

**Évaluation clinique expérimentale de
la validité d'instruments utilisés
pour le diagnostic de lombalgies :
spinoscope,
dynamométrie triaxiale (B-200)
et thermographie,
en regard de l'examen clinique**

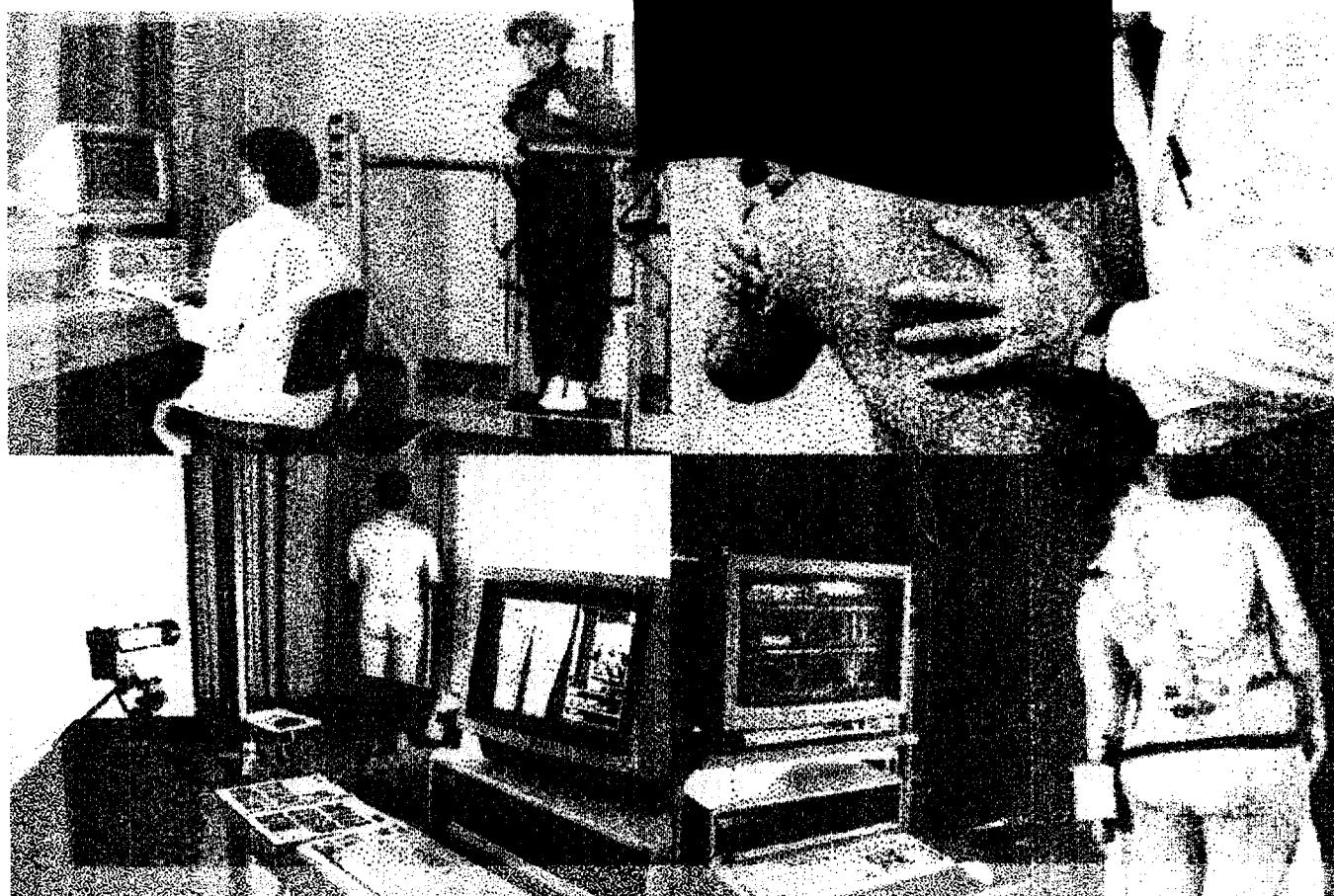
**ÉTUDES ET
RECHERCHES**

Richard Leclaire
John M. Esdaile
James Hanley
Michel Rossignol
Madeleine Bourdouxhe
Jean-Claude Jéquier
Harim Hajian-Tilaki

AOût 1995

R-105

RAPPORT



IRSST

Institut de recherche
en santé et en sécurité
du travail du Québec

La recherche, pour mieux comprendre

L'Institut de recherche en santé et en sécurité du travail du Québec (IRSST) est un organisme de recherche scientifique voué à l'identification et à l'élimination à la source des dangers professionnels, et à la réadaptation des travailleurs qui en sont victimes. Financé par la CSST, l'Institut réalise et finance, par subvention ou contrats, des recherches qui visent à réduire les coûts humains et financiers occasionnés par les accidents de travail et les maladies professionnelles.

Pour tout connaître de l'actualité de la recherche menée ou financée par l'IRSST, abonnez-vous gratuitement au magazine *Prévention au travail*, publié conjointement par la CSST et l'Institut.

Les résultats des travaux de l'Institut sont présentés dans une série de publications, disponibles sur demande à la Direction des communications.

Il est possible de se procurer le catalogue des publications de l'Institut et de s'abonner à *Prévention au travail* en écrivant à l'adresse au bas de cette page.

ATTENTION

Cette version numérique vous est offerte à titre d'information seulement. Bien que tout ait été mis en œuvre pour préserver la qualité des documents lors du transfert numérique, il se peut que certains caractères aient été omis, altérés ou effacés. Les données contenues dans les tableaux et graphiques doivent être vérifiées à l'aide de la version papier avant utilisation.

Dépôt légal
Bibliothèque nationale du Québec

IRSST - Direction des communications
505, boul. de Maisonneuve Ouest
Montréal (Québec)
H3A 3C2
Téléphone : (514) 288-1 551
Télécopieur: (514) 288-7636
Site internet : www.irsst.qc.ca
© Institut de recherche en santé
et en sécurité du travail du Québec,

**Évaluation clinique expérimentale de
la validité d'instruments utilisés
pour le diagnostic de lombalgies :
spinoscope,
dynamométrie triaxiale (B-200)
et thermographie,
en regard de l'examen clinique**

**ÉTUDES ET
RECHERCHES**

RAPPORT

**Richard Leclaire, Hôpital Notre-Dame
Département de médecine, Université de Montréal**

**John M. Esdaile, Hôpital général de Montréal
Division d'épidémiologie clinique, Université McGill**

**James Hanley, Département d'épidémiologie et de biostatistiques,
Université McGill**

**Michel Rossignol, Hôpital général juif Sir Mortimer B. Davis,
Centre d'épidémiologie clinique et de recherche en santé publique,
Université McGill**

Madeleine Bourdouxhe, IRSST

**Jean-Claude Jéquier, employé de l'IRSST au début de l'étude;
actuellement à l'Hôpital de Radio-pédiatrie de Genève**

**Karim Hajian-Tilaki
Département d'épidémiologie et de biostatistiques, Université McGill**

ÉQUIPE DE RECHERCHE

Responsable du projet

- Richard Leclaire, M.D., F.R.C.P.C., physiatrie, Hôpital Notre-Dame de Montréal et Institut de physiatrie du Québec; professeur agrégé, département de médecine, Université de Montréal.

Principaux investigateurs

- Jean-Claude Jéquier, M.D., M.Sc., sciences cliniques, IRSST*.
- John M. Esdaile, M.D., F.R.C.P.C., rhumatologie, M.Ph., Hôpital général de Montréal, division d'épidémiologie clinique; professeur titulaire, Université McGill.
- Michel Rossignol, M.D., F.R.C.P.C., santé communautaire, M.Sc., épidémiologie, Hôpital général juif Sir Mortimer B. Davis, Centre d'épidémiologie clinique et de recherche en santé publique; professeur adjoint, Université McGill.
- James A. Hanley, Ph.D., biostatistiques, professeur titulaire, département d'épidémiologie et de biostatistiques, Université McGill.
- Madeleine Bourdouxhe, M.Sc., IRSST.
- Karim Hajian-Tilaki, M.Sc., biostatistiques, département d'épidémiologie et de biostatistiques, Université McGill.
- * Employé de l'IRSST au début de l'étude; travaille actuellement à l'Hôpital de Radio-pédiatrie de Genève.

Principaux participants

Médecins sélectionneurs de référence

- Lesly Germain, M.D., F.R.C.S.C., orthopédie, Hôpital Notre-Dame de Montréal et Centre hospitalier Saint-Michel.
- Roland Proulx, M.D., F.R.C.P.C., physiatrie, Hôpital Notre-Dame de Montréal; professeur adjoint de clinique, Université de Montréal.

Biostatisticien

- Karim Hajian-Tilaki, M.Sc.

Entraîneure à la motivation

- Caroline Vachon, T.R.P.

Assistante de recherche

- Francine Bujold, T.R.P.

Évaluateurs cliniques

- Luc Fortin, M.D., F.R.C.P.C., physiatrie, Hôpital Notre-Dame de Montréal.
- Guy Bouvier, M.D., F.R.C.S.C., neurochirurgie, Hôpital Notre-Dame de Montréal et Hôpital Sainte-Justine.

Utilisateurs des technologies

- Marcel Morand, M.D., (B-200), F.R.C.P.C., physiatrie, Hôpital Notre-Dame de Montréal et Cité de la santé de Laval.
- Jean Fleury, M.D., (B-200), F.R.C.P.C., physiatrie, Institut de réadaptation de Montréal.
- Patrick Loisel, M.D., (spinoscopie), F.R.C.S.C., orthopédie, Centre hospitalier universitaire de Sherbrooke; professeur adjoint en médecine, Université de Sherbrooke.
- Jean-Paul Ouellette, M.D., (spinoscopie), C.C.F.P., Hôpital général d'Ottawa; chargé d'enseignement, Université d'Ottawa.
- Yves Bergeron, M.D., (thermographie), F.R.C.P.C., physiatrie, Hôpital Notre-Dame de Montréal et Hôpital du Sacré-Coeur.
- Roger Vadeboncoeur, M.D., (thermographie), F.R.C.P.C., physiatrie, Hôpital Saint-Luc.

Spinex : les données présentées à l'Annexe I ont été recueillies pour la plupart par M. Steeve Asselin, M.Sc., alors que M. Victor Lanzo, B.Eng., B.Sc., les a complétées; l'interprétation de ces données a été effectuée par M. Serge Gracovetsky, Ph.D., et le Dr Nicolas Newman, M.D., F.R.C.S.C., orthopédie, Hôtel-Dieu de Montréal, professeur adjoint de clinique, Université de Montréal.

TABLE DES MATIÈRES

	Page
Sommaire	5
1. Objectif	9
1.1 Objectif général	9
1.2 Objectifs spécifiques	9
2. Introduction	9
2.1 Pertinence	9
2.2 Importance de la lombalgie	10
2.3 Thermographie	11
2.4 Dynamométrie triaxiale (B-200)	12
2.5 Spinoscopie	12
3. Méthode	13
3.1 Devis expérimental	13
3.2 Sélection de l'étalon de référence	13
3.3 Mesures de la validité	14
3.4 Phase 1 - Valeurs normales	16
3.4.1 Objectif	16
3.4.2 Population	16
3.4.3 Critères d'inclusion	17
3.4.4 Critères d'exclusion	17
3.4.5 Recueil des données et sélection des sujets normaux	18
3.4.6 Examen des sujets normaux par les technologies	18
3.5 Phase 2 - Évaluation des technologies chez des travailleurs avec lombalgie	19
3.5.1 Objectifs	19
3.5.2 Population	20
3.5.3 Critères d'inclusion et d'exclusion	20
3.5.4 Intervention et examen des sujets	22
3.5.5 Choix méthodologiques et mesures de contrôle	24
4. Résultats	25
4.1 Rappel méthodologique	25
4.2 Résultats	31
5. Discussion	36
6. Conclusion	42
7. Retombées éventuelles	43
Bibliographie	45

TABLE DES MATIÈRES (suite)

	Page
APPENDICE 1 Figures et tableaux	A1
Liste des figures	A3
Liste des tableaux	A11
ANNEXE I Résultats de la spinoscopie obtenus de Spinex	A31
Liste des tableaux	A37
ANNEXE II Références additionnelles	A101
B-200	A103
Thermographie	A107
ANNEXE III Sujets refusés	A117
ANNEXE IV Cahier des formulaires utilisés (en dépôt à l'IRSST)	A121
Liste des formulaires utilisés	A123

SOMMAIRE

Objectif général

Évaluer la validité de trois technologies, soit la spinoscopie, la dynamométrie triaxiale (B-200) et la thermographie, pour diagnostiquer une lombalgie récente chez des travailleurs.

Objectifs spécifiques

Déterminer les caractéristiques diagnostiques de chacune des trois technologies avec et sans les données cliniques, comparer les technologies entre elles, évaluer l'influence de la motivation des sujets sur la précision diagnostique de chacune et déterminer la reproductibilité des résultats selon chacun des utilisateurs des technologies.

Méthode

Dans un premier temps, une étude portant sur les valeurs normales de sujets sains (phase I) a été réalisée afin de permettre à chaque technologie de se calibrer, aux participants de se familiariser avec la méthode et à l'équipe de recherche de s'assurer du déroulement optimal de la phase II.

La phase II a ensuite été réalisée avec l'objectif d'évaluer la capacité de ces technologies à poser un diagnostic de lombalgie.

Deux médecins spécialistes expérimentés (médecins sélectionneurs) ont été appelés à constituer un étalon de référence (*gold standard*) selon des critères d'inclusion et d'exclusion cliniques permettant de classer les sujets en deux groupes, soit les sujets malades et les sujets en santé.

Les deux médecins sélectionneurs ont constitué un ensemble de 41 malades présentant un premier épisode de lombalgie de type entorse lombaire d'une durée de plus de quatre semaines mais ne dépassant pas 10 semaines, sans évidence clinique de discopathie, de radiculopathie ou d'anomalie radiologique.

Un ensemble de 46 sujets sains, dont ni les antécédents, ni l'examen clinique et radiologique ne révélaient d'affection lombaire ou autre antécédent ayant requis une évaluation médicale, a été utilisé comme groupe contrôle.

Par attribution aléatoire, la moitié des malades présentant une entorse lombaire ont été formés pour dissimuler leur lombalgie et la moitié des sujets sains, pour simuler la présence d'une affection lombaire. Quatre groupes ont ainsi été constitués, soit 1) les sujets sains et honnêtes, 2) les sujets sains et simulateurs, 3) les sujets malades et honnêtes et 4) les sujets malades et dissimulateurs.

Les quatre groupes de sujets ont été soumis à une évaluation clinique individualisée à simple insu faite par deux médecins spécialistes (évaluateurs cliniques) différents des médecins sélectionneurs (servant de référence). Des techniciens certifiés, reconnus compétents et choisis par les utilisateurs des technologies, ont procédé aux examens suivants : thermographie, spinoscopie, dynamométrie triaxiale. Toutes les évaluations cliniques et les examens furent réalisés dans un ordre prédéterminé, sur une période maximale de 24 heures. Les données des technologies, portant un numéro de code servant à l'identification anonyme des sujets, ont été emmagasinées sur support informatique. Les techniciens, un pour la thermographie, un pour la dynamométrie triaxiale et deux pour la spinoscopie, ont administré les tests à simple insu et ont suivi un protocole d'enregistrement des données préétabli.

Après la répartition aléatoire de l'ordre des enregistrements et le codage des sujets, deux utilisateurs pour chacune des technologies ont interprété individuellement les données, sans avoir accès à leurs lectures précédentes. Un délai de quatre semaines a été respecté entre cette première lecture et la deuxième, lors de laquelle les utilisateurs des technologies avaient accès aux informations cliniques; les données de cette deuxième lecture ont été présentées dans un ordre aléatoire différent de celui de la première lecture. Cette deuxième lecture a été immédiatement suivie du consensus des deux utilisateurs de chaque technologie sur les données des technologies combinées aux informations cliniques.

L'analyse des résultats tient compte de trois mesures de la validité, soit la méthode COR et celles de l'exactitude simple et double. Les critères objectifs de performance retenus *a priori* pour établir l'utilité clinique des technologies étaient une sensibilité de 85 % et une spécificité de 90 %, attendu qu'il s'agit d'une étude expérimentale donnant le maximum de chance aux technologies de performer.

La première mesure de validité, soit la méthode de la caractéristique opérationnelle des résultats (méthode COR), correspond à la situation la plus facile pour les utilisateurs des technologies et les évaluateurs cliniques.

La deuxième, la méthode de l'exactitude simple, visait à évaluer l'aptitude des technologies et de l'examen clinique à distinguer les sujets sains et malades basée sur l'état de santé seul, alors que l'indécision était considérée comme une mauvaise décision.

La troisième méthode, soit celle de l'exactitude double, avait pour but de distinguer les sujets selon l'exactitude double, soit déterminer leur catégorie d'appartenance réelle basée sur l'état de santé et le rôle dans l'un des quatre groupes : sain et honnête; sain et simulateur; malade et honnête; malade et dissimulateur.

Les pourcentages que l'on obtient pour chaque sous-groupe ou chaque technologie peuvent se comparer directement entre eux, puisqu'ils proviennent du même échantillon expérimental. Par contre, il ne faudrait pas penser retrouver de tels résultats dans une population générale, car les taux de sujets honnêtes ou

tricheurs n'y seraient vraisemblablement pas les mêmes, bien que ce pourcentage ne soit pas clairement établi.

Résultats

Pour l'ensemble des sujets (honnêtes, simulateurs et dissimulateurs), le B-200 a obtenu un indice COR de 0,80, une valeur globale de 70,7 % à l'exactitude simple et de 39,1 % à l'exactitude double; la spinoscopie, un indice COR de 0,75, une valeur globale de 61,4 % à l'exactitude simple et de 38,5 % à l'exactitude double; et l'examen clinique, un indice COR de 0,74, une valeur globale de 63,8 % à l'exactitude simple et de 59,8 % à l'exactitude double. La thermographie est non performante, avec un indice COR de 0,33, une valeur globale de 37,9 % à l'exactitude simple et de 38,0 % à l'exactitude double. Le B-200 avec l'addition des examens cliniques démontre un indice COR de 0,85, une valeur globale de 73,6 % selon l'exactitude simple et de 67,3 % selon l'exactitude double, soit la meilleure performance réalisée dans cette étude.

Pour le sous-groupe de sujets honnêtes, dans le même ordre de présentation, les résultats sont les suivants :

- examen clinique : indice COR de 0,99, exactitude simple de 93,2 % et double de 92,1 %;
- B-200 : indice COR de 0,91, exactitude simple de 80,7 % et double de 68,1 %;
- spinoscopie : indice COR de 0,78, exactitude simple de 62,5 % et double de 59,1 %;
- thermographie : indice COR de 0,30.

