1%) ont présenté une image numérique dont la qualité perçue dépassait la qualité de l'image originale. Ceci est possible grâce au jeu de modifications du contraste et de la brillance de l'image numérique rendu disponible à l'utilisateur par le logiciel de visualisation. L'image rehaussée a donc été jugée subjectivement comme étant de qualité supérieure à la qualité perçue de l'image analogique correspondante par 5 des 7 radiologistes participant à l'étude. Il est à noter que l'image analogique ne peut être rehaussée. Il s'agit là d'un atout intéressant des systèmes numériques.

L'équipe de radiologistes participant à cette étude se disait capable d'identifier les anomalies sur l'image numérique avec au moins autant d'aisance que sur l'image analogique dans 77,7% des évaluations. Malgré cela, ils se disaient confiants

d'effectuer sur l'image numérique un aussi bon diagnostic qu'ils le feraient sur l'image

analogique dans 68,8% des évaluations étudiées.

Tableau 12
Comparaison des images numériques par rapport aux images originales analogiques pour l'ensemble des évaluations effectuées (n = 427).

	Qualité de l'image numérique par rapport à l'image originale				
	Nettement moindre	Légèrement moindre	Égale	Légèrement meilleure	Nettement meilleure
	Non acceptable	Acceptable à la lecture diagnostique			
Identification de l'anomalie	8 (1,9%)	87 (20,4%)	330 (77,3%)	2 (0,4%)	0
Confort / Confiance diagnostique	11 (2,6%)	122 (28,6%)	292 (68,4%)	2 (0,4%)	0

Pour chacun des systèmes anatomiques étudiés, plus de 94% des clichés cliniques étaient, selon les radiologistes de l'équipe, de qualité propice à l'établissement d'un diagnostic : 98,4% des clichés d'abdomen, 94,6% des clichés de poumons et 98,5% des clichés ostéo-articulaires, comme le montre le tableau 13. La qualité des images numériques était jugée égale ou supérieure à la qualité des images originales dans 66,7% des

clichés d'abdomen, 48,8% des clichés de poumons et 72,9% des clichés ostéo-articulaires. La proportion plus faible de 48,8% des clichés pulmonaires se démarque par rapport aux deux autres proportions qui s'approchent des 70%, mais la proportion des clichés pulmonaires de qualité diagnostique correspond à celle des deux autres systèmes anatomiques.

Tableau 13
Comparaison de la qualité des images numériques par rapport aux images originales analogiques pour les trois systèmes étudiés : abdominal, pulmonaire et osseux.

Qualité des clichés du système	Qualité de l'image numérique par rapport à l'image originale			
	Non diagnostique	Diagnostique		
		Moindre	Égale	Meilleur
abdominal (n = 63)	1 (1,6%)	20 (31,7%)	40 (63,5%)	2 (3,2%)
pulmonaire (n = 168)	9 (5,4%)	77 (45,8%)	82 (48,8%)	0
osseux (n = 196)	3 (1,5%)	50 (25,5%)	140 (71,5%)	3 (1,5%)
Total des clichés (n = 427)	13 (3,0%)	147 (34,4%)	262 (61,5%)	5 (1,1%)

Dans le cas du système pulmonaire, les radiologistes se disaient incapables d'identifier suffisamment d'anomalies sur

l'image numérique pour porter un diagnostic dans 3,6% des cas. Mais, ils se disaient être capables d'identifier les anomalies sur

l'image numérique avec au moins autant d'aisance que sur l'image analogique dans 74,4% des cas étudiés. Malgré cette performance, ils se sentaient inconfortables à porter un diagnostic dans 4,8% des cas, mais ils se sentaient confiants d'effectuer un aussi bon diagnostic sur l'image numérique que sur l'image originale dans 59,5% des cas étudiés. Parmi les structures spécifiques à ce type d'images, le niveau d'inconfort diagnostic était le plus élevé pour le parenchyme pulmonaire (2,4% des cas); il était de 1,2% pour les contours cardiaques et la cage thoracique. Les radiologistes se disaient être capables d'identifier sur l'image numérique les contours cardiaques et la cage thoracique avec au moins autant d'aisance que sur l'image analogique dans 66,1% et 59,5% des cas respectivement. Mais la proportion était moindre (42,9%) pour le parenchyme pulmonaire. Finalement, la latitude de l'image numérique, c'est-à-dire l'intervalle des niveaux de gris présenté à l'écran, était jugé trop faible pour porter un diagnostic dans 1,2% des cas, mais elle était égale ou supérieure à celle de l'image analogique dans 44% des cas étudiés (voir tableau 16 de l'annexe 4).

Dans le cas du système osseux, les radiologistes se disaient incapables d'identifier suffisamment d'anomalies sur l'image numérique pour porter un diagnostic dans 1,0% des cas. Mais, ils se disaient être capables d'identifier les anomalies sur l'image numérique avec au moins autant d'aisance que sur l'image analogique dans 82,7% des cas étudiés. Correspondant à cette performance, ils se sentaient inconfortables à porter un diagnostic dans 1,0% des cas, mais ils se sentaient confiants d'effectuer un aussi bon diagnostic sur l'image numérique que sur l'image originale dans 77,5% des cas étudiés. Parmi les structures spécifiques à ce type d'images, le niveau d'inconfort

diagnostic était le plus élevé pour les trabéculations osseuses (2,0% des cas); il était de 0,5% et 0% pour les tissus mous et la ligne graisseuse. Les radiologistes se disaient être capables d'identifier sur l'image numérique les trabéculations osseuses, les tissus mous et les lignes graisseuses avec au moins autant d'aisance qu'ils le faisaient sur l'image analogique dans 52,0%, 92,8% et 92,4% des cas, respectivement. Finalement, la latitude de l'image numérique était jugée trop faible pour porter un diagnostic dans 2,0% des cas, mais elle était égale ou supérieure à celle de l'image analogique dans 67,9% de cas étudiés (voir tableau 17 de l'annexe 4).