Pour le sous-groupe de sujets simulateurs et dissimulateurs (malhonnêtes), les résultats sont :

- B-200 : indice de 0,72, exactitude simple de 60,5 % et double de 9,3 %;
- spinoscopie : indice de 0,73, exactitude simple de 59,3 % et double de 17,5 %;
- examen clinique : indice de 0,30, exactitude simple de 32,6 % et double de 26,8 %;
- thermographie : indice de 0,38.

Conclusion

La conclusion est claire. L'emploi des technologies (thermographie, dynamométrie triaxiale et spinoscopie) à des fins diagnostiques n'est d'aucune utilité lorsqu'il s'agit de sujets sains ou lombalgiques (entorse lombaire) honnêtes. Dans l'état actuel des connaissances, l'évaluation clinique s'avère performante et suffisante lorsque le sujet est honnête et réellement malade, ce qui représente probablement la majorité des lombalgiques accidentés du travail.

Par contre, si le sujet est simulateur ou dissimulateur, ni les technologies ni l'examen clinique ne peuvent actuellement distinguer de façon satisfaisante

l'état de santé de celui de maladie. En effet, malgré la performance supérieure de la dynamométrie triaxiale et de la spinoscopie à celle de l'examen clinique chez les sujets simulateurs ou dissimulateurs, ces technologies sont toutefois inaptes à distinguer de façon fiable ces sujets. Aucune méthode n'arrive à détecter adéquatement si un patient est honnête ou pas, qu'il simule ou dissimule une lombalgie.

La généralisation des résultats de la présente recherche à caractère expérimental comporte certaines limites étant donné la nature de l'échantillon utilisé. Toute généralisation doit donc tenir compte de ce facteur.

1. OBJECTIF

1.1 Objectif général

Évaluer la validité interne de trois nouvelles techniques diagnostiques, soit la spinoscopie, la dynamométrie triaxiale (B-200) et la thermographie, ainsi que l'examen clinique seul, pour détecter une lombalgie récente chez des travailleurs.

1.2 Objectifs spécifiques

Phase 1 : Déterminer les « valeurs normales » de chaque technologie chez 40 sujets en santé.

Phase 2 :

- a) Déterminer les caractéristiques diagnostiques de chaque technologie et de l'examen clinique seul, et les comparer au moyen de la méthode COR (courbe de la caractéristique opérationnelle du résultat).
- b) Évaluer les caractéristiques diagnostiques des technologies lorsqu'on fournit à leurs utilisateurs les informations provenant des cliniciens et des technologies.
- c) Évaluer l'influence de la "motivation" sur les caractéristiques diagnostiques par la comparaison des sujets simulant ou dissimulant une lombalgie avec des sujets vraiment sains ou malades.
- d) Déterminer la reproductibilité des lectures par le même utilisateur et entre les deux utilisateurs d'une même technologie.

2. INTRODUCTION

2.1 Pertinence

La CSST est responsable de l'indemnisation des soins de santé et de perte de revenu des travailleurs atteints d'une lombalgie liée au travail et aux maladies professionnelles aux termes de la *Loi sur les accidents du travail et les maladies professionnelles*. Au Québec, les coûts des examens utilisant les trois technologies au début de cette étude (1989) étaient respectivement de 300 \$ pour la thermographie (Diagnos inc.), 625 \$ pour le B-200 (Diagnos inc.) et de 150 \$ pour le spinoscope (Spinex). Aux États-Unis, les coûts pouvaient atteindre 700 \$ US. Le prix d'achat de ces appareils était respectivement de 101 538 \$ pour le thermographe (modèle Agema Thermovision 890), de 70 100 \$ pour le B-200 (Cidraf Equipment Inc.) et approximativement de 100 000 \$ pour un spinoscope.

Il convenait donc d'évaluer scientifiquement les caractéristiques diagnostiques de ces trois technologies par rapport à l'examen clinique avant que leur usage ne se répande, pour d'abord mieux évaluer leur performance eu égard au diagnostic des lombalgies afin de cerner leur apport clinique. Cette étude s'avérait d'autant plus nécessaire si l'on tient compte des frais prévisibles pour les malades et les organismes d'indemnisation et, surtout, des coûts humains pouvant découler de l'emploi de technologies non validées qui, utilisées sans discernement, pouvaient devenir préjudiciables au malade. Pour répondre à ce défi qu'est l'évaluation des technologies proposées par Fineberg¹ et d'autres², l'IRSST a commandité la présente recherche.

2.2 Importance de la lombalgie

En 1993, 39 733 nouveaux dossiers d'affections vertébrales ont été indemnisés par la CSST. Rapporté sur l'ensemble de la population couverte par le régime de SST* québécois, cela correspondrait à une proportion de 1,8 % des travailleurs québécois atteints de telles affections.

Ce volume de dossiers représente plus de 29 % des nouvelles déclarations rapportées en 1993 et indemnisées par la CSST. Elles ont occasionné 35,1 % des montants versés pour l'ensemble des lésions professionnelles, soit 80 millions de dollars³.

Le Groupe de travail québécois sur les aspects cliniques des affections vertébrales chez les travailleurs accidentés, commandité par l'IRSST, soulignait en 1987 l'absence d'études de bonne qualité relatives au diagnostic et au traitement des affections lombaires. Les conclusions de ce groupe d'étude ont permis de reconnaître les discordances inhérentes à l'évaluation des signes cliniques, de la douleur rapportée et de la perte de fonction des malades⁴. Deyo⁵, dans une revue de la documentation scientifique, a fait la même constatation. Cette situation a amené l'IRSST à déclarer prioritaire la recherche sur les affections lombaires. L'étude actuelle se situe dans ce contexte.

Quant aux procédures diagnostiques, le groupe de travail⁴ reconnaît l'importance des antécédents et de l'examen physique pour tous les malades présentant une lombalgie, de l'examen radiologique conventionnel et de la vitesse de sédimentation chez les lombalgiques symptomatiques pour plus de sept semaines. Il reconnaît aussi l'évidence scientifique justifiant l'utilisation de la tomodensitométrie et de la myélographie dans la sténose spinale de même que de l'électrodiagnostic chez les malades symptomatiques six mois après une chirurgie discale. Aucune preuve scientifique ne démontrait la valeur d'autres procédures diagnostiques. Par la suite, aucune technologie, incluant la thermographie⁶ ou la dynamométrie lombaire^{7,8}, n'a démontré scientifiquement sa valeur diagnostique (Annexe II - Références additionnelles).

De nouvelles techniques ou technologies s'avèrent sûrement nécessaires pour identifier les affections vertébrales lombaires. Elles pourraient inclure de nouvelles méthodes d'évaluation clinique et des technologies non invasives ou la combinaison de méthodes cliniques et de technologies.

* santé et de sécurité du travail

2.3 Thermographie

En 1956, à l'Université McGill, Lawson⁹ utilisa pour la première fois la thermographie afin d'étudier la température cutanée dans les cancers du sein. Depuis lors, cette technologie s'est raffinée par le développement de la thermographie à cristaux liquides^{10,11} et à infrarouge.

La thermographie est une procédure diagnostique permettant de capter et de mesurer l'énergie infrarouge émise par la peau. Cette énergie calorique est enregistrée par une caméra à infrarouge reproduisant sur écran cathodique une image thermique en couleurs, codée selon un gradient thermique. La procédure est non invasive puisqu'elle n'utilise pas de radiation ionisante. Son utilisation clinique, pour fins diagnostiques, repose sur la prémisse que l'émission de chaleur est symétrique sur le corps d'un sujet normal.

Dans certaines pathologies, l'irritation des fibres nerveuses, par l'intermédiaire de l'action réflexe des fibres du système nerveux autonome, induirait une asymétrie calorique pouvant confirmer la présence d'une affection induisant cette dysfonction.

La thermographie pourrait ainsi, dans certaines pathologies, faciliter la détermination de l'atteinte d'un nerf spinal ou d'une de ses branches, soit antérieure comme dans les pathologies discales, soit postérieure comme dans les pathologies vertébrales touchant les articulations postérieures, les ligaments, les muscles ou encore la peau.

Selon la documentation scientifique actuelle^{10 à 40}, cette technique serait potentiellement apte à diagnostiquer une atteinte radiculaire par hernie discale lombaire. De l'ensemble des études, il ressort que la sensibilité de la technique serait grande, soit environ 90 %, alors que sa spécificité serait beaucoup moindre, soit 60 à 80 %. Soulignons que la plupart de ces études sont descriptives et rétrospectives, et qu'elles comportent de nombreux biais; la catégorisation des sujets par l'étalon de référence n'est appliquée qu'après la réalisation de la thermographie⁴¹ et, mis à part quelques rares exceptions⁴², les lecteurs ne sont pas à simple insu. Wexler⁴³ a même suggéré que cet examen puisse servir à refuser une demande d'indemnisation : « may well provide a sound basis for the denial of a medical disability claim ». Dans les affections vertébrales lombaires non radiculaires, les résultats diagnostiques restent conflictuels et n'en sont qu'au stade de l'exploration clinique sans protocole comparatif. Le diagnostic de lombalgie reposerait sur l'observation d'une asymétrie thermique peut-être attribuable à des points de tension musculaire (*trigger points*).

L'équipement utilisé est le Thermovision 890, commercialisé par AGEMA Thermovision. Refroidi par azote liquide, cet appareil a le plus grand pouvoir de résolution et ses données graphiques peuvent être enregistrées sur vidéo, sur film négatif ou sur papier. Pour être valable, la procédure doit faire appel à des normes d'utilisation strictes, telles que recommandées par l'American Medical

Association. Ces normes ont été appliquées intégralement dans la présente recherche.

Cette technologie diagnostique est par ailleurs utilisée partout dans le monde, et l'on compte environ 2000 appareils en usage aux États-Unis.

2.4 Dynamométrie triaxiale (B-200)

Le B-200 est un appareil permettant d'enregistrer les forces musculaires dans trois axes avec des résistances variables.

L'appareil, produit par la maison Isotechnologies Inc., est sur le marché depuis 1987 et son modèle le plus ancien, le B-100, depuis 1984. Il a fait l'objet de plusieurs études cliniques et biomécaniques^{6,44 à 58}.

Selon certains chercheurs et l'entreprise Isotechnologies, l'appareil s'avère utile pour mesurer la capacité fonctionnelle du rachis lombaire, pour mesurer les progrès des malades présentant une affection lombaire et pour détecter une affection lombaire. Selon Spengler et coll.⁵⁹ et Fraser et coll.⁶⁰, il pourrait identifier les sujets simulateurs (*malingers*) par les courbes du mouvement et par la qualité de l'effort fourni par le sujet.

Des études de normalité ont été réalisées, et plus de 2000 de ces appareils sont en usage aux États-Unis.

Cependant, la revue de la documentation scientifique ne démontre pas que cet instrument diagnostique soit actuellement en mesure de catégoriser les sujets en sujets sains ou malades, ou encore de détecter les simulateurs (cf. Annexe II).

Newton et coll. concluent, en 1993, suivant leurs recherches⁶ et une revue approfondie de la littérature, qu'après une décennie d'études, il n'existe toujours pas d'évidence scientifique pour soutenir l'utilité des technologies isokinétiques ou isoinerties d'évaluation de la force de la musculature lombaire lors de l'évaluation préemploi, dans la pratique clinique du diagnostic des lombalgies et dans l'expertise médico-légale.

2.5 Spinoscopie

Cet appareil est le dernier-né des instruments diagnostiques des affections vertébrales lombaires. Produit par la maison Spinex Medical Corporation du Québec^{61,62}, cette technique non invasive permet l'analyse des mouvements du rachis lombaire, particulièrement en flexion-extension. Des électrodes sensibles à la lumière infrarouge, fixées sur les diverses vertèbres lombaires, sont suivies dans leur déplacement à l'aide de deux caméras à infrarouge qui permettent de reproduire et de décomposer tant la mobilité lombaire globale qu'intersegmentaire sur un écran moniteur. L'appareil enregistre aussi le mouvement des hanches, mesure l'angle lombo-sacré ainsi que le rapport entre le mouvement produit par le rachis lombaire et par les hanches. Un appareil électromyographique enregistre de façon concomitante l'activité relative des muscles multifidus et

des ischio-jambiers. Les résultats peuvent être reproduits sur papier.

La technique est nouvelle, et les études de normalité n'en sont qu'à leur début alors qu'on s'en sert déjà en clinique pour objectiver ou non des diagnostics de sujets normaux ou anormaux. L'appareil a été conçu pour servir de complément diagnostique permettant de détecter les sujets sains des pathologiques ou encore, des simulateurs ou dissimulateurs⁶³.

Il n'existait pas d'autres références à cet appareil dans la littérature médicale mis à part celles citées au début de ce rapport.

3. MÉTHODE

3.1 Devis expérimental

Le devis expérimental de cette recherche a été élaboré pour permettre l'évaluation de la validité interne des technologies pouvant être utilisées à des fins diagnostiques chez des sujets sains ou lombalgiques (entorse lombaire). La figure 1, présentée en appendice, représente les étapes de l'étude.

3.2 Sélection de l'étalon de référence

Tel que noté par le Groupe de travail québécois sur les aspects cliniques des affections vertébrales chez les travailleurs accidentés⁶⁴, en l'absence de radiculopathie clinique, il n'existe aucune procédure diagnostique, clinique ou instrumentale pour évaluer une lombalgie d'une durée inférieure à sept semaines. Quoique les constats chirurgicaux, la myélographie ou la tomodensitométrie puissent être utilisés comme étalons de référence pour déterminer la validité diagnostique, il aurait été contre l'éthique d'employer les deux premières techniques chez les sujets sains ou chez les malades présentant une lombalgie sans évidence de radiculopathie⁶⁴, (soit 90 % des malades, selon les données de l'Institut de physiatrie du Québec). De plus, aucun de ces examens ne détecterait les lésions responsables chez les malades présentant une lombalgie attribuable à un traumatisme des tissus mous, c.-à-d. ligaments vertébraux, capsules articulaires, musculature vertébrale, etc. Pour ces raisons, nous avons choisi comme étalon de référence une évaluation clinique complète et standardisée par deux médecins spécialistes, ce qui permettait de comparer la validité diagnostique des technologies à évaluer scientifiquement.

Les symptômes et les signes attribuables à l'entorse lombaire sont habituellement la douleur et des anomalies de la mobilité lombaires ou encore, des points douloureux dont la mesure clinique est habituellement subjective et peu reproductible d'un observateur à un autre. Certaines de ces anomalies peuvent être détectées par des technologies: la spinoscopie permet une évaluation dynamique lombaire qui pourrait être apte à déceler des anomalies des mouvements intersegmentaires du rachis lombaire, le B-200 est utilisé pour évaluer les forces de la musculature avec une charge progressive et la thermographie pourrait

détecter des asymétries thermiques probablement induites par un effet neuronal réflexe. De toute évidence, aucune de ces technologies ne mesure directement la douleur. Les anomalies retrouvées lors de l'évaluation instrumentale sont donc des signes indirects dont l'appréciation repose sur leur pertinence clinique chez des sujets symptomatiques ou non.

Plusieurs facteurs de confusion peuvent être responsables d'un diagnostic erroné dont, entre autres, la douleur générée. La douleur, ressentie d'une façon variable selon les sujets, produit secondairement des anomalies de mouvements, elles aussi variables selon les sujets et probablement selon les tissus lésés. La détection et l'appréciation cliniques de ces anomalies, qui sont souvent silencieuses, ne sont pas faciles pour les évaluateurs cliniques ni pour les utilisateurs de ces technologies. Compte tenu de la variabilité individuelle des anomalies et de la douleur ressentie ou non, il aurait été difficile de tenter d'évaluer la validité externe de ces instruments sur une population moins homogène que celle utilisée dans la présente recherche.

Pour cette étude expérimentale, deux groupes distincts ont été créés, soit les sujets certainement normaux (en santé) et les sujets certainement anormaux (malades). Tout sujet intermédiaire, c.-à-d. présentant un doute quant à sa normalité ou son anormalité, a été exclu. Cette méthode permet donc d'optimiser l'identification des sujets normaux et anormaux au moyen des technologies diagnostiques. De plus, ces sujets étaient parfaitement consentants et coopératifs; ils n'avaient aucun avantage à tricher. La capacité de chaque technologie à identifier la présence ou l'absence d'une lombalgie de type mécanique sans atteinte radiculaire cliniquement décelable a été analysée et comparée.

3.3 Mesures de la validité

Une caractéristique importante de la validité clinique d'un instrument diagnostique repose sur son pouvoir discriminant; c'est ce qui a fait l'objet de la recherche actuelle.

Ainsi, cette étude ne visait pas l'évaluation de la fiabilité (consistance interne) des instruments diagnostiques, ce qui est principalement la responsabilité des fabricants, mais plutôt l'évaluation de la validité interne de l'usage pratique que l'on en fait.

Le pouvoir discriminant des instruments diagnostiques a été mesuré en fonction d'une échelle ordinale à cinq niveaux s'étendant de : 1. Certainement pathologique; 2. Probablement pathologique; 3. Indécis; 4. Probablement sain; 5. Certainement sain.

L'exactitude diagnostique a été déterminée en comparant les interprétations sur cette échelle en cinq points avec le diagnostic véritable de l'étalon de référence établi par les médecins sélectionneurs.

Une échelle similaire mesurait le rôle joué par les sujets, soit simulateurs, soit dissimulateurs : 1. Certainement tricheur; 2. Probablement tricheur; 3. Indécis; 4. Probablement honnête; 5. Certainement honnête.

Trois mesures distinctes ont été construites pour évaluer la validité du pouvoir discriminant des instruments diagnostiques des lombalgies mécaniques communes (entorses lombaires). Ces mesures complémentaires assurent, par la convergence des interprétations, des conclusions plus certaines.

. Méthode COR (courbe de la caractéristique opérationnelle du résultat)

La méthode COR permet de représenter graphiquement la capacité diagnostique d'une procédure d'évaluation. Cette méthode est établie objectivement dans le sens qu'elle est indépendante des fréquences des événements étudiés et des critères de décision utilisés pour définir la présence ou l'absence de la maladie. Elle permet aussi la comparaison des performances diagnostiques d'instruments de mesure au moyen d'une échelle commune.

Chaque point de la courbe COR correspond à un couple représentant respectivement la proportion de faux positifs (FP) et de vrais positifs (VP). La sensibilité et la spécificité de l'instrument peuvent ainsi être déduites en chaque point. Une mesure globale est attribuée à chaque courbe en considérant la surface sous la courbe. Elle varie de 0 à 1. Une valeur entre 0 et 0,5 indique une discrimination négative, c.-à-d. inverse au résultat attendu, tandis qu'une valeur entre 0,5 et 1 montre une discrimination positive; une valeur de 0,5 représente une discrimination nulle, c.-à-d. que les diagnostics sont attribuables au hasard. Swets⁶⁵ donne des renseignements supplémentaires sur les courbes opérationnelles relatives et leurs applications pour les diagnostics radiologiques, les pronostics météo, les sources d'information, les polygraphes détecteurs de mensonge et les tests d'aptitudes. Par des méthodes statistiques appropriées, les mesures globales des différentes procédures d'évaluation seront comparées^{66 à 69}.

Sur le plan expérimental, comme chaque sujet est évalué sur une échelle ordinale à cinq niveaux, la proportion de faux positifs et de vrais positifs est calculée séparément en cumulant les valeurs croissantes de l'échelle. Quatre couples distincts de proportion de faux positifs et de vrais négatifs sont obtenus; le dernier couple correspondant au cumul des cinq valeurs de l'échelle offre une proportion de faux positifs et de vrais positifs égale à 1. Un programme informatique, développé par Metz⁶⁶ et utilisé par Hanley^{67,68}, permet de tracer une courbe à travers ces points et d'obtenir la COR expérimentale de même que la surface sous cette courbe expérimentale. Par des méthodes statistiques paramétriques (Metz⁶⁶) ou non paramétriques (DeLong⁶⁹), les surfaces sous les courbes sont comparées.

Cette analyse a été réalisée par James A. Hanley, au département d'épidémiologie et de biostatistiques de l'Université McGill.

. Méthode d'évaluation par exactitude simple

Cette méthode d'analyse évalue l'aptitude des utilisateurs des technologies et des évaluateurs cliniques à distinguer les sujets sains et malades indépendamment de leur état de simulateur ou de dissimulateur. Toute indécision constitue une réponse non conforme à l'état du sujet, donc une erreur.

. Méthode d'évaluation par exactitude double

Cette méthode évalue l'aptitude des instruments diagnostiques et de l'évaluation clinique à distinguer les sujets selon les quatre groupes possibles : les sujets sains honnêtes, les sujets sains malhonnêtes (simulateurs), les sujets malades honnêtes et les sujets malades malhonnêtes (dissimulateurs). Les réponses des évaluateurs doivent donc être exactes selon les deux échelles (échelle de santé et échelle d'honnêteté). Toute indécision constitue une erreur.

L'analyse des résultats doit donc tenir compte des trois mesures de la validité, soit la méthode COR et celles de l'exactitude simple et double. L'approche selon les mesures d'exactitude est exigeante puisque l'indécision constitue une mauvaise réponse. La méthode de l'exactitude double amplifie l'exigence puisque l'indécision se situe à deux niveaux de critères, soit la condition de lombalgique ou non et l'état de simulateur/dissimulateur. Les critères objectifs de performance retenus *a priori* pour établir l'utilité clinique des technologies étaient une sensibilité de 85 % et une spécificité de 90 %, attendu qu'il s'agit d'une étude expérimentale donnant le maximum de chance aux technologies de performer.

3.4 Phase 1 - Valeurs normales

3.4.1 Objectif

La phase I avait pour objectif de permettre la calibration de chaque instrument et de permettre à l'équipe d'apporter les correctifs méthodologiques et techniques pour réaliser la phase II de façon optimale. La première étape était d'autant plus nécessaire que peu d'études ont été publiées sur les valeurs de normalité de ces technologies.

3.4.2 Population

Quarante sujets ont été recrutés à l'Institut de psychiatrie du Québec, un centre de soins externes affilié au Centre hospitalier universitaire de l'Hôpital Notre-Dame (Université de Montréal). Il s'agissait de personnes admises pour des affections mineures non vertébrales et non apparentes, ou de volontaires. Pour évaluer l'influence de l'âge et du sexe, les sujets ont été répartis en un nombre égal d'hommes et de femmes, divisé en quatre strates d'âge : 18 à 30, 31 à 40, 41 à 50 et 51 à 65 ans. Chaque strate comportait une population identique de sujets en fonction de leur classification : travail lourd, modéré, léger et sédentaire.

3.4.3 Critères d'inclusion

Les sujets devaient correspondre aux caractéristiques suivantes :

1. être un travailleur;
2. âgé de 18 à 65 ans;
3. francophone;
4. absence de douleur vertébrale⁷⁰;
5. autoquestionnaire d'Oswestry⁷¹ = 0;
6. autoquestionnaire de Roland-Morris⁷² = 0;
7. consensus des deux médecins sélectionneurs sur la vraisemblance d'une absence d'affection lombaire incluant :
 - a) courbure vertébrale normale;
 - b) mobilité lombaire active et passive normale selon l'âge du sujet;
 - c) rythme lombo-pelvien normal;
 - d) indice de Schober en flexion modifié selon Moll et Wright⁷³ ($\geq 21/15$ cm) (R);
 - e) indice de Schober en extension ($\geq 19/15$ cm);
 - f) absence de douleur à la palpation des ligaments et de la musculature lombaire;
 - g) épreuve jambe tendue bilatéralement négative;
 - h) épreuve de Lasègue bilatéralement négative;
 - i) épreuve de Ely bilatéralement négative;
 - j) longueur des membres inférieurs normale (< 1 cm de différence);
 - k) réflexes rotuliens et achilléens normaux;
 - l) sensibilité des membres inférieurs normale;
 - m) force musculaire segmentaire des membres inférieurs normale;
 - n) démarche normale;
 - o) examen des hanches, des genoux et des chevilles normal;
 - p) pulsations des artères périphériques des membres inférieurs normales;
 - q) examen abdominal normal;
8. examen radiologique (A.P. et latéral) normal tenant compte de l'âge du sujet (examen fait uniquement chez les sujets préalablement admissibles en fonction des critères d'inclusion et d'exclusion);
9. disponibilité et consentement éclairé.

3.4.4 Critères d'exclusion

La présence d'un des éléments suivants excluait le sujet :

1. chirurgie vertébrale;
2. chirurgie des membres inférieurs ou autres pouvant interférer avec une évaluation clinique ou technologique (ex. : chirurgie articulaire, abdominale ou vasculaire des membres inférieurs);
3. évaluation médicale antérieure pour une affection lombaire;
4. antécédents de lombalgie durant la dernière année;
5. antécédents de lombalgie, incluant sciatalgie, d'une durée supérieure à sept jours;
6. antécédents de perte de journée(s) de travail pour lombalgie ou sciatalgie;

7. toute affection médicale pouvant interférer avec les procédures de l'étude (ex. : maladie cardiovasculaire, cancer, maladie pulmonaire chronique, etc.);
8. antécédents d'arthropathie chronique des membres inférieurs;
9. grossesse.

3.4.5 Recueil des données et sélection des sujets normaux

Une assistante de recherche a évalué les candidats potentiels à la lumière des premiers critères d'inclusion. Ils ont tous par la suite été examinés indépendamment par un physiatre et par un orthopédiste possédant chacun une grande expérience (plus de 10 ans) dans l'évaluation des affections lombaires. Chacun de ces spécialistes a procédé à une évaluation clinique globale et standardisée des candidats permettant de conclure à l'absence d'une affection lombaire. Ces évaluations ont été révisées par l'assistante de recherche. Pour être inclus, un sujet devait satisfaire pleinement aux exigences des évaluations effectuées séparément par les deux médecins sélectionneurs et faire consensus.

Les données suivantes ont été recueillies pour tous les sujets : autoquestionnaire de Santé Québec, questionnaire sur les habitudes de vie, antécédents professionnels et type de travail selon les exigences physiques, classées en quatre groupes : sédentaire, travail léger, modéré et lourd (modifié de 74).

3.4.6 Examens des sujets normaux par les technologies

Tous les sujets satisfaisant aux critères d'inclusion et d'exclusion ont passé une radiographie lombaire (cf. les critères d'inclusion) après leur consentement écrit éclairé. Un point de référence radio-opaque a été placé en L4 lors de la radiographie afin de permettre le repérage anatomique précis pour satisfaire aux exigences de la spinoscopie. Les sujets présentant une anomalie radiologique lombo-sacrée ou une arthrose lombaire évaluée plus importante que celle attendue en fonction de leur âge, selon l'avis des médecins sélectionneurs, ont été exclus.

Les sujets dès lors admissibles ont d'abord été soumis soit à la thermographie, soit à la spinoscopie. Il était reconnu par l'équipe de recherche, après réception des informations transmises par Spinex concernant la spinoscopie, que l'examen de dynamométrie triaxiale pouvait « potentiellement » nuire à l'examen optimal des deux premières technologies. Dès cette information connue, la dynamométrie a toujours été faite en dernier lieu.

Les examens au moyen des technologies ont été réalisés par des techniciens choisis par les utilisateurs des appareils. La technicienne de la thermographie était certifiée par l'Association américaine de thermographie. La technicienne du B-200 a été formée par un instructeur recommandé par le fabricant et selon ses normes. Le technicien de la spinoscopie était un employé de Spinex. Chacune des techniques était donc administrée par un technicien expérimenté.

Les données recueillies par chaque technicien ont été interprétées par deux utilisateurs de chacune des technologies (les mêmes qui ont participé à la phase 2) afin d'établir pour chaque technologie les valeurs observées chez des sujets évalués en santé. Toutes les données informatiques des technologies ont été enregistrées sur disquettes. Un document supplémentaire imprimé a été créé avec les données du B-200 et de la spinoscopie alors qu'un négatif couleur a été utilisé comme document de référence pour la thermographie.

Le but premier de la phase I était de constituer une banque de données dites "valeurs normales" tirées des technologies. Un support informatique a été prévu pour évaluer les résultats qui ont été attribués à chaque technologie. Une analyse statistique des données du B-200 a été réalisée et transmise à ses utilisateurs alors que les usagers de la spinoscopie ont procédé à leur propre analyse des résultats. Ceux-ci ont été publiés sans le consentement explicite du commanditaire, l'IRSST. La thermographie étant un examen qui repose sur l'évaluation visuelle, les utilisateurs ont mis au point une grille d'appréciation par consensus issue des valeurs normales et de leur expérience avec cette technologie.

3.5 Phase 2 - Évaluation des technologies chez des travailleurs avec lombalgie

3.5.1 Objectifs

Objectif général

Évaluer la validité interne de trois techniques diagnostiques, soit la spinoscopie, la dynamométrie triaxiale (B-200) et la thermographie, ainsi que l'examen clinique seul, pour détecter une lombalgie récente chez des travailleurs.

Objectifs spécifiques

- Comparer le pouvoir discriminant de chaque technologie et de l'examen clinique seul au moyen de la méthode COR (courbe de la caractéristique opérationnelle du résultat).
- Évaluer les caractéristiques diagnostiques des technologies lorsqu'on fournit aux utilisateurs les informations des évaluateurs cliniques en plus de l'évaluation instrumentale.
- Évaluer l'influence de la "motivation" sur les caractéristiques diagnostiques par la comparaison des sujets simulant ou dissimulant une lombalgie avec des sujets véritablement sains ou malades.
- Déterminer la reproductibilité des lectures par le même utilisateur et entre deux utilisateurs d'une même technologie.

3.5.2 Population

Comme pour la phase I, les participants ont été recrutés à l'Institut de physiothérapie du Québec.

Sujets malades

Du 10 mai 1990 au 13 mai 1993, 288 personnes souffrant d'une lombalgie et potentiellement admissibles ont été évaluées en fonction des critères d'inclusion et d'exclusion de l'étude (voir la section 3.5.3). Après un examen par le premier médecin sélectionneur, 236 d'entre elles ont été refusées et 11 autres l'ont été après l'examen du deuxième médecin sélectionneur, pour un total de 247 malades inadmissibles faute de consensus. Un ensemble de 41 sujets malades a ainsi été formé.

Sujets sains

L'assistante de recherche a effectué la présélection des sujets sains selon les critères d'inclusion et d'exclusion (voir la section 3.5.3). Leur appariement a été fait avec la même méthode que dans la phase I, soit selon le sexe, l'âge et la catégorie de travail (tableau 1). Ces sujets sains ont ensuite été évalués par les médecins sélectionneurs selon les mêmes critères. D'un total de 54, six ont été refusés par manque de consensus des médecins et deux en raison de leur inaptitude à exécuter les tests demandés (voir l'Annexe III pour les informations sur chaque refus). Un ensemble de 46 sujets sains a ainsi été formé.

Sujets admis à l'étude

C'est ainsi que 87 sujets ont été considérés admissibles selon les critères d'inclusion et d'exclusion, dont 41 souffrant d'une lombalgie et 46 en santé, sans lombalgie. Des 41 malades, 21 ont été choisis par randomisation pour jouer le rôle de sujets réellement malades et les 20 autres devaient dissimuler leur lombalgie. Des 46 sujets sains, 23 ont été choisis de la même façon pour jouer le rôle de sujets sains et les 23 autres devaient simuler une lombalgie.

3.5.3 Critères d'inclusion et d'exclusion

Normaux

Les critères d'inclusion et d'exclusion des sujets normaux étaient identiques à ceux décrits pour la phase I de l'étude des valeurs normales.

Lombalgiques

Les critères d'inclusion étaient les suivants :

1. travailleur;
2. âgé de 18 à 60 ans inclusivement;
3. francophone;
4. antécédents d'une lombalgie liée à un accident;

5. lombalgie présente⁷⁰ depuis moins de 10 semaines;
6. lombalgie quotidienne depuis le début;
7. autoquestionnaire d'Oswestry⁷¹ ≥ 1 ;
8. autoquestionnaire de Roland-Morris⁷² ≥ 1 ;
9. deux des quatre éléments suivants devaient ressortir de l'examen des médecins sélectionneurs :
 - a) posture lombaire anormale;
 - b) asymétrie des mouvements lombaires actifs;
 - c) asymétrie des mouvements lombaires passifs;
 - d) douleur à la palpation lombaire (ligament, vertèbre ou muscle) présumément attribuable à une affection lombaire;
10. accord des deux médecins sélectionneurs sur les points suivants :
 - a) lombalgie ne dépassant pas le genou;
 - b) égalité des membres inférieurs (moins de 1 cm);
 - c) réflexes rotuliens et achilléens normaux;
 - d) sensibilité des membres inférieurs normale;
 - e) force musculaire segmentaire des membres inférieurs normale;
 - f) pulsations artérielles des membres inférieurs normales;
 - g) examen abdominal normal;
 - h) examen articulaire des membres inférieurs normal;
11. disponibilité des sujets après consentement écrit éclairé;
12. radiographie lombo-sacrée récente (deux clichés) normale considérant l'âge des sujets.

Les critères d'exclusion étaient les suivants :

1. chirurgie vertébrale;
2. chirurgie des membres inférieurs ou autres pouvant interférer avec une évaluation clinique ou technologique (ex. : chirurgie articulaire, abdominale ou vasculaire des membres inférieurs);
3. évaluation médicale antérieure pour une affection lombaire;
4. antécédents de lombalgie durant la dernière année;
5. antécédents de lombalgie, incluant sciatalgie, d'une durée supérieure à sept jours;
6. antécédents de perte de journée(s) de travail pour lombalgie ou sciatalgie;
7. toute affection médicale pouvant interférer avec les procédures de l'étude (ex. : maladie cardiovasculaire, cancer, maladie pulmonaire chronique, etc.);
8. antécédents d'arthropathie chronique des membres inférieurs;
9. grossesse.

Les 41 sujets lombalgiques présentaient les caractéristiques suivantes :

	Médiane	(25 ^e percentile, 75 ^e percentile)
Durée de la lombalgie	3 semaines	(2,5)
Douleur ⁷⁰	46 %	(37,54)
Roland-Morris ⁷²	46 %	(35,62)
Oswestry ⁷¹	35 %	(25,44)

3.5.4 Intervention et examen des sujets

Entraînement en fonction de la "motivation" des sujets

Considérant que la "motivation" peut constituer un facteur en mesure d'influer sur la performance lors de l'évaluation, tel qu'il arrive parfois dans la pratique quotidienne, les sujets ont été divisés par répartition aléatoire selon un code préétabli, puis séparés en deux groupes. Vingt-trois sujets sains ont été entraînés à jouer le rôle de simulateur et 20 malades, celui de dissimulateur. Ils ont été formés par une assistante selon une méthode standardisée impliquant un entraînement maximal de cinq minutes. Celui-ci consistait à familiariser le sujet avec une histoire plausible, appropriée au rôle à jouer et au comportement à adopter au cours des examens médicaux et ceux des technologies. Chaque sujet a été informé de la nécessité de faire les examens en entier dans la mesure de sa capacité physique. Cette technique utilisant des simulateurs et des dissimulateurs a déjà été employée par d'autres chercheurs^{75 à 78}.

Examen des sujets

Examens cliniques

Les évaluateurs cliniques choisis (un neurochirurgien et un physiatre) n'étaient impliqués dans aucune autre étape de cette étude. Ces médecins possèdent une grande expérience dans l'évaluation des lombalgiques. Ils ont rempli le même formulaire (cf. Mesures principales, page 23) que tous les autres lecteurs des technologies. Un second formulaire standardisé (fiches d'informations 9A et 10A, Annexe IV) et détaillé permettait de recueillir toutes les données cliniques que l'on s'attend à trouver lorsqu'un médecin examine un patient lombalgique. Ces deux formulaires ont été remis aux utilisateurs des technologies pour les étapes de 2^e lecture (avec informations médicales) et de 3^e lecture (consensus avec informations médicales).

Examens par les technologies

Enregistrement des données

L'enregistrement des données des technologies a été réalisé par les mêmes

techniciens qui ont participé à la phase I, à l'exception des 16 derniers sujets, pour lesquels la spinoscopie a été effectuée par un autre technicien de la maison Spinex.

Les enregistrements informatiques ont été transférés sur des disquettes individuelles et chaque sujet a été codé au moyen d'un numéro de six chiffres de façon à préserver son anonymat auprès des techniciens et des lecteurs. Les enregistrements du B-200 et de la spinoscopie ont été conservés sur un document écrit et ceux de la thermographie, sur un film. On a conservé une copie de tous les documents. Le codage, permettant l'anonymat des enregistrements, a été fait par l'IRSST, qui en est le dépositaire. Le décodage a été réalisé par James A. Hanley, statisticien au département d'épidémiologie et de biostatistiques de l'Université McGill, à la fin de l'étude, soit au moment de l'analyse des résultats.

Tous les sujets consentants ont été soumis aux évaluations décrites dans la phase I.

Interprétation des données

Toutes les lectures ont été réalisées en conformité avec le protocole de recherche établi mis à part Spinex qui a interprété les données de spinoscopie selon une méthode différente, expliquée à l'Annexe I.

Ainsi, à l'exception de Spinex, les lectures ont été effectuées en cas par cas, à intervalle de quatre semaines, sans possibilité de consulter les lectures antérieures. L'ordre de présentation a été déterminé aléatoirement.

La première lecture a été faite par chaque utilisateur individuellement avec les seules données des technologies; la deuxième, avec ces mêmes données auxquelles s'ajoutaient les informations des évaluateurs cliniques, et toujours sur une base individuelle; la troisième consistait en un consensus des deux utilisateurs de chaque technologie en possession des informations des évaluateurs cliniques.