Finalement, dans le cas du système abdominal, aucune structure particulière n'a pu être identifiée, que ce soit par son contraste marqué ou que ce soit pour son utilité clinique à identifier des pathologies particulières. Une particularité des images d'abdomen est que toutes les structures, sauf le côlon, ont une densité tissulaire qui s'approche de celle de l'eau. Il n'y avait pas lieu de faire de la latitude un critère spécifique de qualité d'image, comme ce fût le cas pour les deux autres systèmes anatomiques. Deux critères ont donc été retenus : la capacité d'identifier les anomalies et la confiance diagnostique. Dans 71,4% des cas étudiés, les radiologistes se disaient être capables d'identifier les anomalies sur l'image numérique avec au moins autant d'aisance que sur l'image analogique, mais ils se disaient être capables d'identifier suffisamment d'anomalies pour porter un diagnostic dans 100% des cas. Ils se sentaient inconfortables à porter un diagnostic dans 1,6% des cas, mais ils se sentaient confiants d'effectuer un aussi bon diagnostic sur l'image numérique que sur l'image originale, dans 66,7% des cas étudiés (voir tableau 18 de l'annexe 4).

CHAPITRE 5

5. DISCUSSION

5.1 Méthodologie utilisée

Cette étude consistait à évaluer la capacité des radiologistes d'identifier les anomalies qui mènent à l'établissement d'un diagnostic par la visualisation objective d'images radiologiques, faisant ainsi abstraction du processus mental complexe mis en branle dans l'élaboration du diagnostic. Un tel processus est basé sur la connaissance et tout le savoir faire du médecin établissant le diagnostic. Cette façon de procéder diffère de celles publiées dans la littérature où la capacité d'effectuer un diagnostic est directement évaluée [Johnston & Worrall, 2003]. De nombreux résultats semblent montrer qu'une certaine réduction de la qualité des images visualisées ne semble pas interférer avec la précision du diagnostic [O'Sullivan et coll., 1997].

La méthodologie utilisée comporte cependant une limitation : les deux types d'images numériques et analogiques ont été lus consécutivement à l'intérieur de la même session en commençant par la lecture des images numériques. Notre hypothèse de départ était que la capacité de détection des anomalies lors de la lecture des images numériques serait moindre que la capacité de détection lors de la lecture des films. Cependant, la lecture du film a pu être effectuée en complément à la lecture à l'écran et non de façon indépendante. Ainsi, la connaissance de l'image numérique aurait pu influencer sur la manière de lire la même image sur film. Malgré l'insistance de l'assistant de recherche auprès des radiologistes de lire

chaque image de façon indépendante, nous ne pouvons prétendre à la complète indépendance de la lecture de chacune des images durant la session de visualisation.

Une révision de toutes les images numériques et analogiques, postérieurement à l'étude clinique, nous a permis de constater que les fausses anomalies détectées par les radiologistes provenaient d'une mauvaise interprétation de structures anatomiques visualisées sur les images, qu'elles soient numériques ou analogiques, et non d'artéfacts causés par le développement des films ou par la numérisation des films.

5.2 Chaîne d'imagerie améliorée

La méthode utilisée pour évaluer techniquement le système de numérisation des films est conceptuellement similaire à la méthode utilisée par l'équipe de Davidson [Davidson et coll., 2001] dans l'établissement de la courbe de transfert d'un numériseur de films. Notre méthode d'évaluation a dépassé le cadre du numériseur en considérant l'ensemble de la chaîne d'imagerie. La valeur maximale de l'intensité lumineuse du moniteur, inférieure à celle du négatoscope, réduit l'intervalle des niveaux de gris perceptible au lecteur. Il s'agit là d'une limitation du moniteur qui a été notée par l'ensemble des radiologistes, particulièrement lors de la visualisation des images du système pulmonaire. Le remplacement du moniteur conventionnel par un moniteur de haute brillance, tel qu'utilisé dans les systèmes PACS,

permettrait d'augmenter la valeur maximale d'intensité lumineuse et donc, de superposer davantage la fonction de transfert du moniteur sur la fonction de transfert du négatoscope. Il est cependant essentiel que le moniteur choisi soit capable d'enregistrer une table de visualisation, de manière à l'appliquer sur l'image juste avant son affichage. Le coût d'un tel moniteur est de l'ordre de 12 000 \$, comparativement à environ 1 000 \$ pour un moniteur conventionnel de bonne qualité.