Cette méthode permettait d'évaluer l'amélioration potentielle des performances diagnostiques des lecteurs.

Mesures principales

Une échelle ordinale de cinq niveaux de certitude a été utilisée pour établir le degré de certitude du diagnostic, à savoir :

- 1) certainement pathologique;
- 2) probablement pathologique;
- 3) indécis;
- 4) probablement sain;
- 5) certainement sain.

Une échelle similaire a été utilisée pour évaluer la capacité des lecteurs à identifier les sujets simulateurs ou dissimulateurs, à savoir :

- 1) certainement tricheur;
- 2) probablement tricheur;
- 3) indécis;
- 4) probablement honnête;
- 5) certainement honnête.

Autres mesures

Des informations additionnelles pouvant être utiles à des fins de recherche éventuelle (recherche en développement), soit la gravité des anomalies, le site de l'affection, la latéralisation, etc., ont aussi été colligées sur une grille établie par les évaluateurs cliniques et les utilisateurs des technologies.

3.5.5 Choix méthodologiques et mesures de contrôle

Toutes les évaluations technologiques ont été réalisées à l'Institut de physiatrie du Québec, alors que les évaluations cliniques l'ont été soit à l'Institut, soit à l'Hôpital Notre-Dame. Elles ont été effectuées sur une période maximale de six heures, à l'exception de celles de quatre sujets. Pour deux d'entre eux, la spinoscopie a été faite 48 heures après les examens des médecins sélectionneurs, en raison de la non-disponibilité de Spinex à procéder et, dans le cas de deux autres sujets, respectivement à 48 heures et à 72 heures, à cause de défaillances d'équipements informatiques relatifs à la thermographie et à la spinoscopie. De nombreux efforts ont donc été déployés par tous les participants à cette recherche pour que les sujets présentent le même état à tous les évaluateurs (évaluateurs cliniques et utilisateurs des technologies), étant donné que leur condition pouvait évoluer.

Les malades choisis comme sujets présentaient des caractéristiques cliniques correspondant à des affections musculo-squelettiques vertébrales considérées comme légèrement invalidantes et difficiles à objectiver. Ce choix fondamental des chercheurs était dicté par le fait que, lorsque les évidences cliniques sont présentes, les cliniciens expérimentés peuvent aisément établir le diagnostic ou constater l'anormalité des malades. Les examens radiologiques de plus en plus sophistiqués (tomodensitométrie et résonance magnétique nucléaire) confirment habituellement le diagnostic à des coûts inférieurs à ceux des technologies fonctionnelles en voie de développement.

Le défi à relever consistait donc à évaluer les possibilités diagnostiques des technologies chez les lombalgiques sans discopathie manifeste. Ces malades représentent la vaste majorité des sujets à évaluer (90 %) puisque nous ne possédons pas actuellement d'instrument diagnostique ou de mesures pouvant répondre à ce besoin dans un domaine où l'on exige de plus en plus de démonstrations objectives de la maladie. Ainsi, un maximum d'efforts a été consenti au choix des sujets pour que ceux-ci soient représentatifs, avec le moins de biais possible, de la population à étudier. Certains des critères d'exclusion sont restrictifs et pourraient limiter la portée des résultats sur

une population moins homogène. Cependant, compte tenu du fait que nous avons étudié des technologies pour la plupart en voie de développement et pour lesquelles les limites et les indications restent à définir, il a été décidé de leur donner le maximum de chance de démontrer leur efficacité.

Les affirmations de certains fabricants de ces technologies quant aux possibilités diagnostiques discriminantes chez des sujets simulateurs ont constitué un motif additionnel qui nous a conduit à étudier l'aspect de la "motivation" des sujets. Même si le pourcentage de ces malades n'est pas bien établi et même s'il est probablement minime¹⁹, les organismes payeurs sont fort sensibles à cette question controversée⁴³.

Deux des technologies, soit la spinoscopie et la dynamométrie, sont les plus répandues dans les évaluations de lombalgies. C'est pourquoi nous les avons choisies pour cette recherche. Certaines études (Annexe II) suggèrent que la thermographie serait potentiellement apte à diagnostiquer une radiculopathie secondaire à une affection discale. Certains auteurs croient, en outre, que cette technologie serait apte à identifier, dans certains cas, des affections mécaniques moins graves (Annexe II) et, d'un point de vue méthodologique, qu'elle pourrait servir de point de référence négatif (aucun effet) quant à un pouvoir diagnostique; l'équipe de recherche l'a donc incluse dans son protocole.

4. RÉSULTATS

4.1 Rappel méthodologique

Un diagnostic résulte d'une série d'évaluations. Celles-ci intègrent de façon complexe plusieurs éléments qui concernent tant l'évaluateur que l'évalué. Il est relativement facile de mesurer un résultat global : il est juste ou faux. Par contre, il s'avère plus difficile de connaître la contribution relative des éléments qui proviennent de l'évaluateur ou de l'évalué parce que ces derniers sont intimement liés tout au cours de la procédure diagnostique. C'est l'objectif que cette étude a visé pour mesurer de façon quantitative la contribution des technologies seules ou en combinaison avec l'évaluation clinique traditionnelle.

La lecture des résultats doit se faire de la même façon, c.-à-d. en décomposant les résultats globaux selon les différentes sources de variation pouvant produire une erreur. Celles-ci peuvent provenir soit de la technique de l'évaluation (évaluation clinique, évaluation technique et en combinaison), soit de l'évalué, qui présente toujours une double contrainte selon son intégrité physique (en santé ou malade) et morale (honnête ou pas). Les sujets sains malhonnêtes sont des simulateurs et les sujets malades malhonnêtes, des dissimulateurs. Les sujets honnêtes sont soit vraiment sains, soit vraiment malades.

Il n'y a qu'une seule réponse exacte mais plusieurs réponses erronées possibles. Il est donc facile de distinguer l'importance relative du processus diagnostique

qui doit être largement supérieur au hasard pour être valable.

Pour rendre cet exercice encore plus discriminant, quatre autres niveaux dans le processus diagnostique ont été imposés aux évaluateurs cliniques et aux utilisateurs des technologies. Le premier niveau concerne la différence des évaluations lorsque celles-ci sont répétées indépendamment par deux évaluateurs cliniques ou utilisateurs de technologies (différences interévaluateurs). Le deuxième niveau mesure la stabilité de l'évaluation de chaque évaluateur clinique ou utilisateur des technologies entre une première et une deuxième lecture (différences intraévaluateurs). Le troisième mesure la capacité de juger de l'information fournie par l'autre méthode d'évaluation, à savoir la valeur accordée par les utilisateurs des technologies à l'évaluation clinique et réciproquement, la valeur accordée par les évaluateurs cliniques aux deux technologies (le B-200 et la thermographie). Ce troisième niveau est combiné au deuxième, car la seconde lecture incluait les évaluations cliniques. Enfin, le quatrième niveau concerne la capacité de chaque évaluateur clinique ou utilisateur de technologie à intégrer l'évaluation de son collègue pour obtenir un consensus.

Quelques-uns de ces niveaux diagnostiques n'ont pu être imposés à certaines des mesures en raison de l'impossibilité de répéter l'examen clinique du patient, qui change au cours du temps, ou de faire de la simulation ou de la dissimulation en thermographie. Ces mesures sont donc absentes.

Un tel découpage du processus diagnostique présente deux avantages majeurs.

D'une part, il permet de mesurer la valeur absolue de chaque méthode d'évaluation (clinique ou instrumentale) seule ou en combinaison avec d'autres, ainsi que la contribution de l'évaluateur clinique ou de l'utilisateur de technologie en tant qu'individu par rapport à la contribution de la technique d'évaluation, que celle-ci soit clinique ou instrumentale.

D'autre part, en multipliant les niveaux du processus diagnostique, la probabilité d'erreur augmente du simple au double à chaque niveau de subdivision. Inversement, la probabilité de trouver une bonne réponse « par hasard » diminue. Il n'y a en effet qu'une chance sur quatre de poser le bon diagnostic avec les contraintes liées à l'évalué (intégrités physique et morale). En ajoutant celles qui sont liées à l'évaluateur clinique ou à l'utilisateur de technologies (double lecteur, double lecture, information clinique ou technologique et consensus), la probabilité d'obtenir au hasard une réponse exacte diminue.

Il ne s'agit pas ici d'étudier la lombalgie avec différentes techniques diagnostiques, mais bien d'étudier différentes techniques diagnostiques avec des cas de lombalgie soigneusement triés. Les résultats sont donc présentés selon deux modes d'évaluation. Le premier mode est du type « vrai ou faux », le deuxième, du type « certainement vrai, probablement vrai, incertain, probablement faux, certainement faux ».

Le premier mode d'évaluation permet de voir où se trouve la ligne de démarcation

entre ce qui est « juste ou faux ». Il inclut les indécis dans les faux, car l'indécision ne sert à rien pour agir ou trancher une situation floue.

Le deuxième mode est plus raffiné, car il considère que chaque niveau d'évaluation contient une information équivalente. En d'autres termes, la certitude est aussi importante que la franche indécision. Cette approche permet de savoir si tous les niveaux d'évaluation sont en continuité ou si le degré de certitude est biaisé vers le faux ou vers le vrai. Cette information s'avère importante pour juger si les évaluateurs cliniques ou les utilisateurs de technologies ont une propension vers un groupe plutôt que vers un autre.

Le premier mode permet d'obtenir un pourcentage de cas diagnostiqués « juste », selon la définition simple du « vrai ou faux », mais ne dit rien sur les cas indécidables et rien non plus sur les cas incertains.

Le deuxième mode, qui comporte cinq niveaux d'évaluation, se lit mieux sur des courbes qui expriment l'importance relative des succès selon le niveau de certitude. Plus spécifiquement, il exprime la proportion qui existe entre la capacité de détecter (la sensibilité) et la capacité d'exclure (la spécificité) selon le niveau de certitude.

La possibilité de quantifier l'incertitude peut sembler futile puisqu'elle ne sert à rien pour agir. Il faut au contraire inverser le raisonnement et se dire qu'un examen qui se révèle incertain dans une trop grande proportion de cas est inutile sinon dangereux, car en plus d'occasionner des coûts, il ne distingue rien. Dans une telle situation, il vaut mieux faire confiance au hasard, qui ne juge de rien ni de personne. Un tel essai retarde par contre le développement d'un test vraiment discriminant.

Application pratique

Les quatre tableaux de contingence et les explications qui suivent ont pour objectif de permettre au lecteur de comprendre les tableaux des résultats.

1. Répartition des sujets selon les facteurs étudiés

Pour bien comprendre le travail de discrimination imposé aux évaluateurs cliniques et aux utilisateurs des technologies, il faut retenir que chaque sujet choisi pour participer à l'étude porte une double condition. Il est d'une part « sain ou malade », ce qui représente le facteur « état de santé », et il est aussi désigné aléatoirement pour être « honnête ou malhonnête », ce qui représente le facteur « motivation ». L'ensemble de ces conditions donne le tableau de contingence 1. Ce tableau présente la répartition des sujets selon leur groupe d'appartenance.

Tableau de contingence 1

		FACTEUR « État de santé »		
		Malades	Sains	
FACTEUR « Motivation »	Malhonnêtes	20	23	43 Malhonnêtes
	Honnêtes	21	23	44 Honnêtes
		41	46	87 Total
		Malades	En santé	

On y distingue facilement les groupes de sujets selon les facteurs qu'ils représentent. Ces groupes sont importants, car ils servent de dénominateurs de départ aux différents rapports qui seront obtenus par les évaluateurs cliniques et par les utilisateurs des technologies.

2. Répartition des sujets après discrimination par les évaluateurs cliniques et les utilisateurs des technologies

Le vrai tableau de contingence qui fait l'objet de l'étude est constitué des rapports obtenus, pour chaque technique, entre le facteur de référence, appelé aussi "étalon de référence", et le facteur « diagnostic », qui est la conséquence du tri effectué par les évaluateurs cliniques et les utilisateurs de technologies.

Tableau de contingence 2

		RÉFÉRENCE (Étalon de référence)	
		Malades +	Sains -
		(a) vrais positifs	(b) faux positifs
DIAGNOSTIC	Malades +		
	Sains -	(c) faux négatifs	(d) vrais négatifs
		somme des malades a+c	somme des sains b+d
		sensibilité $\frac{a}{a+c} \times 100$	spécificité $\frac{d}{b+d} \times 100$
			somme totale a+b+c+d
			valeur globale $\frac{a+d}{a+b+c+d} \times 100$

Les résultats des interprétations sont « juste ou faux » en fonction de l'état de santé du patient qui est « en santé ou malade ». On les classe en « vrai positif ou vrai négatif », si le tri est exact. Par contre, si le résultat de l'évaluation est erroné, on les classe en « faux positif » lorsque le sujet est en santé mais que l'évaluateur clinique ou l'utilisateur de technologie le croit malade, et en « faux négatif » lorsque le sujet est malade mais que l'évaluateur clinique ou l'utilisateur des technologies le croit en santé.

Le rapport entre le nombre de diagnostics justes portés sur l'ensemble des sujets malades au départ est appelé « sensibilité ». C'est la capacité d'une technique diagnostique à détecter les sujets malades. Inversement, le rapport entre le nombre de diagnostics justes posés sur l'ensemble des sujets en santé au départ est appelé « spécificité ». C'est la capacité d'une technique diagnostique à détecter les sujets sains ou, plus précisément, à exclure les sujets malades (tableau 2).

Ce qui fait l'originalité de cette étude, c'est qu'au départ deux contraintes sont imposées à l'évaluateur clinique ou à l'utilisateur de technologie, et non une seule, comme c'est le cas habituellement : l'état de santé et la "motivation". Quelle que soit la technique que le clinicien ou le lecteur de technologie utilise pour établir son diagnostic, il a donc deux chances de se tromper et non une seule. Il peut se tromper sur l'état de santé ou sur la "motivation", ou sur les deux à la fois. Il a donc seulement une chance sur quatre de viser juste plutôt qu'une sur deux. Mais les deux sources d'erreur peuvent être identifiées en considérant le tableau de contingence complet (tableau 3).

Tableau de contingence 3

		RÉFÉRENCE			
		Malades		Sains	
		+		-	
DIAGNOSTIC		Honnêtes	Malhonnêtes	Honnêtes	Malhonnêtes
		Code		Code	
Malades +		10	11 dissimulateurs	faux positifs	
Sains -		faux négatifs		00	01 simulateurs

Les chiffres qui apparaissent dans ce tableau correspondent aux codes utilisés pour identifier les sujets selon leur groupe d'appartenance.

On peut donc déterminer la performance de l'évaluateur clinique et de l'utilisateur de technologie en sachant si l'état de santé ou la motivation du sujet se trouve à l'origine de son erreur.

Les résultats présentés au tableau de contingence 4 reproduisent les résultats du consensus des cliniciens à la première évaluation (voir tableau 9A en appendice).

Tableau de contingence 4
(se référer au tableau 9A, consensus
des cliniciens, à titre d'exemple)

		RÉFÉRENCE		N ^{bre} au départ
		Malades	Sains	
		Honnêtes Malhonnêtes	Honnêtes Malhonnêtes	
DIAGNOSTIC	Malades +	21	20	
		19	5	
		dissimulateurs		
		3	18	
		faux positifs		
	Sains -	23	23	N ^{bre} au départ
		2	15	
		faux négatifs		
		20	5	
		simulateurs		

49/87

Sensibilité totale	24/41	Sensibilité malhonnêtes	5/20
Spécificité totale	25/46	Spécificité malhonnêtes	5/23
Sensibilité honnêtes	19/21	Valeur globale	49/87
Spécificité honnêtes	20/23	Valeur honnêtes	39/44
		Valeur malhonnêtes	10/43

Toutes les évaluations ayant fait appel aux mêmes patients et suivi le même protocole expérimental, les résultats sont directement comparables. Il est également possible de comparer la performance de chaque évaluateur individuellement et à chacune des lectures. On peut en déduire l'importance relative de l'information disponible à chaque étape. On peut également comparer les techniques entre elles (l'examen clinique étant ici considéré comme l'une d'elles).

4.2 Résultats

Le tableau 1 en appendice fait état des caractéristiques des sujets malades et normaux de la population étudiée.

Évaluation de la lombalgie

1. La méthode COR

L'indice COR des différentes modalités dans l'évaluation de la lombalgie en fonction des lectures et des lecteurs pour l'ensemble des sujets est illustré à la figure 2.

Le tableau 2 présente la performance diagnostique mesurée par l'indice COR des différentes technologies et de l'examen clinique dans l'évaluation de la lombalgie selon les évaluateurs, conditionnelle à l'état réel des sujets malhonnêtes ou honnêtes. La figure 3 montre les intervalles de confiance (95 %).

Pour l'ensemble des sujets et pour la première lecture, les indices COR obtenus par le meilleur des lecteurs de chaque méthode, sans les informations cliniques, étaient, par ordre décroissant, de 0,85 pour le B-200 2^e lecteur, de 0,79 pour la spinoscopie 2^e lecteur, de 0,76 pour l'examen clinique 1^{er} lecteur et de 0,35 pour la thermographie 1^{er} lecteur.

Pour la deuxième lecture, alors que les informations des évaluateurs cliniques étaient fournies aux utilisateurs des technologies, les résultats étaient, par ordre décroissant pour le meilleur des lecteurs de chaque méthode, de 0,85 pour le B-200 2^e lecteur, de 0,74 pour la spinoscopie 1^{er} lecteur et de 0,55 pour la thermographie 2^e lecteur.

Lors de la troisième lecture découlant du consensus de deux utilisateurs des technologies avec les informations cliniques, les résultats étaient respectivement de 0,80 pour le B-200, de 0,79 pour la spinoscopie et de 0,50 pour la thermographie.

La combinaison la plus performante et la plus consistante était le B-200 avec les informations cliniques (lecteur 1 = 0,84 et lecteur 2 = 0,85).

Pour les sujets honnêtes, les indices COR des technologies lors de la première lecture sans les informations cliniques étaient les suivants, par ordre de performance des lecteurs : en regard de l'examen clinique de 0,99 pour le 1^{er} lecteur et le 2^e lecteur, de 0,92 pour le B-200 2^e lecteur et de 0,89 pour le 1^{er} lecteur, de 0,81 pour la spinoscopie 1^{er} lecteur et de 0,74 pour le 2^e lecteur, de 0,35 pour la thermographie 1^{er} lecteur et de 0,24 pour le 2^e lecteur. Quant aux technologies, l'ajout des informations cliniques lors de la deuxième lecture augmentait ou diminuait l'indice COR de certains lecteurs : 0,99 pour le B-200 2^e lecteur et 0,98 pour le 1^{er} lecteur, 0,92 pour la spinoscopie 1^{er} lecteur et 0,72 pour le 2^e lecteur, 0,55 pour la thermographie 2^e lecteur et 0,48 pour le 1^{er} lecteur. La troisième lecture par consensus n'améliorait pas les performances.

Pour les sujets malhonnêtes, la première lecture sans les informations cliniques donnait ce qui suit : 0,84 pour la spinoscopie 2^e lecteur et 0,61 pour le

1^{er} lecteur, 0,80 pour le B-200 2^e lecteur et 0,64 pour le 1^{er} lecteur, 0,40 pour la thermographie 2^e lecteur et 0,36 pour le 1^{er} lecteur, en regard de l'examen clinique 0,36 pour le 1^{er} lecteur et 0,24 pour le 2^e lecteur. Pour les sujets malhonnêtes, l'ajout des informations cliniques diminuait les performances respectivement à 0,73 pour la spinoscopie 2^e lecteur, à 0,65 pour le B-200 1^{er} lecteur et augmentait à 0,65 pour la thermographie 1^{er} lecteur. La troisième lecture par consensus n'améliorait pas les performances.

Évaluation du rôle de simulateur et de dissimulateur

Le tableau 3 présente les indices COR des différentes modalités dans l'évaluation de l'état de simulateur selon les lecteurs, conditionnelle à l'état réel des sujets sains ou malades, alors que la figure 4 illustre les intervalles de confiance (95 %) correspondants. La thermographie est absente puisque sans objet, les sujets ne pouvant alors simuler ou dissimuler.

Pour l'ensemble des sujets, les indices COR obtenus par le meilleur des lecteurs pour chaque méthode, sans les informations cliniques étaient, par ordre décroissant, de 0,80 pour l'examen clinique 2^e lecteur, de 0,67 pour le B-200 1^{er} lecteur et de 0,60 pour la spinoscopie 1^{er} lecteur.

Pour la deuxième lecture, alors que les informations des évaluateurs cliniques étaient fournies aux utilisateurs des technologies, les résultats étaient, par ordre décroissant pour le meilleur lecteur de chaque technologie, de 0,84 pour le B-200 2^e lecteur et de 0,76 pour la spinoscopie 1^{er} lecteur.

Lors de la troisième lecture découlant du consensus de deux utilisateurs des technologies avec les informations cliniques, les résultats étaient respectivement de 0,79 pour le B-200 et de 0,79 pour la spinoscopie.

Pour les sujets sains, les indices COR des technologies lors de la première lecture sans les informations cliniques étaient les suivants, par ordre de performance des lecteurs : de 0,68 pour la spinoscopie 1^{er} lecteur et de 0,67 pour le 2^e lecteur, de 0,64 pour le B-200 1^{er} lecteur et de 0,49 pour le 2^e lecteur et, en regard de l'examen clinique, de 0,85 pour le 2^e lecteur et de 0,75 pour le 1^{er} lecteur.

Pour la deuxième lecture, alors que les informations des évaluateurs cliniques étaient fournies aux utilisateurs des technologies, les résultats étaient, par ordre décroissant, de 0,95 pour la spinoscopie 1^{er} lecteur, de 0,81 pour le B-200 2^e lecteur et de 0,79 pour le 1^{er} lecteur et de 0,51 pour la spinoscopie 2^e lecteur.

Lors de la troisième lecture découlant du consensus de deux utilisateurs des technologies avec les informations cliniques, les résultats étaient respectivement de 0,82 pour la spinoscopie et de 0,78 pour le B-200.

Pour les sujets malades, les indices COR des technologies lors de la première lecture sans les informations cliniques étaient les suivants, par ordre de

performance des lecteurs : de 0,73 pour le B-200 1^{er} lecteur et de 0,53 pour le 2^e lecteur, de 0,51 pour la spinoscopie 1^{er} lecteur et de 0,50 pour le 2^e lecteur et, en regard de l'examen clinique, de 0,82 pour le 1^{er} lecteur et de 0,73 pour le 2^e lecteur.

Pour la deuxième lecture, alors que les informations des évaluateurs cliniques étaient fournies aux utilisateurs des technologies, les résultats étaient, par ordre décroissant, de 0,87 pour le B-200 2^e lecteur, de 0,82 pour la spinoscopie 2^e lecteur, de 0,79 pour le B-200 1^{er} lecteur et 0,49 pour la spinoscopie 1^{er} lecteur.

Lors de la troisième lecture découlant du consensus de deux utilisateurs des technologies avec les informations cliniques, les résultats étaient respectivement de 0,80 pour le B-200 et de 0,76 pour la spinoscopie.

Les COR (figures 5A et 5B) illustrent clairement les performances différentes des technologies et de l'examen clinique dans l'évaluation de la lombalgie en fonction de l'ajout des informations et en tenant compte de la moyenne des deux lecteurs pour les sujets honnêtes et les sujets malhonnêtes.

Les tableaux 4 et 5 font état des indices COR pour l'évaluation de la lombalgie selon les lecteurs en fonction respectivement du sexe et de l'âge. L'on observe une tendance à obtenir une moins bonne performance dans l'identification de la lombalgie chez les femmes, surtout marquée chez les évaluateurs cliniques, tendance qui se transpose au moment des lectures avec les informations cliniques pour les technologies, mais avec plus de variations.

Le tableau 6 illustre les comparaisons statistiques des indices COR de l'évaluation de la lombalgie entre les différentes technologies pour les sujets honnêtes de même que pour les sujets malhonnêtes. Ce tableau présente aussi les comparaisons statistiques des indices COR pour l'évaluation du rôle de simulateur entre les différentes technologies pour les sujets sains de même que pour les sujets malades.

Mise en garde : Le niveau des tests de comparaison est soumis à titre indicatif (objectifs exploratoires), car la multiplicité des comparaisons (non orthogonales) non corrigées statistiquement comporte le risque de surestimer les différences.

Évaluation de la lombalgie

- **Sujets honnêtes :** l'examen clinique est supérieur à la thermographie 1^{er} lecteur ($p < 0,01$) et 2^e lecteur ($p < 0,01$); au B-200 1^{er} lecteur ($p < 0,01$); à la spinoscopie 1^{er} lecteur ($p < 0,01$) et 2^e lecteur ($p < 0,02$); à la thermographie + examen clinique 1^{er} lecteur ($p < 0,01$) et 2^e lecteur ($p < 0,01$); à la spinoscopie + examen clinique 1^{er} lecteur ($p < 0,03$) et 2^e lecteur ($p < 0,01$).

- . **Sujets malhonnêtes** : le B-200 1^{er} lecteur ($p < 0,02$) et 2^e lecteur ($p < 0,01$) ainsi que la spinoscopie 2^e lecteur ($p < 0,01$) sont supérieurs à l'examen clinique, de même que le B-200 + examen clinique 1^{er} lecteur ($p < 0,03$) et 2^e lecteur ($p < 0,01$) et la spinoscopie + examen clinique 2^e lecteur ($p < 0,01$). Le B-200 2^e lecteur ($p < 0,01$) est supérieur au B-200 + examen clinique.
- . Le B-200, comparé à la spinoscopie, ne montre pas de différences statistiquement significatives tant pour les sujets honnêtes que malhonnêtes. Avec les informations cliniques, le B-200 est supérieur à la spinoscopie 2^e lecteur ($p < 0,01$) pour les sujets honnêtes, mais ne montre pas de différence pour les sujets malhonnêtes.

Évaluation du rôle de simulateur

- . **Sujets sains (simulateurs)** : l'examen clinique est supérieur au B-200 2^e lecteur ($p < 0,01$); le B-200 + examen clinique est supérieur au B-200 2^e lecteur ($p < 0,01$); la spinoscopie + examen clinique est supérieure à la spinoscopie 1^{er} lecteur ($p < 0,01$); la spinoscopie + examen clinique est supérieure à l'examen clinique 1^{er} lecteur ($p < 0,02$) et l'examen clinique 2^e lecteur ($p < 0,01$) est supérieur à la spinoscopie + examen clinique.
- . **Sujets malades (dissimulateurs)** : l'examen clinique est supérieur à la spinoscopie 1^{er} lecteur ($p < 0,01$) et 2^e lecteur ($p < 0,04$). L'examen clinique 1^{er} lecteur est supérieur à la spinoscopie + examen clinique ($p < 0,01$). Le B-200 + examen clinique est supérieur au B-200 2^e lecteur ($p < 0,01$). La spinoscopie + examen clinique est supérieure à la spinoscopie 2^e lecteur ($p < 0,01$).
- . Le B-200, comparé à la spinoscopie, ne montre pas de différences statistiquement significatives tant pour les sujets sains que malades. Avec les informations cliniques, pour les sujets sains, la spinoscopie 1^{er} lecteur est supérieure au B-200 ($p < 0,03$) alors que la spinoscopie 2^e lecteur est inférieure au B-200 ($p < 0,01$). Pour les sujets malades, le B-200 est supérieur à la spinoscopie 1^{er} lecteur ($p < 0,01$).

2. Les méthodes d'exactitude simple et double

Nous avons aussi procédé à une analyse de la sensibilité (aptitude à détecter la lombalgie parmi les sujets malades), de la spécificité (aptitude à détecter l'absence de la lombalgie parmi les sujets sains) et de la valeur globale ou l'efficacité d'un test (proportion de décisions correctes).

Les résultats selon le premier niveau de classification, à savoir l'exactitude simple (aptitude à distinguer les sujets sains et malades basée sur l'état de santé seul) en rapport avec l'état réel des sujets, honnêtes ou malhonnêtes, sont résumés dans les tableaux 7 et 8.

Ces tableaux permettent de constater la variabilité entre les lecteurs de la spinoscopie particulièrement en ce qui concerne la sensibilité chez les sujets honnêtes et malhonnêtes et la spécificité chez les sujets honnêtes, de même que la valeur globale. Ces variations sont moindres pour le B-200 et peu marquées pour l'examen clinique.

Les tableaux 9A à 9G donnent les informations brutes selon le deuxième niveau de classification, à savoir l'exactitude double (aptitude à distinguer les sujets selon les quatre groupes).

Pour les deux niveaux, une réponse classée indécise par les évaluateurs cliniques et les utilisateurs des technologies était considérée comme une mauvaise réponse.

Le tableau 10 fait état de l'aptitude à distinguer les sujets des quatre groupes (exactitude double basée sur l'état de santé et le rôle) quant à la sensibilité, la spécificité et la valeur globale. Il permet de constater l'inaptitude de l'examen clinique et de toutes les technologies, avec ou sans les informations cliniques, à distinguer les sujets qui simulent ou dissimulent une entorse lomboire.

Les tableaux 11A et 11B présentent les résultats selon les méthodes d'analyses utilisées pour l'évaluation de la lombalgie suivant l'état réel de l'ensemble des sujets honnêtes et malhonnêtes en tenant compte de la moyenne des deux évaluateurs. Ils synthétisent l'essentiel des informations colligées et permettent d'évaluer les performances respectives quant aux valeurs globales des instruments diagnostiques et de l'examen clinique dans les lombalgies mécaniques communes (entorses lombaires).

Ces tableaux établissent la concordance des méthodes de mesures utilisées en rapport avec le niveau de difficulté exigé. Le tableau 11A montre que même si l'on fait une moyenne des deux lecteurs, les technologies, soit le B-200 et la spinoscopie, sont non performantes dans le cas des sujets malhonnêtes à l'exactitude double lorsqu'il s'agit de détecter le groupe réel d'appartenance par rapport au double état (simulateur/dissimulateur et lombalgique ou non). Cette lacune des technologies seules se reflète sur la valeur globale inférieure de l'exactitude double (tableau 11B).

5. DISCUSSION

L'équipe de recherche reconnaît que les exigences de performance attendues des instruments diagnostiques étaient élevées. On demandait aux technologies de statuer sur un état de santé ou de maladie et sur un état de simulateur et ou de dissimulateur pour lequel il n'existe actuellement aucun standard autre que l'expertise clinique, dont la valeur absolue n'est pas connue.

L'équipe de recherche, la documentation scientifique et le consensus médical reconnaissent aussi que l'ajout des informations cliniques aux résultats des

techniques d'imagerie médicale dans les radiculopathies lombaires secondaires à des hernies discales ne prêtent pas à un long débat sur l'existence de ce type d'affections par les experts médicaux. Par ailleurs, l'imagerie médicale possède ses limites et ses zones grises⁸⁰. Elle est entre autres dépendante, comme plusieurs instruments diagnostiques, d'une interprétation temporelle.

Cette lacune admise, la présente recherche avait pour objectif d'évaluer le pouvoir discriminant diagnostique des technologies dans une sphère de pathologie beaucoup moins identifiable, soit l'entorse lombaire. Certaines de ces technologies prétendent pouvoir identifier les malades présentant ce type d'affection, les distinguant ainsi des sujets sains et, plus encore, pouvoir distinguer les sujets simulateurs ou dissimulateurs, soit deux niveaux de questionnement qui représentent un défi jusqu'à présent non résolu. Il s'agissait de déterminer les caractéristiques diagnostiques des technologies et de l'examen clinique, puis de comparer les résultats. La méthode d'analyse a permis de tenir compte des niveaux de difficulté introduits par l'utilisation de sujets malhonnêtes simulant ou dissimulant une lombalgie en employant, en plus, des méthodes d'analyse conventionnelle de la sensibilité, de la spécificité et de la valeur globale sur une échelle ordinale graduée évaluant le niveau de certitude. Les résultats ont aussi été analysés en tenant compte de l'apport des informations cliniques. Ce modèle établi d'utilisation progressive des informations cliniques et du consensus de deux évaluateurs cliniques est couramment utilisé en clinique et en recherche. Les résultats démontrent normalement une amélioration de la sensibilité et de la spécificité des résultats. Cette méthode permet en outre l'analyse des variations entre les lectures pour le même évaluateur et entre les évaluateurs.

La valeur diagnostique potentielle des technologies repose essentiellement sur le résultat clinique global (validité)⁸¹. Il va de soi qu'une technique fiable constitue une prémisse nécessaire à la validation d'un instrument diagnostique. L'équipe de recherche était d'avis que la responsabilité première de produire un instrument fiable était du ressort des fabricants de ces appareils.

L'efficacité clinique, par ailleurs, constitue le champ d'expertise de la recherche clinique en médecine. Évaluer cette efficacité (validité clinique) constituait l'objet même de cette étude. Celle-ci s'avérait d'autant plus nécessaire qu'elle était susceptible d'apporter des éléments qui contribueraient à documenter les coûts-bénéfice reliés à l'utilisation de ces technologies.

Cette étude compare les performances diagnostiques de trois technologies - la thermographie, la dynamométrie triaxiale et la spinoscopie - par rapport à leurs aptitudes à détecter une lombalgie mécanique commune récente (entorse lombaire). Leurs performances ont aussi été comparées à celles d'évaluateurs cliniques. L'étalon de référence pour tous était le consensus de deux cliniciens sélectionneurs autres que les évaluateurs cliniques. Il n'est donc pas inattendu que l'examen clinique puisse être performant, comme le démontrent les résultats pour les sujets honnêtes. Nonobstant ce constat méthodologique pouvant les favoriser (étalon de référence clinique), la performance réduite dans le cas des sujets malhonnêtes suggère plutôt que les évaluateurs cliniques n'ont pas eu un

avantage par rapport aux technologies. Il est en effet probable qu'ils aient accordé le bénéfice du doute aux sujets, jusqu'à preuve du contraire, comme cela est habituel, puisque leur rôle consiste à traiter des malades.

Les résultats de l'étude démontrent que la thermographie n'est d'aucune utilité clinique dans le diagnostic d'une lombalgie mécanique commune récente. Son emploi avait d'ailleurs comme objectif de servir de point de référence inférieur probable permettant la comparaison avec les autres technologies et avec l'examen clinique. Sa pauvre performance, soupçonnée au départ et confirmée par l'étude, ne justifie pas l'emploi de cette technologie dans le diagnostic des affections lombaires de ce type. Les performances diagnostiques de la dynamométrie triaxiale, de la spinoscopie et de l'examen clinique sont cliniquement et statistiquement supérieures aux siennes. L'analyse n'a toutefois pas démontré de différence statistique favorisant l'emploi de la dynamométrie ou de la spinoscopie.

Un des aspects importants de cette étude a été l'inclusion de sujets sains simulant une lombalgie et de malades la dissimulant. Ce procédé méthodologique a permis de constater que la dynamométrie triaxiale et la spinoscopie paraissent offrir un certain avantage dans de telles circonstances. Ces deux technologies semblent être plus performantes que l'examen clinique chez les sujets simulateurs et dissimulateurs selon l'analyse COR des sujets malhonnêtes, mais pas dans le cas des sujets honnêtes ni pour la totalité des sujets (tableau 2 en appendice). Par ailleurs, les résultats obtenus par la méthode de l'exactitude simple démontrent une baisse de cette performance (tableau 7 en appendice), alors que les technologies s'avèrent non performantes (tableaux 9A à 9D en appendice) lorsqu'on utilise la méthode de l'exactitude double, plus contraignante.

Ces baisses de performance, présentées dans les tableaux 11A et 11B, sont attribuables au fait que l'indécision pénalise les évaluateurs cliniques et les utilisateurs des technologies employant les deux méthodes. De plus, la méthode de l'exactitude double réduit davantage la performance totale puisque les évaluateurs cliniques et les utilisateurs des technologies doivent obligatoirement classer chaque sujet dans la catégorie à laquelle il appartient. Ces contraintes sont cependant très représentatives de l'exigence diagnostique imposée à un instrument ou à un clinicien. Les résultats de cette étude expérimentale sur une population de lombalgiques présentant une entorse lombaire ne permettent pas de confirmer l'utilité du B-200 et de la spinoscopie comme instruments pouvant être aptes à distinguer les sujets malhonnêtes, tel que le rapportent certains auteurs^{59,60}.

Par ailleurs, l'addition des informations cliniques améliore la consistance des résultats pour des lectures du B-200, mais a un effet négatif sur la performance de la spinoscopie. Cette dernière technologie démontre une plus grande variation entre le même lecteur et entre les lecteurs, que l'on utilise la méthode COR, l'exactitude simple ou l'exactitude double. L'ajout d'un consensus entre les utilisateurs des technologies a un effet égalisateur sur les performances de ces technologies.

Par contre, l'ajout des informations cliniques à la méthode de l'exactitude double améliore les performances autrement faibles du B-200 et de la spinoscopie.

Les résultats permettent aussi de constater que les évaluateurs cliniques ne modifient pas leurs décisions une fois connus les résultats des technologies (c.-à-d. ceux de la dynamométrie triaxiale et de la thermographie). Ces mêmes évaluateurs cliniques semblent moins performants pour identifier la lombalgie chez les femmes.

Les résultats de la recherche permettent donc de confirmer les multiples sources de variations entre le même utilisateur de technologie et entre les divers utilisateurs, avec ou sans l'ajout des informations cliniques, mais sans pouvoir mesurer statistiquement ces variations. En effet, chaque utilisateur de technologie a employé des critères qui lui étaient propres et, par définition, non comparables puisqu'ils pouvaient changer au cours des lectures, ce qui fut probablement le cas. Par ailleurs, les résultats bruts de l'exactitude simple et double illustrent l'importance de ces variations (tableaux 8 et 10 en appendice).