Le processus de numérisation des films réduit la résolution spatiale de l'ordre de 8 paires de lignes par millimètre pour le film à 5,1 paires de lignes par millimètre pour l'image numérisée, si la numérisation se fait sur la largeur totale du film, soit 35 centimètres. Malgré cette perte de résolution spatiale calculée, la résolution spatiale effective mesurée avec une mire Nuclear Associates est la même pour les deux types d'images, soit 2,5 paires de lignes par millimètre (voir tableau 6). Deux radiologistes ont mentionné que pour quelques cas seulement, leur perception des détails sur l'image numérique semblait réduite par rapport au film. Considérant la même valeur de résolution spatiale effective pour les deux modalités de lecture, soit 2,5 paires de lignes par millimètre, il est possible que cette difficulté de percevoir certains détails exprimés par les deux radiologistes de l'équipe, soit due au niveau de brillance moins élevé du moniteur et non pas à une perte supplémentaire de la résolution due à la numérisation des films.

5.3 Qualité des images numériques

Les résultats de notre étude indiquent que 96% des images numérisées avec un numériseur de films et lues sur un moniteur

conventionnel rendu linéaire et bien calibré, sont considérées de qualité propice à l'établissement d'un diagnostic. De plus, la comparaison des deux modes de lecture évalués dans cette étude montre des valeurs statistiques qui dépassent les 90% et donc une capacité de lecture à l'écran comparable à celle effectuée sur le négatoscope. Par conséquent, un centre hospitalier qui n'a pas les moyens de se procurer d'une haute technologie de numérisation pourrait utiliser ce type de système pour numériser ses films radiologiques, à la condition que 1) la chaîne d'imagerie soit linéaire et que 2) chaque composante du système soit bien calibrée.

Les valeurs de sensibilité et de spécificité obtenues dans notre étude concordent avec celles d'autres études similaires [Calder et coll., 1999; Eng et coll., 2000; Tachakra, 2002; Johnston et Worrall, 2003], mais diffèrent études visant des pathologies précises [Larson et coll., 1997; Calder et coll., 1999; Palombo et coll., 2002; Abdoh et coll., 2003]. Dans ces dernières études, les valeurs de sensibilité et de spécificité s'approchent ou dépassent les 90%, sauf pour l'étude de Eng ($\geq 85\%$). La différence entre les valeurs de sensibilité et les valeurs de spécificité est positive ou négative, selon l'étude. Aucune tendance n'a été notée.

Certaines études visaient à comparer deux technologies sur la capacité des lecteurs à détecter des éléments particuliers de l'image, appelés anomalies [Scott et coll., 1995; Larson et coll., 1997]. Les valeurs de sensibilité étaient systématiquement plus basses que les valeurs de spécificité qui dépassaient les 90% en général, comme c'est le cas pour notre étude. Les valeurs de sensibilité obtenues dans notre étude, pour chaque mode de lecture, sont de l'ordre de 57% alors que les valeurs de spécificité avoisinent les 90% et ce, pour les deux modes de lecture étudiée. La différence de

2% à 4% entre les valeurs obtenues pour la lecture à l'écran par rapport au film n'est pas significative ($P > 0,05$).

Les valeurs statistiques obtenues lors de la comparaison des deux méthodes de lecture montrent que l'utilisation de l'imagerie numérique provenant du système étudié ne diminue pas la capacité des radiologistes à détecter des anomalies sur les images radiologiques, une conclusion que ne partage pas Calder et collègues [Calder et coll., 1999] qui évaluèrent un système similaire. Cependant, ces auteurs ne mentionnent pas s'ils ont effectué des tests de linéarité de la chaîne d'imagerie du système utilisé dans leur étude. Nous pouvons présumer que ce système affichait un comportement non linéaire qui n'a pas été corrigé avant le début de l'étude clinique de comparaison des images.

5.4 Niveau de confiance des radiologistes face à la lecture d'images numériques

Le jugement apporté par les radiologistes de l'équipe concernant leur capacité d'identifier des anomalies et leur confiance à établir un diagnostic à partir des images lues est de nature très subjective. Globalement, ce jugement reflète les résultats objectifs de performance obtenus lors de l'évaluation des deux modes de lecture, mais il ne reflète pas ces résultats si on ne considère que le seul fait de se sentir aussi à l'aise face à la lecture de l'image numérique qu'à la lecture du film. Ce dernier est le médium sur lequel les radiologistes sont habitués de travailler depuis de nombreuses années. Les radiologistes les plus méfiants étaient les plus expérimentés, mais en même temps, ils étaient les moins familiers à manipuler un ordinateur. Par contre, les radiologistes les

plus confiants étaient les plus jeunes qui ont grandi dans l'ère de l'informatique, mais qui avaient moins d'expérience en radiologie, bien que leur expérience était non négligeable.

Les radiologistes de l'équipe expriment deux niveaux de confiance en leur capacité de détecter des anomalies sur l'image numérique. La première expression du confort diagnostique correspond à leur niveau de confiance d'être capable d'établir un diagnostic sur les images numériques. Ils se disent confiants de pouvoir établir un diagnostic sur plus de 95% de tous les clichés étudiés : 95,2% des clichés de poumons, 98,4% des clichés d'abdomen et 99% des clichés ostéo-articulaires. Ils se disent être capables d'identifier suffisamment d'anomalies pour établir un diagnostic sur les images numériques de plus de 96% des cas étudiés : 96,4% des cas de poumons, 100% des cas d'abdomen et 99% des cas ostéo-articulaires. Cette première expression du confort diagnostique concorde avec les résultats objectifs de performance obtenus dans cette étude.