L'étude des résultats bruts et des résultats obtenus après les analyses descriptives et statistiques permet de formuler des commentaires de trois natures : quantitative, qualitative et pragmatique.

1. Commentaires de nature quantitative

En regardant les résultats en détail, on est frappé par l'importance des différences existant entre eux, les résultats couvrant toute l'échelle des possibilités de mesure pour les rapports et les proportions de 0 à 100 %. Cela a pour conséquence de montrer que les choix offerts par la structure de l'étude ont couvert toute la gamme des réponses. Cette constatation signifie que les comparaisons disponibles ne sont pas réduites à des différences insignifiantes sur une partie d'une échelle de quantification, mais sur la presque totalité des possibilités de numérisation.

Cette observation permet d'interpréter directement la valeur des « résultats bruts » obtenus de chaque évaluation individuelle pour chaque sous-groupe. Il s'agit de la quantification de la difficulté à distinguer entre l'état clinique de chaque sous-groupe. Chacun peut en tirer des conclusions et agencer les résultats. Par contre, les « moyennes » obtenues après regroupement, ou la moyenne générale, ne représentent vraisemblablement pas les valeurs auxquelles on pourrait s'attendre si l'on étudiait une population générale, car les proportions de sujets sains ou malades, aussi bien que de sujets honnêtes ou malhonnêtes, n'y sont pas égales, comme dans ce plan expérimental. Il faut donc interpréter avec une certaine prudence ces moyennes qui ne se retrouveraient pas dans une étude portant sur un échantillon d'une population générale.

Cette remarque prend de l'importance si l'on tente de déduire les chances de réussite ou d'échec en évaluant le coût de l'utilisation des technologies sur une telle population. La proportion d'individus en santé ou malades susceptibles de

tricher (les simulateurs et les dissimulateurs), bien qu'elle soit inconnue, est fort probablement différente de celle des individus malades honnêtes.

En outre, dans la pratique clinique quotidienne, la frontière n'est pas toujours bien délimitée entre la malhonnêteté flagrante, un certain degré d'amplification consciente des symptômes et des signes d'une pathologie réelle et l'incapacité fonctionnelle démesurée par rapport à une pathologie donnée.

2. Commentaires de nature qualitative

Les résultats quantitatifs obtenus proviennent de la constitution d'« informations » tirées de signes observés par chaque évaluateur selon sa méthode de mesure, clinique ou technologique. Cela veut dire que l'évaluation commence par la recherche de signes, qui sont ensuite interprétés en fonction de valeurs de normalité pour déduire l'état du sujet. Il existe donc au moins trois sources d'erreur possibles sur le plan du « signe », qui peut être faussement présent ou absent, ou sur le plan de son « interprétation ». Un troisième niveau pourrait être celui de valeurs de « normalité » erronées. L'étude ne dit pas où l'erreur s'est produite. La présence de tricheurs vise précisément à fausser le niveau du signe, mais la possibilité d'erreur sur le plan de l'interprétation ou des normes reste possible et probable, comme le démontrent les résultats de la recherche. Il faut se poser les questions suivantes : À quel niveau doit-on intervenir pour améliorer la performance diagnostique ? Faut-il imaginer des cliniciens ou des appareils faisant plus de mesures (signes plus abondants et plus précis), un protocole d'interprétation plus rigoureux (par exemple, avec un programme d'intelligence artificielle), ou une révision des valeurs de normalité avec des critères plus détaillés tenant compte, par exemple, du sexe, de l'âge et de la morphométrie ?

Ces résultats mettent en évidence un autre aspect du traitement de l'information. Il s'agit essentiellement de la non-addition de l'information obtenue par des méthodes d'examen différentes. Cela se constate au moment de l'ajout des données cliniques aux trois technologies, et vice-versa. L'information des technologies ne modifie pour ainsi dire pas les résultats des évaluateurs cliniques alors que l'inverse est vrai : les résultats des lecteurs de technologies sont modifiés en mieux ou en pire par l'information clinique.

Le consensus est aussi la combinaison de deux informations, pourtant de même nature. Celles-ci ne s'additionnent pas systématiquement pour donner un résultat meilleur ou pire. Malgré l'utilisation de grilles, de procédures et d'algorithmes propres à chaque méthode, les informations de chaque évaluateur ne se combinent pas de manière prévisible.

Par ailleurs, il faut être bien conscient du fait que chaque sujet de l'étude présentait toujours deux informations en interaction : l'une liée au facteur d'état de santé, l'autre au facteur d'honnêteté. Une erreur sur l'un de ces facteurs se répercute automatiquement sur l'autre. Par exemple, l'incapacité à déterminer si un sujet est un simulateur (en santé mais tricheur) amène l'évaluateur clinique et l'utilisateur des technologies à le classer parmi les

malades honnêtes. Or, on constate de nettes différences entre l'erreur provoquée par chacun des facteurs, selon la méthode d'évaluation. Les évaluateurs cliniques sont beaucoup plus influencés par le facteur d'honnêteté. Leur performance est basse si le sujet est malhonnête. Les évaluations pratiquées avec le B-200 et le spino-scopie sont, elles aussi, peu performantes pour les dissimulateurs qui sont malades et les simulateurs qui sont en santé. Il est probable que le traitement des signes observés dans chacune de ces catégories de sujets, pour en tirer de l'information, soit très différent selon la méthode. Chaque méthode engendre ses signes particuliers, les interprète différemment et utilise forcément des critères de normalité qui lui sont propres. Cet effet pourrait constituer un avantage pour distinguer les sujets selon des sources variées d'information et d'interprétation, si ces sources s'additionnaient lorsqu'on les réunit. Ce n'est manifestement pas le cas.

3. Commentaires de nature pragmatique

Cette recherche a permis de clarifier plusieurs aspects de la problématique, même si elle n'apporte pas de réponse à toutes les questions. L'absence de réponse tranchée à certaines interrogations constitue en soi une réponse valable : il n'existe pas de réponse tranchée à certaines questions. Que ce soit à cause de l'avancement des connaissances ou des caractéristiques intrinsèques à chaque méthode (y compris à la clinique) à procurer des signes, des procédures d'interprétation suffisantes pour engendrer de l'information ou enfin, des normes de références indiscutables, le pouvoir discriminant moyen de toutes ces méthodes est au mieux médiocre. En effet, ce pouvoir s'avère élevé si les patients sont malades et honnêtes, mais il chute considérablement dès que ces conditions ne sont pas remplies. En plus, il faut composer avec la variation de performance entre les évaluateurs cliniques et les utilisateurs des technologies ainsi qu'avec leur capacité à s'entendre par consensus pour améliorer la performance.

De savoir que des instruments sont parfois meilleurs, parfois moins bons que des cliniciens, selon les conditions imposées, ne change pas le niveau de discrimination global de chaque méthode puisque les conditions des patients (malade ou en santé, simulateur ou dissimulateur) ne sont pas connues dans la vie réelle. Voilà la vraie réponse. C'est préférable que de savoir si une méthode technologique est un peu meilleure ou moins bonne qu'une autre. On peut, par contre, déterminer qu'une méthode (la thermographie) n'est jamais bonne.

L'analyse de l'ensemble des résultats démontre qu'aucune technologie, pas plus que l'examen clinique, n'a une performance vraiment fiable pour détecter les sujets malhonnêtes, quel que soit le type d'analyse utilisée. Toutefois, l'examen clinique a une performance optimale chez les sujets honnêtes (tableaux 11A et 11B en appendice).

Aucune technologie n'a révélé une performance permettant d'établir son utilité clinique, à savoir une sensibilité de 85 % et une spécificité de 90 %, tel que déterminé *a priori* pour cette étude. Ce seuil de performance, qui peut paraître élevé, a été choisi compte tenu des incidences tant financières que morales pour les individus et la population en général. En effet, statuer erronément qu'un

malade est malhonnête se traduirait par la perte de l'indemnité à laquelle il a droit et le stigmatiserait sans preuve.

L'on pourrait toujours argumenter, comme une des limites à la portée des résultats de cette recherche, que des utilisateurs plus expérimentés de ces technologies ou encore des médecins différents auraient pu obtenir de meilleurs résultats. Soulignons que les utilisateurs des technologies et les évaluateurs cliniques ayant participé à cette étude sont représentatifs des personnes qui font usage de ces technologies.

En outre, il existe maintenant des programmes informatisés d'analyse des données de ces technologies, dont pour le B-200 et la spinoscopie, qui n'étaient pas disponibles lorsque cette recherche a été entreprise. Peut-être que les résultats auraient alors été différents et qu'ils pourraient l'être à l'avenir. Mais aucun de ces programmes n'a jusqu'à maintenant été validé scientifiquement.

À ce propos, une analyse des résultats selon une méthode utilisant un système programmé pour la spinoscopie a été réalisée par la maison Spinex. Les résultats, présentés en détail à l'Annexe I, ne montrent pas une performance supérieure à celle des utilisateurs choisis pour la spinoscopie dans le cours de cette recherche.

6. CONCLUSION

Cette étude expérimentale avait pour objectif d'évaluer la validité interne d'instruments diagnostiques des lombalgies. Compte tenu des limites de cette étude en ce qui concerne les malades présentant une lombalgie mécanique commune (entorse lombaire) d'une durée inférieure à 10 semaines avec absence de radiculopathie, de déficit neurologique ou de maladie radiologique importante, les constats suivants s'imposent :

- . L'emploi à des fins diagnostiques de la thermographie, de la dynamométrie triaxiale et de la spinoscopie chez un sujet sain ou présentant une lombalgie mécanique commune n'est d'aucune utilité lorsque ce dernier est honnête puisque, dans ces conditions, aucune méthode n'est supérieure à l'examen clinique.
- . Les résultats de l'étude actuelle ne démontrent pas l'utilité de l'emploi des technologies à des fins diagnostiques chez des sujets sains mais pouvant simuler une lombalgie, ou encore chez des sujets malades désirant dissimuler leur maladie.

7. RETOMBÉES ÉVENTUELLES

Les résultats de cette étude suggèrent qu'une évaluation ponctuelle de sujets sains ou de malades lombalgiques n'est actuellement ni fiable ni valide s'il s'agit de simulateurs ou de dissimulateurs (état réel). L'emploi des technologies à ces fins peut induire en erreur et porter un préjudice financier et moral important à la population.

Soulignons en outre, au regard de ce constat, qu'il s'agissait d'une recherche expérimentale, c.-à-d. que les catégories de malades étaient clairement définies. Dans la pratique courante, cela est inhabituel. L'ensemble de la population de lombalgiques est beaucoup moins homogène et compte des malades dont l'état de maladie réel est obscurci par plusieurs déterminants plus ou moins conscients. De plus, la lombalgie varie dans le temps et est fonction de diverses mises en situation.

On aurait donc probablement intérêt à utiliser une méthode permettant des évaluations observationnelles multiples dans le temps au cours du processus de réhabilitation. Cela serait probablement plus juste et plus pertinent. Les ressources financières pourraient sans doute servir plus utilement à d'autres fins qu'à celle de rechercher des simulateurs.

Remerciements

L'équipe de recherche désire remercier les techniciens des instruments diagnostiques qui ont participé à cette recherche, soit M^{mes} Rolande Soulière et Lise St-Laurent ainsi que MM. Steeve Asselin, M.Sc. et Victor Lanzo, B.Eng., B.Sc. Merci également à M. Thierry Petitjean Roget, de l'IRSST, pour le choix par randomisation et le codage des sujets.

BIBLIOGRAPHIE

1. Fineberg HV, Hiatt HH. Evaluation of medical practices: the case for technology assessment. *N Engl J Med* 1979;301:1086-91.
2. Health Care Technology. Effectiveness, Efficiency and Public Policy. Feeny D, Guyatt G, Tugwell P, editors. Institute for Research on Public Policy, Montreal, 1986.
3. CSST. Statistiques sur les affections vertébrales 1990-1993, DSGI, Service de la statistique, tableaux 10 et 11, septembre 1994;19.
4. Quebec Task Force on Spinal Disorders. Scientific approach to the assessment and management of activity-related spinal disorders. *Spine* 1987;12 Suppl:1-59.
5. Deyo RA. Measuring the functional status of patients with low back pain. *Arch Phys Med Rehabil* 1988;69:1044-53.
6. Hoffman RM, Kent DL, Deyo RA. Diagnostic accuracy and clinical utility of thermography for lumbar radiculopathy: a meta-analysis. *Spine* 1991;16:623-8.
7. Newton M, Waddell G. Trunk strength testing with Iso-machines. Part 1: Review of a decade of scientific evidence. *Spine* 1993;18:801-11.
8. Newton M, Thow M, Somerville D, Henderson I, Waddell G. Trunk strength testing with Iso-machines. Part 2: Experimental evaluation of the Cybex II back testing system in normal subjects and patients with chronic low back pain. *Spine* 1993;18:812-24.
9. Lawson RN. Implications of surface temperatures in the diagnosis of breast cancer. *Can Med Assoc J* 1956;75:309.
10. Barnes RB. The early history of thermography. In: Uematsu S, editor. *Medical thermography: theory and clinical applications*. Los Angeles: Brentwood Publishing Co., 1976:1-14.
11. Pochaczewsky R, Wexler CE, Meyers PH, Epstein JA, Marc JA. Liquid crystal thermography of the spine and extremities. Its value in the diagnosis of spinal root syndromes. *J Neurosurg* 1982;56:386-95.
12. Edeiken J, Wallace JD, Curley RF, Lee S. Thermography and herniated lumbar disks. *A J R* 1968;102:790-6.
13. Duensing F, Becker P, Rittmeyer K. Thermographic findings in lumbar disk protrusions. *Arch Psychiatr Nervenkr* 1973;217:53-70.

14. Raskin MM, Martinez-Lopez M, Sheldon JJ. Lumbar thermography in discogenic disease. *Radiology* 1976;119:149-52.
15. Tichauer ER. The objective corroboration of back pain through thermography. *J Occup Med* 1977;19:727-31.
16. Ching C, Wexler CE. Peripheral thermographic manifestations of lumbar disc disease. *Applied Radiology* 1978;100:53-8.
17. Wexler CE. An overview of liquid crystal and electronic lumbar, thoracic and cervical thermography. Tarzana (CA): Thermographic Services Inc., 1981.
18. Uematsu S, Hendler N, Hungerford D, et al. Thermography and electromyography in the differential diagnosis of chronic pain syndromes and reflex sympathetic dystrophy. *Electromyogr Clin Neurophysiol* 1981;21:165-82.
19. Pochaczewsky R, Feldman F. Contact thermography of spinal root compression syndromes. *Neuroradiology* 1982;3:243-50.
20. Rubal BJ, Traycoff RB, Ewing KL. Liquid crystal thermography: a new tool for evaluating low back pain. *Physical Therapy* 1982;62:1593-6.
21. Hendler N, Uematsu S, Long D. Thermographic validation of physical complaints in "psychogenic pain" patients. *Psychosomatics* 1982;23:283-7.
22. Wexler CE. Atlas of thermographic lumbar patterns. Tarzana (CA): Thermographic Services Inc., 1983.
23. Pochaczewsky R. Assessment of back pain by contact thermography of extremity dermatomes. *Orthop Rev* 1983;12:45-58.
24. Fischer AA, Chang CH, Kuo JC. Value of thermography in the diagnosis of radiculopathy as compared with electrodiagnosis. *Arch Phys Med Rehabil* 1983;64:526.
25. Nakano K. Liquid crystal contact thermography (LCT) in the clinic evaluation of traumatic low back pain (LBP). *J Neurol Orthop Med Surg* 1984;5:207-11.
26. Newman RI, Seres JL, Miller EB. Liquid crystal thermography in the evaluation of chronic back pain: a comparative study. *Pain* 1984;20:293-305.
27. Hubbard JE, Hoyt C. Pain evaluation in 805 studies by infrared imaging. *Thermology* 1986;1:161-6.

28. Chafetz N, Wexler CE, Kaiser JA. Neuromuscular thermography of the lumbar spine with CT correlation. *Spine* 1988;13:922-5.
29. Mahoney L, McCulloch J, Csima A. Thermography in back pain.
1. Thermography as a diagnostic aid in sciatica. *Thermology* 1985;1:43-50.
30. Weinstein SA, Weinstein G. Trends in clinical practice induced by standardized electronic thermography: a review of 2,000 patients. *Mod Med* 1987;55 Suppl:58-64.
31. Hubbard JE, Hoyt C. Pain evaluation by electronic infrared thermography: Correlations with symptoms, EMG, myelogram and CT scan. *Thermology* 1985; 1:26-35.
32. Addatto KA, Phillips SI, Manale BL, Watermeier JJ, King AG, Bartholomew PM. CT & thermogram, a comparison of 91 patients. *Ortho Trans* 1985; 9:215-6.
33. Leroy PL, Christian CR, Filasky R. Diagnostic thermography in low back pain syndromes. *Clin J Pain* 1985;1:4-13.
34. Ishigaki T, Sakuma S, Okae S, et al. Infra-red imaging in syringomyelia: Correlation with the laterality of the syrinx. *Thermology* 1986;1:235-41.
35. Perelman RB, Adler D, Humphreys M. Comparison of lumbosacral thermograms with CT scans. In: Abernathy M, Uematsu S, editors. *Medical thermography*. Anover (PA): Sheridan Press, 1986:127-33.
36. Brown RKJ, Bassett LW, Exler CE, Gold RH. Thermography as a screening modality for nerve fiber irritation inpatients with low back pain. *Mod Med* 1987;55 Suppl:86-8.
37. Chafetz N, Wexler CE, Kaiser JA. Neuromuscular thermography of the lumbar spine with CT correlation. *Spine* 1988;13:922-5.
38. Mahomey GH, McCulloch J. Thermography in back pain. Thermography as a diagnostic aid in sciatica. *Thermology* 1985;1:43-50.
39. Mills GH, Davies GK, Getty CJM, Conway J. The evaluation of liquid crystal thermography in the investigation of nerve root compression due to lumbosacral lateral spinal stenosis. *Spine* 1986;11:427-32.
40. Green SA. Reply to Rothschild. *Adult Ortho* 1987;10:265.
41. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. *N Eng J Med* 1978;299:926-30.
42. Hubbard J, Maultsby J, Wexler CE. Lumbar and cervical thermography for nerve fiber impingement: A critical review. *Clin J Pain* 1986;2:131-7.