La seconde expression du confort diagnostique correspond à leur sentiment personnel de leur capacité de lire l'image numérique avec autant d'aisance que de lire un film. En effet, les radiologistes indiquent qu'ils se sentent confiants à établir un aussi bon diagnostic à la lecture des images numériques qu'à la lecture des films dans 59,5% des cas de poumons, 66,7% des cas d'abdomen et 77,5% des cas ostéo-articulaires. Similairement, ils se disent être capables d'identifier les anomalies sur les images numériques avec au moins autant d'aisance qu'ils le feraient sur les films dans 74,4% des cas de poumons, 71,4% des cas d'abdomen et 82,7% des cas ostéo-articulaires. Les valeurs de cette seconde expression du confort diagnostique sont beaucoup plus faibles que celles

correspondant à la première expression du confort diagnostique. Ceci reflète le manque d'expérience des lecteurs avec le médium numérique et non pas leur capacité moins grande d'établir un diagnostic. Les radiologistes expriment leur malaise face à l'habitude de manipuler et de travailler avec l'image numérique. Ceci est tout à fait normal considérant que peu d'entre eux avaient déjà travaillé avec un PACS.

La considération des trois éléments importants de la qualité d'une image, soit la résolution spatiale, le contraste et la latitude, permettrait l'amélioration possible de la performance de l'équipement. D'abord, la perte de résolution spatiale de 8 à 5,1 paires de lignes par millimètre due au processus de numérisation n'a pas eu de conséquence sur la qualité visuelle des images. En effet, la résolution spatiale effective était la même sur les deux types d'images, soit 2,5 paires de lignes par millimètre, ce qui permet de distinguer des structures de 200µm de diamètre séparées par une distance de 200 µm. Ensuite, l'algorithme de correction appliquée sur les images numériques juste avant leur affichage a permis de retrouver les niveaux de contrastes du film. Finalement, l'image numérique corrigée peut afficher la quasi-totalité de la latitude du film test, telle que démontré dans l'évaluation technique du système, mais avec un niveau de brillance inférieur à celui du négatoscope. La latitude perceptible de l'image numérique était jugée inférieure à celle du film éclairé avec un négatoscope de brillance plus élevée.

Ce phénomène s'est reflété dans les trois systèmes anatomiques étudiés. Par exemple, les radiologistes ont indiqué que dans 45,9% des cas ostéo-articulaires étudiés, la qualité des trabéculations osseuses affichées sur les images numériques était légèrement moindre que celle affichée sur les films. La différence entre le blanc et le noir était inférieure sur l'image numérique à celle affichée sur le

film. Cette même affirmation diminution de la qualité s'appliquait aux tissus mous dans 6,6% des cas et aux lignes graisseuses dans 7,7% des cas (voir tableau 17 de l'annexe 4). Ces résultats concordent avec ceux de l'équipe de Larson [Larson et coll., 1998] dans laquelle les trabéculations osseuses affichées à l'écran étaient jugées de qualité égale ou supérieure à celle de ces mêmes structures anatomiques affichées sur les films (56% vs. 52% dans notre étude), mais pas pour les tissus mous (69% vs. 92,8% dans notre étude).

Ce phénomène se reflète aussi dans les images dans lesquelles les radiologistes ont indiqué que le parenchyme pulmonaire sur l'image numérique était de qualité légèrement inférieure à celle du film dans 54,8% des cas. Il était plus difficile de visualiser les fins changements des niveaux de gris car les images numériques paraissaient plus foncées que les films. La réduction de la latitude perceptible sur l'image numérique aurait augmenté globalement la densité optique du poumon, rendant plus difficile la visualisation du parenchyme pulmonaire. Ce niveau de difficulté accru a diminué de façon marquée la confiance diagnostique des radiologistes. Ces résultats concordent avec ceux obtenus par Larson et son équipe pour le parenchyme pulmonaire (44% vs. 42,8% pour notre étude), mais les dépassent pour les autres éléments (45% pour le médiastin (vs. 66,1% pour notre étude) et la résolution spatiale (vs. 59,5% pour notre étude).

Finalement, les images d'abdomen exigent davantage en fin contraste et latitude qu'en résolution spatiale. L'intervalle des niveaux de gris utilisé pour afficher ce type d'images est relativement petit puisque l'ensemble des structures anatomiques de l'abdomen possède une densité qui s'approche de celle de l'eau. La qualité de l'image numérique était jugée par les

radiologistes comme étant légèrement moindre, mais tout à fait acceptable, dans 28,6% des cas. La luminosité maximale du moniteur étant plus faible que celle du négatoscope, augmentait artificiellement la densité optique globale de l'image. Ce phénomène réduisait ainsi la capacité des radiologistes d'identifier les faibles contrastes.

Face à ces résultats, deux actions pourraient être prises pour améliorer le niveau de confiance. D'abord, il serait indiqué de former adéquatement les utilisateurs à la manipulation de l'appareil et de leur donner suffisamment de temps pour qu'ils s'habituent à lire les images numériques avant d'utiliser l'appareil de façon routinière en clinique. Ensuite, le moniteur utilisé pourrait être remplacé par un moniteur de haute brillance. Ceci permettrait d'augmenter la brillance maximale, rapprochant ainsi la fonction de transfert du moniteur de celle du négatoscope, et par conséquent, augmenterait la latitude perceptible des niveaux de gris. Cette action améliorerait la performance et le niveau de confiance des radiologistes qui utiliseraient cette technologie.

5.5 Généralisation des résultats

L'étude ne comportait pas d'évaluation du système de téléradiologie à distance. Il ne s'avérait pas nécessaire d'effectuer ce test car il n'influerait pas sur la qualité des images. Le format d'images utilisé pour la transmission sur un réseau informatique est DICOM, un format universellement accepté par toutes les firmes d'imagerie médicale et de PACS. Une image DICOM est difficile à

corrompre et si cela survient, ça se voit facilement. Nous estimons que nos résultats auraient été les mêmes au plan de la qualité des images. L'impact majeur de la transmission sur réseau des images DICOM se manifesterait par un délai de transmission et non sur la qualité des images. Nos résultats peuvent donc s'appliquer à tout processus de lecture à distance d'images radiographiques numérisées avec un numériseur de films et transmises sur un réseau informatique.

Les résultats de cette étude peuvent aussi se généraliser à toute activité de lecture à distance d'images radiologiques numériques provenant d'une salle d'urgence de n'importe quel hôpital régional pour les raisons suivantes. D'abord, les cas cliniques utilisés dans cette étude sont typiques des cas rencontrés dans les salles d'urgence des hôpitaux régionaux du Québec. Les trois systèmes anatomiques choisis sont ceux qui sont les plus fréquemment radiographiés dans les salles d'urgence. Ensuite, notre équipe de radiologistes est composée d'individus qui lisent régulièrement des radiographies provenant de salles d'urgence dans leur activité clinique. La variabilité de leur expérience reflète l'expérience professionnelle de l'ensemble des radiologistes qui lisent des cas de salles d'urgence. Finalement, les radiographies utilisées ont été prises avec des moyens techniques qui reflètent ceux utilisés dans les hôpitaux régionaux du Québec. Les équipements, les techniques de positionnement des patients, ainsi que les techniques radiologiques, sont typiques des examens faits à l'urgence et sont généralisés à travers les hôpitaux du Québec. L'ensemble de ces éléments réunit nous permet de conclure à la généralisation de nos résultats.

CHAPITRE 6

6. CONCLUSION

Nous pouvons conclure de cette étude que l'utilisation d'un système modeste de téléradiologie, basé sur la numérisation des films et l'affichage des images sur un moniteur conventionnel, peut produire des images de qualité propice à l'établissement d'un diagnostic à distance, à condition que :

1. la chaîne d'imagerie ait un comportement linéaire, et
2. chaque composante du système soit calibrée avant d'être utilisée.

Cette conclusion est basée sur deux résultats. D'abord, les valeurs statistiques de comparaison de la lecture d'images numérisées par rapport au standard établi, le film, dépassent les 93%, indiquant que la performance des radiologistes est équivalente quelque soit le mode de lecture des images. Ensuite, le niveau de confiance des radiologistes à établir un diagnostic sur les images numérisées dépasse les 95%.

Cependant, le niveau de confiance des radiologistes à établir sur l'image numérisée est aussi bon diagnostic que sur le film est nettement inférieur (59,5% - 82,7%),

reflétant 1) un malaise face à leur inexpérience à utiliser l'outil, plutôt que leur inhabilité à faire un diagnostic sur une image numérique, et 2) une avenue possible d'amélioration de l'outil.

Deux actions sont proposées pour améliorer le niveau de confiance des lecteurs :

1. former adéquatement les utilisateurs à l'utilisation de l'outil et leur donner suffisamment de temps pour qu'ils s'habituent à lire les images numériques avant d'utiliser l'appareil de façon routinière en clinique;
2. remplacer le moniteur conventionnel de l'équipement par un moniteur de haute brillance. Ceci rapprocherait la courbe de transfert du moniteur de celle du négatoscope, améliorant ainsi suffisamment le degré de latitude perceptible des niveaux de gris affichés à l'écran pour augmenter de façon appréciable la confiance diagnostique des radiologistes.

CHAPITRE 7

7. RECOMMANDATION

Un centre hospitalier qui ne peut se procurer une haute technologie de numérisation des images radiologiques telle un PACS pourrait faire l'acquisition de ce type de système pour numériser ses films radiologiques, selon les critères suivants :

1. le moniteur fournit en soi un de haute brillance, capable d'enregistrer une nouvelle table de calibrage et pouvant fournir une résolution minimale de 1 600 x 1 200 pixels;
2. le numériseur de films fournit avec l'équipement ait une résolution qui satisfait les critères de l'Association canadienne de radiologie pour la plus petite dimension de films à numériser;
3. la chaîne d'imagerie ait un comportement linéaire;
4. chaque composante du système soit bien calibrée selon les spécifications des manufacturiers;
5. le logiciel de visualisation permet les fonctions utiles aux utilisateurs telles qu'évaluées par le radiologiste membre du comité d'acquisition des équipements médicaux.

ANNEXE I

TYPES DE CAS CLINIQUES UTILISÉS

Poumons (24)

Hémithorax	Nodules
Sinus	Atélectasie
Pneumonie	Pneumothorax
Pneumothorax & emphysème	Pneumothorax & fracture de côtes
Insuffisance cardiaque	Tumeur & fractures de côtes
Œdème pulmonaire	Opacité alvéolaire diffuse
Stade précoce de pneumothorax	Infection pleurale & fracture de côtes
Pneumonie avec accumulation de liquide pleural	
Dislocation de l'épaule droite avec fracture de la clavicule	
Normal (4)	

Os (28)

Épaule & clavicule	Thorax & côtes & rachis dorsal
Coude	Poignet & main
Cheville & pied	Genou
Jambe	Normal (4)

Abdomen (9)

Os iliaque	Urolithiase
Abcès	Colite
Ligament falciforme	Obstruction
Normal (1)	