43. Wexler CE, Chafetz N. Cervical thoracic and lumbar thermography in the evaluation of symptomatic worker's compensation patients. *Mod Med* 1987;55 Suppl:53-7.
44. Cooke C, Menard MR, Beach GN, Locke SR, Hirsch GH. Serial lumbar dynamometry in low back pain. *Spine* 1992;17:653-62.
45. Dillard J, Trafimow J, Andersson GB, Cronin K. Motion of the lumbar spine: reliability of two measurement techniques. *Spine* 1991;16:321-4.
46. Levene JA, Seeds RH, Goldberg HM, Frazier M, Fuhrman GA. Trends in isodynamic and isometric trunk testing on the Isostation B200. *J Spinal Disord* 1989;2:20-35.
47. Gomez T, Beach G, Cooke C, Hrudehy W, Goyert P. Normative database for trunk range of motion, strength, velocity, and endurance with the Isostation B-200 lumbar dynamometer. *Spine* 1991;16:15-21.
48. Hirsch G, Beach G, Cooke C, Menard M, Locke S. Relationship between performance on lumbar dynamometry and Waddell score in a population with low-back pain. *Spine* 1991;16:1039-43.
49. Kroemer KH. An isoinertial technique to assess individual lifting capability. *Hum Factors* 1983;25:493-506.
50. Foreman TK, Baxter CE, Troup JDG. Ratings of acceptable load and maximal isometric lifting strengths: the effects of repetition. *Ergonomics* 1984;27:1283-8.
51. Freivalds A, Chaffin DB, Garg A, Lee KS. A dynamic biomechanical evaluation of lifting maximum acceptable loads. *J Biomech* 1984;17:251-62.
52. Mital A, Karwowski W, Mazouz AK, Orsarah E. Prediction of maximum acceptable weight of lift in the horizontal and vertical planes using simulated job dynamic strengths. *Am Ind Hyg Assoc J* 1986;47:288-92.
53. Marras WS, Rangarajulu SL, Wong Sam PE. Trunk force development during static and dynamic lifts. *Hum Factors* 1987;29:19-29.
54. Parnianpour M, Buchalter D, Nordin M, Kahanovitz N. A non-invasive in-vivo measurement technique for assesement of the six degrees of freedom of trunk motion. In: Lantz SA and King JA, editor. *Advances in bioengineering*. New York: American Society of Mechanical Engineers, 1986:28-9.
55. Seeds RH, Levene J, Goldberg HM. Normative data for IsoStation B100. *J Orthop Sports Phys Ther* 1987;9:141-55.

56. Seeds RH, Levene JA, Goldberg HM. Abnormal patient data for the IsoStation B100. *J Orthop Sports Phys Ther* 1987;9:121-33.
57. Parnianpour M, Nordin M, Kahanovitz N, Frankel V. The triaxial coupling of torque generation of trunk muscles during isometric exertions and the effect of fatiguing isoinertial movements on the motor output and movement patterns. *Spine* 1988;13:982-92.
58. Seeds RH, Levene J, Goldberg HM, Shirke AM. Electronic equipment provides accuracy in back injury analysis. *Occup Health Saf* 1988;57:38-42.
59. Spengler DM, Szpalski M. Newer assessment approaches for the patient with low back pain. *Contemporary Orthopaedics* 1990;21:371-8.
60. Frazer MB, Guy DP, Spengler DM. Lumbar spine flexion and extension torque in normal subjects: an objective study using a multiaxis testing device. Isotechnologies Inc. [res. abstracts]. Pebble Beach (CA): Assoc Bone Joint Surg, April 1988;2.
61. Gracovetsky S. The spinal engine. New York: Springer-Verlag, 1988.
62. Gracovetsky S, Newman N, Kary M. The principles of spinoscopy. Spinex Medical Technologies Inc., Document I.
63. Spinex. The Spinoscope (R): non-invasive imaging system for dynamic functional assessment of the spine. Spinex Medical Technologies Inc., Montréal, Canada.
64. Porter RW, Hibbert C, Evans C. The natural history of root entrapment syndrome. *Spine* 1984;9:418-21.
65. Swets JA. Measuring the accuracy of diagnostic systems. *Science* 1988;240:1285-93.
66. Metz CE. Basic principles of ROC analysis. *Seminars Nucl Med* 1978;8:283-98.
67. Hanley JA, MacNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982;143:29-36.
68. McNeil BJ, Hanley JA, Funkenstein HH, Wallman J. The use of paired ROC curves in studying the impact of history on radiographic interpretation: CT of the head as a case study. *Radiology* 1983;149:75-7.
69. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837-45.

70. Scott J, Huskisson EC. Graphic representation of pain. *Pain* 1976;2:175-84.
71. Fairbank JCT, Couper J, Davies JB, O'Brien JP. The Oswestry low back pain disability questionnaire. *Physiotherapy* 1980;66:271-3.
72. Roland M, Morris R. A study of the natural history of back pain. Part I: Development of a reliable and sensitive measure of disability in low-back pain. *Spine* 1983;8:141-4.
73. Moll JMH, Wright V. Normal range of spinal mobility. An objective clinical study. *Ann Rheum Dis* 1971;30:381-6.
74. Mayer TG, Gatchel RJ, Mayer H, Kishino ND, Keeley J, Mooney V. A prospective two-year study of functional restoration in industrial low back injury. *JAMA* 1987;258:1763-7.
75. Leavitt F. Pain and deception: use of verbal pain measurement as a diagnostic aid in differentiating between clinical and simulated low-back pain. *J Psychosomatic Res* 1985;29:495-505.
76. Leavitt F. Detection of simulation among persons instructed to exaggerate symptoms of low back pain. *J Occup Med* 1987;29:229-33.
77. Leavitt F. Predicting disability time using formal low back pain measurement: the low back pain simulation scale. *J Psychosomatic Res* 1991;35:599-607.
78. Russell ML, Ghee KL, Probstfield JL, Insull W Jr. Development of standardized simulated patients for quality control of the clinical interview. *Controlled Clin Trials* 1983;4:197-208.
79. Miller H, Cartlidge N. Simulation and malingering after injuries to the brain and spinal cord. *Lancet* 1972;1:580-5.
80. Jensen MC, Brant-Zawadzki MN, Obuchowski N, Modic MT, Malkasian D, Ross JS. Magnetic resonance imaging of the lumbar spine in people without back pain. *N Engl J Med* 1994;331:69-73.
81. Sheps SB, Schechter MT. The assessment of diagnostic tests: a survey of current medical research. *JAMA* 1984;252:2418-22.

APPENDICE 1

Figures et tableaux

LISTE DES FIGURES

Figure 1	Protocole : "Instruments diagnostiques".	A4
Figure 2	Indice COR des différentes modalités dans l'évaluation de la lombalgie pour l'ensemble des sujets	A5
Figure 3	Intervalle de confiance au niveau 95 % pour l'indice COR des différentes modalités dans l'évaluation de la lombalgie pour l'ensemble des sujets	A6
Figure 4	Intervalle de confiance au niveau 95 % pour l'indice COR des différentes modalités dans l'évaluation du rôle de simulateur/dissimulateur pour l'ensemble des sujets	A7
Figure 5A	Courbes COR de la moyenne de deux lecteurs pour les différentes modalités dans l'évaluation de la lombalgie pour les sujets honnêtes	A8
Figure 5B	Courbes COR de la moyenne de deux lecteurs pour les différentes modalités dans l'évaluation de la lombalgie pour les sujets malhonnêtes.	A9

Figure 1. Protocole : "Instruments diagnostiques"

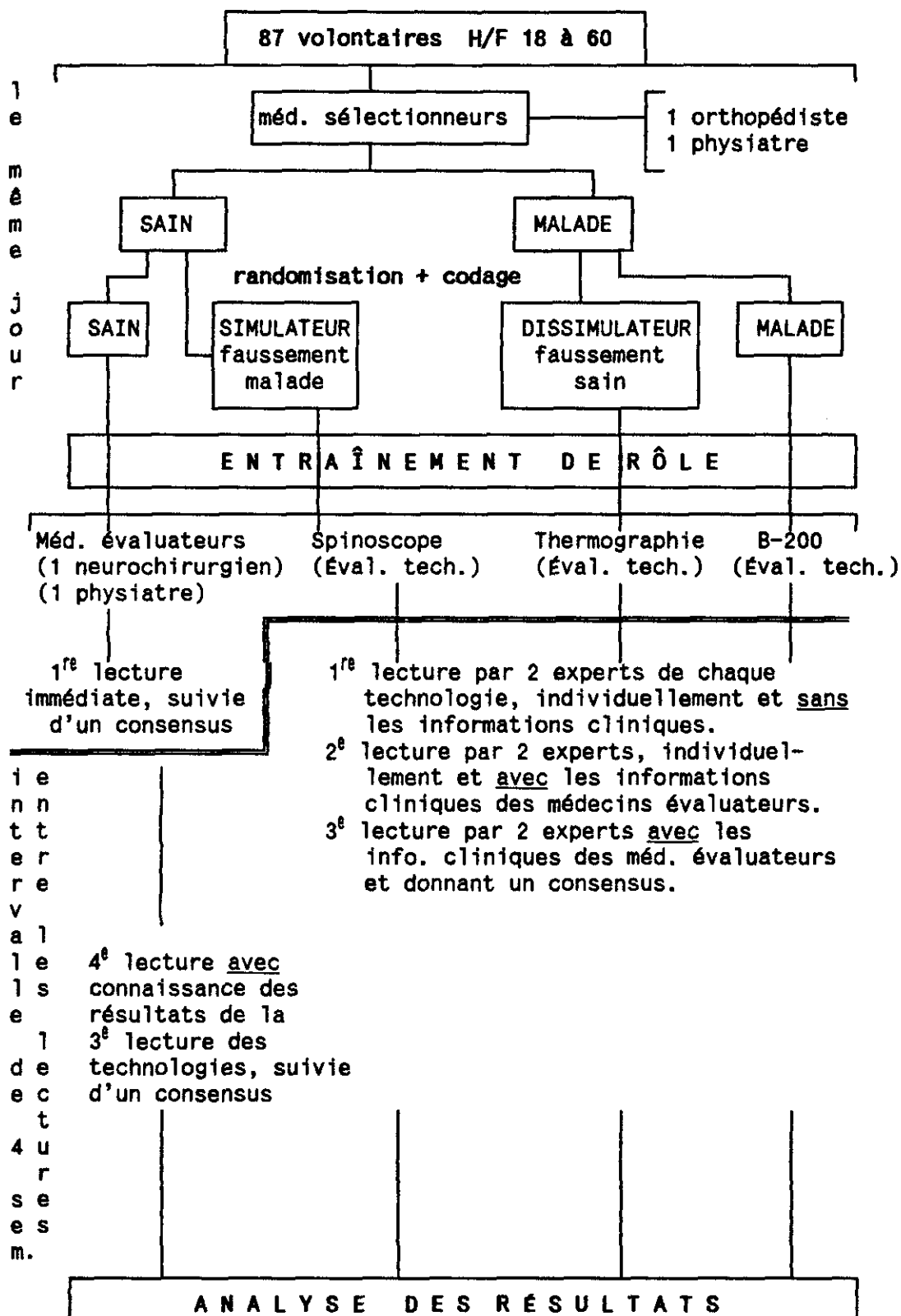
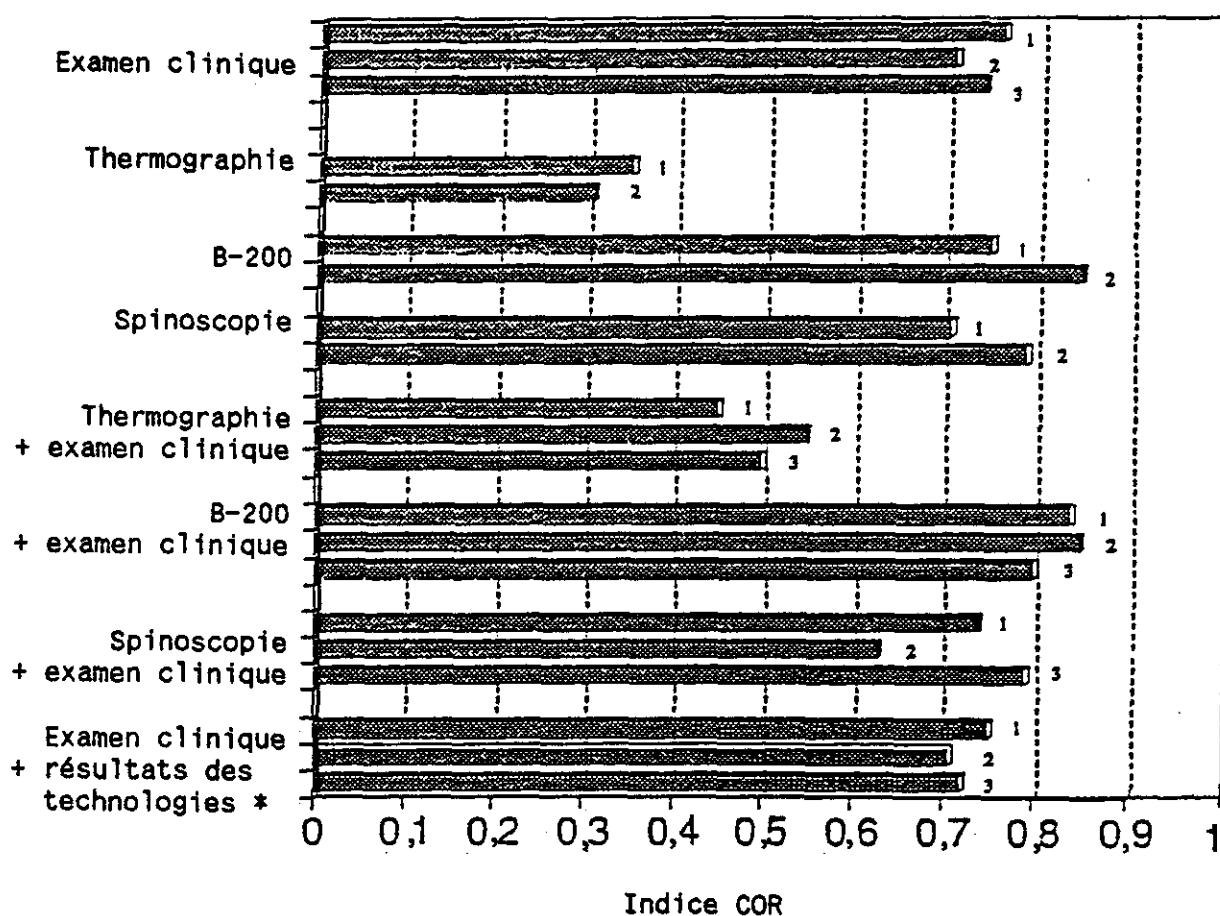


Figure 2. Indice COR des différentes modalités dans l'évaluation de la lombalgie pour l'ensemble des sujets



* Les résultats des technologies étaient disponibles pour le B-200 et la thermographie

1 = 1^{er} lecteur
 2 = 2^e lecteur
 3 = consensus

Figure 3. Intervalle de confiance au niveau 95 % pour l'indice COR des différentes modalités dans l'évaluation de la lombalgie pour l'ensemble des sujets

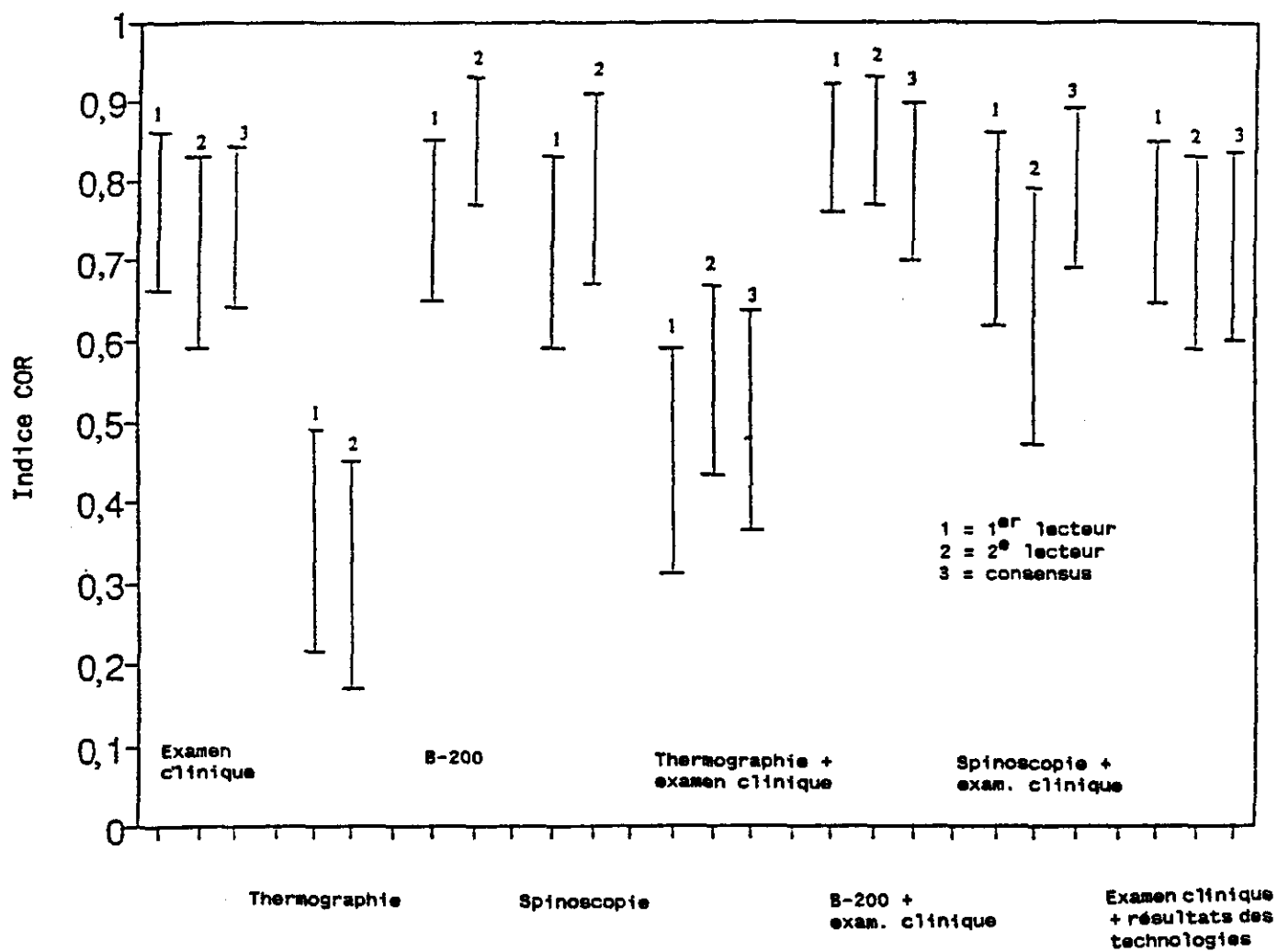


Figure 4. Intervalle de confiance au niveau 95 % pour l'indice COR des différentes modalités dans l'évaluation du rôle de simulateur/dissimulateur pour l'ensemble des sujets