ANNEXE II FORMULAIRES

FORMULAIRES

IMAGES POUMONS _____

image.tif

Anomalies sur image numérique	Cocher
Opacité alvéolaire dans le poumon droit	
Cicatrice à l'apex cardiaque	

Anomalies supplémentaires sur image analogique

Comparaison de l'image numérique par rapport à l'image analogique :					
	Moins bon		Égal	Meilleur	
	--	-	0	+	++
Identification de l'anomalie					
Confort / Confiance diagnostique					
Parenchyme pulmonaire					
Ligne médiastinale & transparence rétrocardiaque					
Cage thoracique					
Latitude générale de l'image					
Qualité de l'image globale					

Qualité des clichés	Qualité diagnostique	Cocher
Analogique > Numérique	Diagnostique	
Analogique = Numérique	Diagnostique	
Analogique < Numérique	Diagnostique	
Analogique > Numérique	NON Diagnostique	

Commentaires :

Anomalies sur image numérique	Cocher
Fractures portion inférieure des vertèbres D6, D8 – D12	
Air rétropéritonéal	

Anomalies supplémentaires sur image analogique

Comparaison de l'image numérique par rapport à l'image analogique :					
	Moins bon		Égal	Meilleur	
	--	-	0	+	++
Identification de l'anomalie					
Confort / confiance diagnostique					
Trabéculations osseuses					
Tissus mous					
Lignes graisseuses					
Latitude générale de l'image					
Qualité de l'image globale					

Qualité des clichés	Qualité diagnostique	Cocher
Analogique > Numérique	Diagnostique	
Analogique = Numérique	Diagnostique	
Analogique < Numérique	Diagnostique	
Analogique > Numérique	NON Diagnostique	

Commentaires :

Anomalies sur image numérique	Cocher
Scoliose dorsolombaire gauche positionnelle	
Spina bifida occulta	

Anomalies supplémentaires sur image analogique

Comparaison de l'image numérique par rapport à l'image analogique :					
	Moins bon		Égal	Meilleur	
	--	-	0	+	++
Identification de l'anomalie					
Confort / confiance diagnostique					
Qualité de l'image globale					

Qualité des clichés	Qualité diagnostique	Cocher
Analogique > Numérique	Diagnostique	
Analogique = Numérique	Diagnostique	
Analogique < Numérique	Diagnostique	
Analogique > Numérique	NON Diagnostique	

Commentaires :

ANNEXE III

PERFORMANCE DE CHAQUE RADIOLOGISTE

Tableau 14

Analyse statistique de la détection de l'ensemble des anomalies (essentiels et accessoire) de chaque radiologiste participant à l'étude, lors de la lecture des images numériques sur écran et des images originales sur films.

Paramètres statistiques	Radiologistes participant à l'étude							(95% I.C.)
	#1	#2	#3	#4	#5	#6	#7	
Sensibilité écran (%)	66,8%	51,4%	61,4%	64,1%	55,0%	49,1%	43,6%	55,9% ± 5,8%
Spécificité écran (%)	97,6%	100,0%	85,4%	92,7%	77,5%	89,7%	90,2%	91,8% ± 5,7%
VPP écran (%)	99,3%	100,0%	95,7%	97,9%	93,1%	96,4%	96,0%	96,9% ± 1,8%
VPN écran (%)	35,4%	27,7%	29,2%	32,5%	23,8%	23,8%	23,0%	27,9% ± 3,5%
Exactitude écran (%)	71,6%	59,0%	65,1%	68,6%	58,5%	55,2%	51,0%	61,3% ± 5,5%
Kappa écran	0,38	0,25	0,26	0,32	0,17	0,18	0,15	0,25 ± 0,06
Sensibilité film (%)	71,4%	53,6%	62,7%	68,2%	57,3%	54,1%	46,8%	59,2% ± 6,0%
Spécificité film (%)	97,6%	100,0%	82,9%	92,7%	80,0%	84,6%	87,8%	89,4% ± 5,2%
VPP film (%)	99,4%	100,0%	95,2%	98,0%	94,0%	95,2%	95,4%	96,7% ± 1,6%
VPN film (%)	38,8%	28,7%	29,3%	35,2%	25,4%	24,6%	23,5%	29,4% ± 3,9%
Exactitude film (%)	75,5%	60,9%	65,9%	72,0%	60,8%	58,7%	53,3%	63,9% ± 5,3%
Kappa film	0,46	0,28	0,28	0,40	0,22	0,25	0,19	0,30 ± 0,07

Tableau 15
Comparaison directe de la lecture des images numériques à l'écran par rapport à la lecture des images originales du film sur le négatoscope, pour chacun des radiologistes participant à l'étude. La lecture de l'image originale est considérée comme la méthode standard.

Paramètres statistiques	Radiologistes participant à l'étude							(95% I.C.)
	#1	#2	#3	#4	#5	#6	#7	
Sensibilité (%)	93,7%	95,8%	97,2%	94,1%	96,3%	89,6%	92,6%	94,2% ± 1,8%
Spécificité (%)	100,0%	100,0%	100,0%	100,0%	99,2%	100,0%	100,0%	99,9% ± 0,2%
VPP (%)	100,0%	100,0%	100,0%	100,0%	99,2%	100,0%	100,0%	99,9% ± 0,2%
VPN (%)	91,2%	96,6%	96,7%	92,3%	96,2%	91,2%	95,0%	94,2% ± 1,7%
Exactitude (%)	96,2%	98,1%	98,5%	96,6%	97,7%	95,0%	96,9%	97,0% ± 0,8%
Kappa	0,92	0,96	0,97	0,93	0,95	0,90	0,94	0,94 ± 0,02

ANNEXE IV

NIVEAU DE CONFIANCE DES RADIOLOGISTES FACE À LA LECTURE DES IMAGES NUMÉRIQUES POUR CHACUN DES SYSTÈMES ANATOMIQUES

Tableau 16
Comparaison des images numériques par rapport aux images originales analogiques pour le système pulmonaire (n = 168).

Système pulmonaire	Qualité de l'image numérique par rapport à l'image originale				
	Nettement moindre	Légèrement moindre	Égale	Légèrement meilleure	Nettement meilleure
	Non acceptable	Acceptable à la lecture diagnostique			
Identification de l'anomalie	6 (3,6%)	37 (22,0%)	124 (73,8%)	1 (0,6%)	0
Confort / Confiance diagnostique	8 (4,8%)	60 (35,7%)	100 (59,5%)	0	0
Capacité d'identifier					
le parenchyme pulmonaire	4 (2,4%)	92 (54,8%)	72 (42,8%)	0	0
les lignes médiastinales et la transparence rétrocardiaque	2 (1,2%)	55 (32,7%)	110 (65,5%)	1 (0,6%)	0
la cage thoracique	2 (1,2%)	66 (39,3%)	100 (59,5%)	0	0
Latitude générale de l'image	2 (1,2%)	92 (54,8%)	74 (44,0%)	0	0

Tableau 17
Comparaison des images numériques par rapport aux images originales analogiques pour le système osseux (n = 196).

Système osseux	Qualité de l'image numérique par rapport à l'image originale				
	Nettement moindre	Légèrement moindre	Égale	Légèrement meilleure	Nettement meilleure
	Non acceptable	Acceptable à la lecture diagnostique			
Identification de l'anomalie	2 (1,0%)	32 (16,3%)	162 (82,7%)	0	0
Confort / Confiance diagnostique	2 (1,0%)	42 (21,5%)	151 (77,0%)	1 (0,5%)	0
Capacité d'identifier					
les trabéculations osseuses	4 (2,0%)	90 (46,0%)	102 (52,0%)	0	0
les tissus mous	1 (0,6%)	13 (6,6%)	169 (86,2%)	13 (6,6%)	0
la ligne graisseuse	0	15 (7,7%)	176 (89,8%)	5 (2,5%)	0
Latitude générale de l'image	4 (2,0%)	59 (30,1%)	124 (63,3%)	9 (4,6%)	0

Tableau 18
Comparaison des images numériques par rapport aux images originales analogiques pour le système abdominal (n = 63).

Système abdominal	Qualité de l'image numérique par rapport à l'image originale				
	Nettement moindre	Légèrement moindre	Égale	Légèrement meilleure	Nettement meilleure
	Non acceptable	Acceptable à la lecture diagnostique			
Identification de l'anomalie	0	18 (28,6%)	44 (69,8%)	1 (1,6%)	0
Confort / confiance diagnostique	1 (1,6%)	20 (31,7%)	41 (65,1%)	1 (1,6%)	0

BIBLIOGRAPHIE

- Abdoh, A. A., Krousel-Wood, M. A., Re, R. N., *Accuracy of telemedicine in detecting uncontrolled hypertension and its impact on patient management.* Telemedicine Journal and e-Health, 9(4):315-323, 2003.
- American College of Radiology, ACR guidelines, www.acr.org, 2002.
- Bodemeyer, J., Mariss, G., Gursky, S., *Teleradiology-blessing or curse? 4 years routine in the North Hestia Teleradiology Data Network.*, Radiology, 42(2):71-81, 2002 (Abstract).
- Bouyer, J. Hémon, D., Cordier, S., Derriennic, F., Stücker, Il, Stengel, B., Clavel, J., *Épidémiologie : principes et méthodes quantitatives*, Éditions INSERM, 1995, 498 pages.
- Brumage, M. R., Chinn, S., Cho, K., *Teleradiology in a military training Area*, Journal of Telemedicine & Telecare, 7(6):348-52, 2001.
- Bushberg, J. T., Seibert, J. A., Leidholdt Jr, E. M., Boone, J. M., *The Essential physics of medical imaging*, Éditions Williams & Wilkins, 1994, 742 pages.
- Calder, L. D., Maclean, J. R., Bayliss, A. Pl, Gilbert, F. J., Grant, A. M., *The diagnostic performance of a PC-based teleradiology link.*, Clinical Radiology, 54(10):659-64, 1999.
- Canadian Association of Radiologists, *CAR guidelines*, www.car.ca, 1999.
- Crowe, B. L., Hailey, D. M., de Silva, M., *Teleradiology at a children's hospital: a pilot study*, Journal of Telemedicine & Telecare, 2(4):210-6, 1996.
- Darkins, A., Dearden, C. H., Rocke, L. G., Martin, J. B., Sibson, L., Wooton, R., *An Evaluation of telemedical support for a minor treatment centre.*, Journal of Telemedicine & Telecare, 2(2):93-9, 1996.
- Davidson, H. C., Johnston, D. J., Christian, M. E., Harnsberger, H. R., *comparison of radiographic image quality from four digitization devices as viewed on computer monitors.*, Journal of Digital Imaging, 14(1):24-29, 2001.
- Eng, J., Mysko, W. K., Weller, G. E., Renard, R., Gitlin, J. N., Bluemke, D. A., Magid, D., Kelen, G. D., Scott, W. W. Jr., *Interpretation of emergency department radiographs: a comparison of emergency medicine physicians with radiologists, residents with faculty, and film with digital display.*, American Journal of Roentgenology, 175(5):1233-8, 2000.
- Glowacki, M., Czernicki, Z., Mierzejewski, M., *The first system of image transmission and neurosurgical telecommunication in Poland*, Neurologia & Neurochirurgia Polska, 36(2):363-9, 2002 (Abstract).

- Jithoo, R., Govender, P.V., Corr, P., Nathoo, N., *Telemedicine and neurosurgery: experience of a regional unit based in South Africa.*, Journal of Telemedicine & Telecare, 9 (2):63-6, 2003.
- Johansen, I., *Experiences with teleradiology in general practice in Oppland*, Tidsskrift for Den Norske Laegeforening, 120(16):1896-8, 2000, (Abstract).
- Johnston, K.C., Worrall, B.B., *Teleradiology assessment of computerized tomographs online reliability study (TRACTORS) for acute stroke evaluation.*, Telemedicine Journal & E-Health, 9(3):227-33, 2003.
- Kato, H., *Image reading: Conversion to digital signals*, Computed radiography, éditeurs: Tateno, Y., Iinuma, T. and Takano, M., Édition: Springer-Verlag, 185 pages, 1987.
- Kiuru, M., J., Paakkala, T., A., Kallio, T. T., Aalto, J., Rajamaki, M., *Effect of teleradiology on the diagnosis, treatment and prognosis of patients in a primary care centre.*, Journal of Telemedicine & Telecare, 8(1):25-31, 2002.
- Kondo, Y., *Medical image transfer for emergency care utilizing internet and mobile phone.*, Nippon Hoshasen Gijutsu Gakkai Zasshi, 58(10):1393-401, 2002, (Abstract).
- Larson, A., Lynch, D. A., Zeligman, B., Harlow, C., Vanoni, C., Thieme, G., Kilcoyne, R., *Accuracy of diagnosis of subtle chest disease and subtle fractures with a teleradiology system.* American Journal of Radiology, 170:19-22, 1998.
- Lemieux, R., *Évaluation d'un système de téléradiologie composé d'un numériseur de films*, Mémoire de maîtrise, Université de Montréal, 2004.
- Lim, A.J., *Image quality in film digitizers: testing and quality assurance*, Syllabus: A course in Physics, Radiological Society of North America, December 1996.
- Maass, M., Kosonen, M., Kormano, M., *Transportation savings and medical benefits of a teleneuroradiological network*, Journal of Telemedicine & Telecare, 6(3):142-6, 2000.
- O'Sullivan, D. C., Averch, T. D., Cadeddu, J. A., Moore, R. G., Beser, N., Breitenbach, C., Khazan, R., Kavoussi, L. R., *Teleradiology in urology: comparison of digital image quality with original radiographic films to detect urinary calculi.*, Journal of Urology, 158:2216-2220, 1997.
- Palombo, A., Haigh, T., Furguson, J., Pedley, D., *Can pediatric radiographs be accurately interpreted using an inter-hospital telemedicine system?* Journal of Telemedicine & Telecare, 8 Suppl 2:70-2, 2002.
- Ricci, W. M., Borrelli, J., *Teleradiology in orthopedic surgery: impact on clinical decision making for acute fracture management*, Journal of Orthopaedic Trauma, 16(1):1-6, 2002.

- Scott, W., Bluemke, D., Mysko, W., et al., *Interpretation of emergency department radiographs by radiologists and emergency medicine physicians: teleradiology versus radiograph readings*. Radiology, 195: 223-229, 1995.
- Siebert, J. A., Barnes, G. T., Gould, R. G., *Specification, acceptance testing, and quality control of diagnostic X-ray imaging equipment*, Medical physics monograph No. 20, American Institute of Physics, Inc., 1994, 1129 pages.
- Stranzinger, E., Treumann, T.C., Dreier, D., Allgayer, B., *Teleradiology for emergency cranial computed tomography.*, ROFO-Fortschritte auf dem Gebiet der Rontgenstrahlen und der Bildgebenden V, 175(5):646-53, 2003, (Abstract).
- Tachakra, S., *Level of diagnostic confidence, accuracy, and reasons for mistakes in teleradiologic for minor injuries*. Telemedicine Journal and e-Health, 8(1):111-121, 2002.
- Vuletic, S., *Teleradiology in Croatia.*, Journal of Telemedicine & Telecare, 7 Suppl 1:73-5, 2001.
- Wadley, B. D., Hayward, Ul, Trambert, M., Kywi, A., Hartzman, S., *Are referring doctors ready for enterprise and community wide immediate image and report access?*, Journal of Digital Imaging, 15 Suppl 1:140-3, 2002.



Des gens de soins et d'excellence

Centre hospitalier universitaire de Sherbrooke

Unité d'évaluation des technologies et des modes d'intervention en santé

Hôpital Fleurimont, 3001, 12^e Avenue Nord, Fleurimont
(Québec) J1H 5N4

Hôtel-Dieu, 580, rue Bowen Sud, Sherbrooke
(Québec) J1G 2E8

Tél. : (819) 346-1110, poste 13802