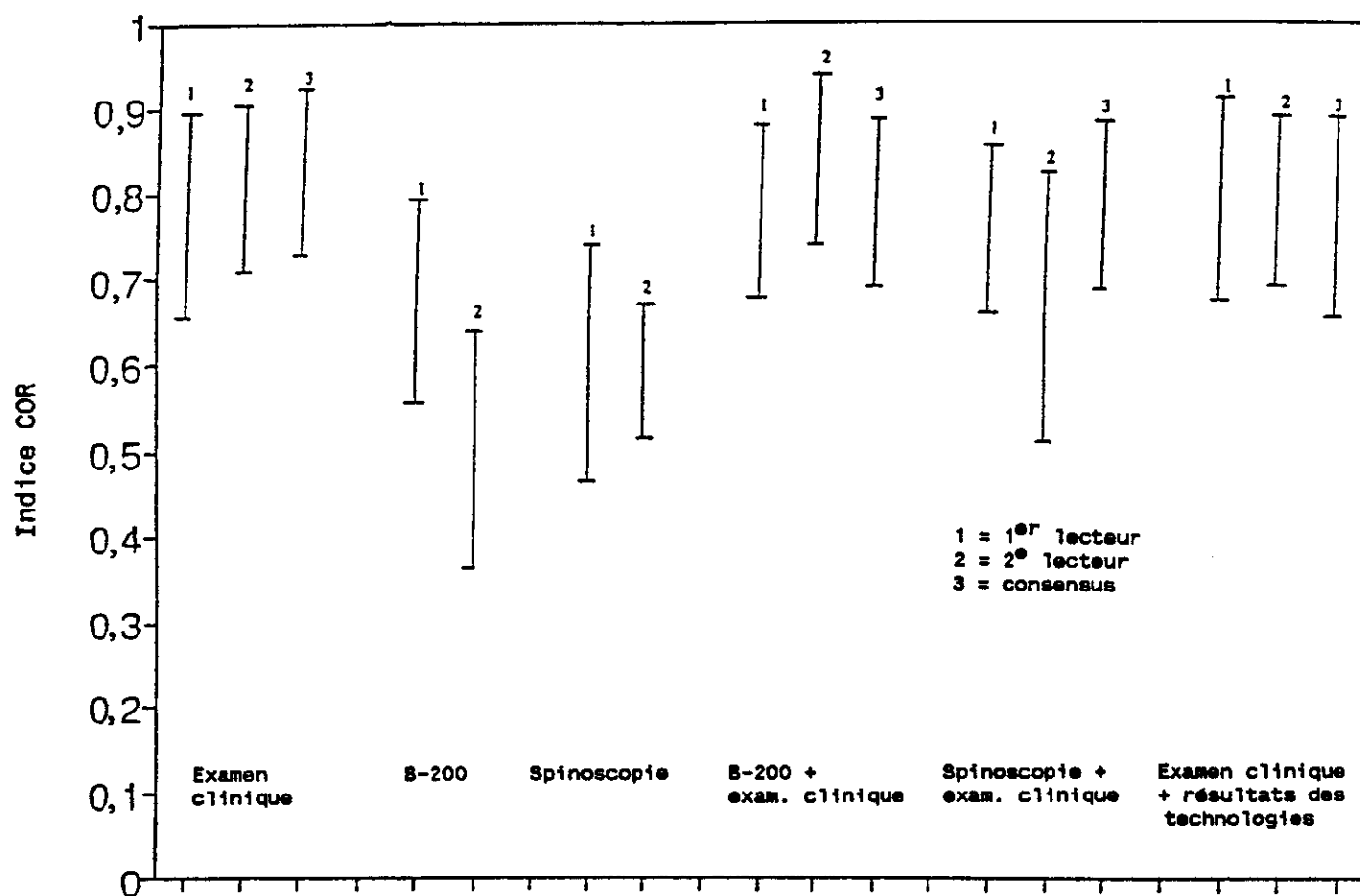


Figure 5A. Courbes COR de la moyenne de deux lecteurs pour les différentes modalités dans l'évaluation de la lombalgie pour les sujets honnêtes

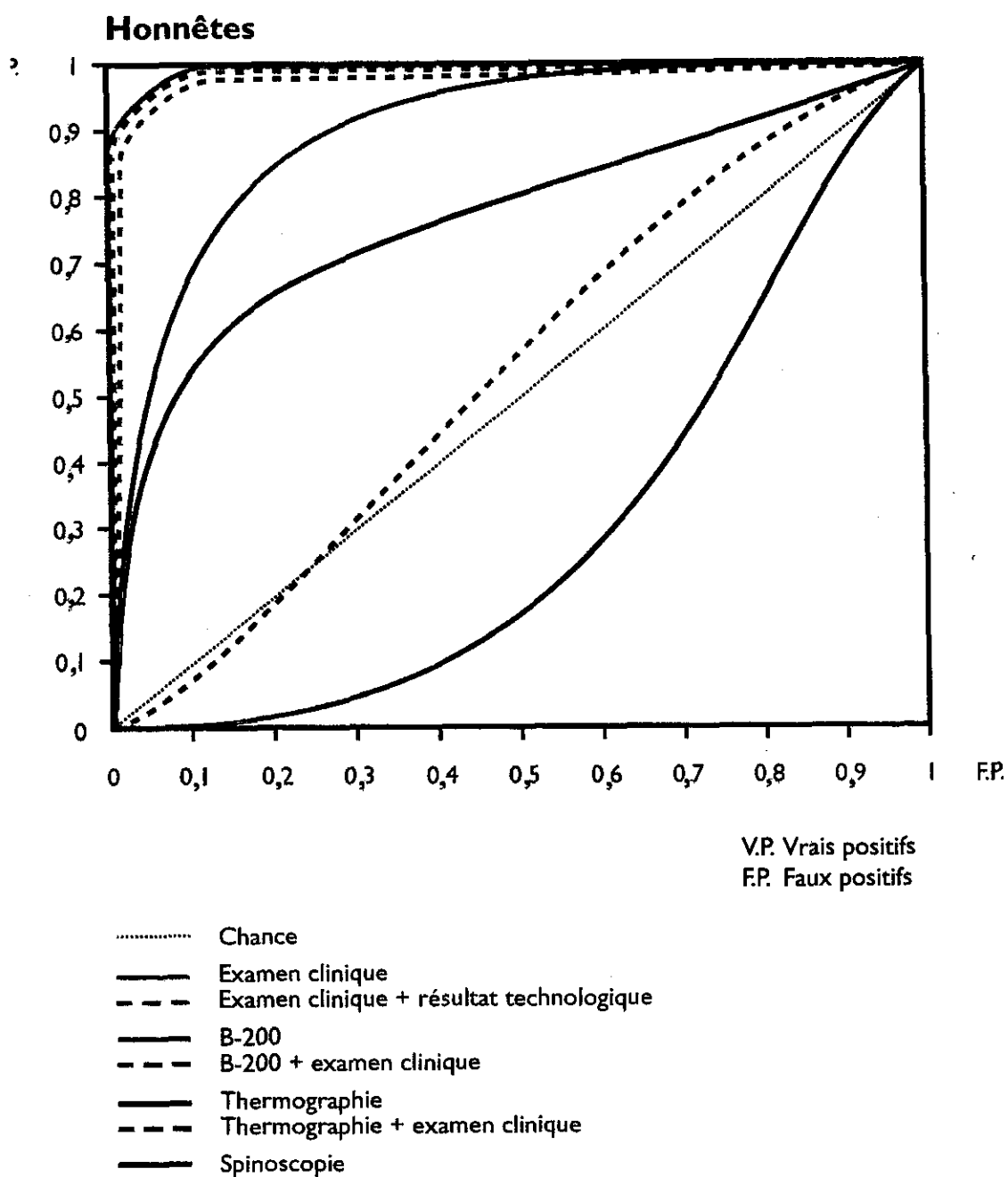
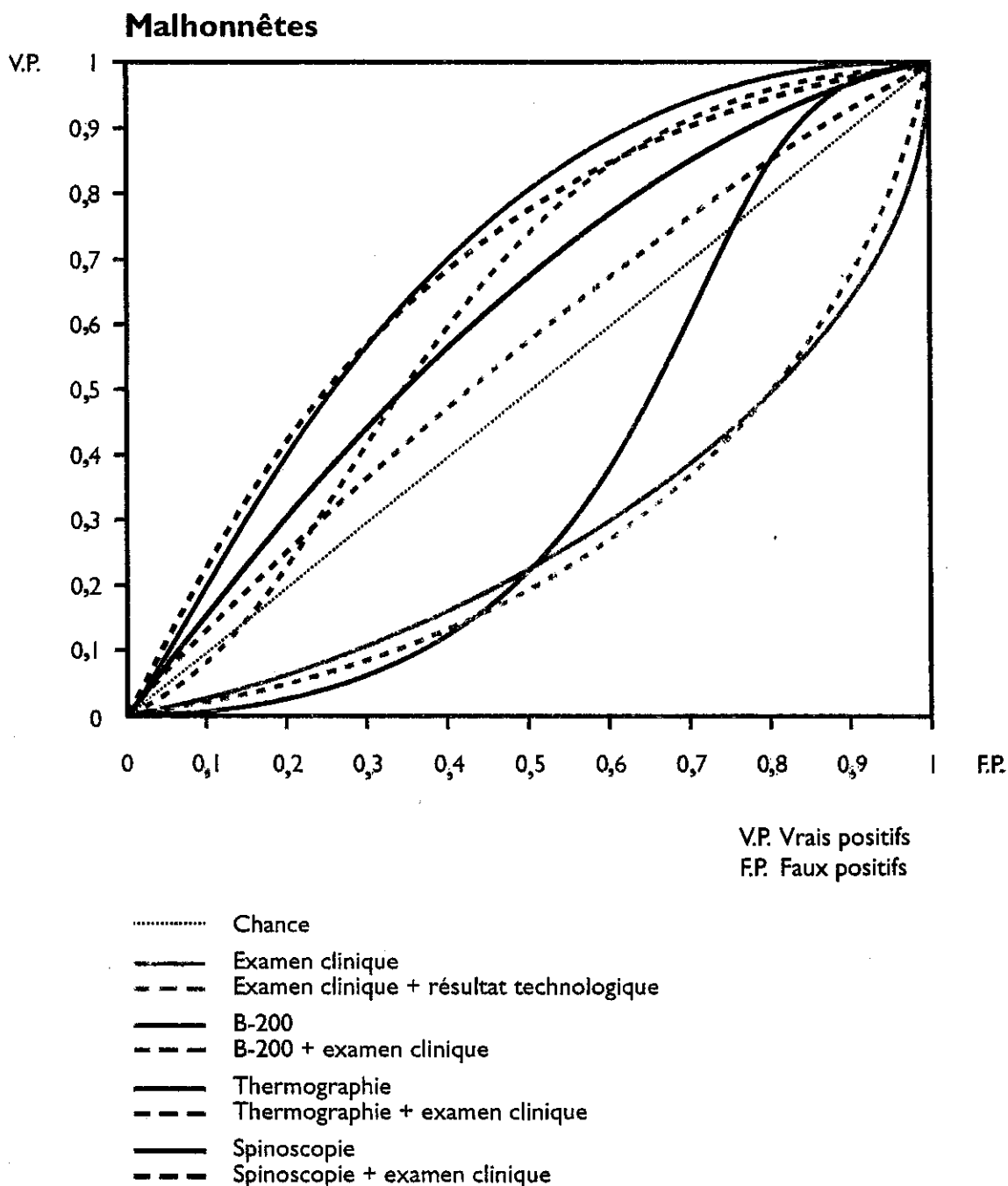


Figure 5B. Courbes COR de la moyenne de deux lecteurs pour les différentes modalités dans l'évaluation de la lombalgie pour les sujets malhonnêtes



LISTE DES TABLEAUX

Tableau 1	Description des sujets malades et sains	A13
Tableau 2	Évaluation de la lombalgie chez les sujets malhonnêtes (simulateurs et dissimulateurs) et honnêtes : indices COR pour les différentes modalités et selon les lecteurs	A14
Tableau 3	Évaluation de l'état de simulateur/dissimulateur chez les sujets sains et malades : indices COR pour les différentes modalités et selon les lecteurs	A15
Tableau 4	Évaluation de la lombalgie selon le sexe du sujet : indices COR pour les différentes modalités et selon les lecteurs	A16
Tableau 5	Évaluation de la lombalgie selon le groupe d'âge des sujets : indices COR pour les différentes modalités et selon les lecteurs	A17
Tableau 6	Évaluation de la lombalgie et de l'état de simulateur/dissimulateur : comparaison des indices COR pour les différentes modalités et selon les lecteurs	A18
Tableau 7	Exactitude simple : aptitude à distinguer les sujets sains et malades en rapport avec l'état réel, honnête ou malhonnête, en %	A19
Tableau 8	Exactitude simple : variations entre les lecteurs pour les sujets honnêtes et malhonnêtes, en %	A20
Tableau 9A	Exactitude double : aptitude à distinguer les sujets	A21
Tableau 9B	Exactitude double : aptitude à distinguer les sujets sains ou malades par la thermographie	A22
Tableau 9C	Exactitude double : aptitude à distinguer les sujets par le B-200	A23
Tableau 9D	Exactitude double : aptitude à distinguer les sujets par la spinoscopie	A24

LISTE DES TABLEAUX (suite)

Tableau 9E	Exactitude double : aptitude à distinguer les sujets, en % (sommaire)	A25
Tableau 9F	Exactitude double : aptitude à distinguer les sujets, performance chez les sujets honnêtes ou malhonnêtes, en % (sommaire)	A26
Tableau 9G	Exactitude double : aptitude à distinguer les sujets, performance chez les sujets sains ou malades, en % (sommaire).	A27
Tableau 10	Exactitude double : aptitude à distinguer les sujets quant à la sensibilité, la spécificité et la valeur globale pour les sujets honnêtes ou malhonnêtes, en %	A28
Tableau 11A	Résultats selon les méthodes d'analyse COR, d'exactitude simple et d'exactitude double dans l'évaluation de la lombalgie selon l'état réel pour les sujets honnêtes et malhonnêtes (moyenne des résultats obtenus pour chaque lecture)	A29
Tableau 11B	Résultats selon les méthodes d'analyse COR, d'exactitude simple et d'exactitude double dans l'évaluation de la lombalgie selon l'état réel pour l'ensemble des sujets honnêtes et malhonnêtes (moyenne des résultats obtenus pour chaque lecture)	A30

Tableau 1. Description des sujets malades et sains

Caractéristiques			État		Total
			Malades	Sains	
Nbre de sujets (%)			41 (47)	46 (53)	87
Âge	Moyenne		33	33	33
	Min.- Max.		18-57	19-55	18-57
	Écart type		9,8	9,0	9,4
Sexe					
Nbre de sujets (%)	Homme		21 (51,2)	23 (50)	44 (50,6)
Catégorie de travail					
Nbre de sujets (%)	Lourd		25 (61,0)	20 (43,5)	45 (51,7)
	Modéré		12 (29,3)	10 (21,7)	22 (24,3)
	Léger		4 (9,7)	10 (21,7)	14 (16,1)
	Sédentaire		0 (0)	6 (13,1)	6 (6,9)
	Total		41 (100)	46 (100)	87 (100)
Taille (cm)					
Hommes	Moyenne		173	173	173
	Min.- Max.		165-188	155-185	155-188
	Écart type		6,5	6,5	6,5
Femmes	Moyenne		159	160	160
	Min.- Max.		150-170	148-178	148-178
	Écart type		6,1	7,3	6,7
Total	Moyenne		166	167	166
	Min.- Max.		150-188	148-185	148-188
	Écart type		9,7	9,3	9,4
Poids (kg)					
Hommes	Moyenne		74	78	76
	Min.- Max.		59-93	63-95	59-95
	Écart type		11	9,9	10,5
Femmes	Moyenne		60	56	58
	Min.- Max.		45-78	40-79	40-79
	Écart type		9,4	8	8,7
Total	Moyenne		66	67	67
	Min.- Max.		45-93	40-95	40-95
	Écart type		12,4	14,1	13,3

Tableau 2. Évaluation de la lombalgie chez les sujets malhonnêtes (simulateurs et dissimulateurs) et honnêtes : indices COR pour les différentes modalités et selon les lecteurs

Modalités	Lecteur	Malhonnêtes n = 43		Honnêtes n = 44		Total n = 87	
		Indice	ET†	Indice	ET	Indice	ET
Examen clinique	1er	0,36	0,09	0,99†		0,76	0,05
	2e	0,24	0,08	0,99	0,02	0,71	0,06
	Consensus	0,36	0,08	0,99	0,01	0,74	0,05
Thermographie	1er	0,36	0,09	0,35	0,09	0,35	0,07
	2e	0,40	0,12	0,24	0,09	0,31	0,07
B-200	1er	0,64	0,08	0,89	0,05	0,75	0,05
	2e	0,80	0,07	0,92	0,04	0,85	0,04
Spinoscopie	1er	0,61	0,09	0,81	0,07	0,71	0,06
	2e	0,84	0,07	0,74	0,10	0,79	0,06
Thermographie + examen clinique	1er	0,65	0,09	0,48	0,09	0,45	0,07
	2e	0,40	0,10	0,55	0,10	0,55	0,06
	Consensus	0,48	0,10	0,53	0,09	0,50	0,07
B-200 + examen clinique	1er	0,65	0,09	0,98	0,02	0,84	0,04
	2e	0,60	0,10	0,99	0,02	0,85	0,04
	Consensus	0,59	0,09	0,97	0,02	0,80	0,05
Spinoscopie + examen clinique	1er	0,48	0,09	0,92†	0,04†	0,74	0,06
	2e	0,73	0,13	0,72†	0,07†	0,63	0,08
	Consensus	0,69	0,08	0,88	0,05	0,79	0,05
Examen clinique + résultats des technologies*	1er	0,33	0,08	0,99†		0,75	0,05
	2e	0,24	0,07	0,99	0,02	0,71	0,06
	Consensus	0,28	0,08	0,99	0,01	0,72	0,06

ET† = Erreur type

† = Indice estimé selon une approche non paramétrique⁶⁹

|| = Variance observée non calculable, effectif non suffisant

* = Les résultats des technologies étaient disponibles pour le B-200 et la thermographie.

Tableau 3. Évaluation de l'état de simulateur/dissimulateur chez les sujets sains et malades : indices COR pour les différentes modalités et selon les lecteurs

Modalités	Lecteur	Sains n = 46		Malades n = 41		Total n = 87	
		Indice	ET†	Indice	ET	Indice	ET
Examen clinique	1er	0,75	0,09	0,82	0,09	0,77	0,06
	2e	0,85	0,06	0,73	0,09	0,80	0,05
	Consensus	0,82	0,07	0,80	0,08	0,82	0,05
B-200	1er	0,64	0,09	0,73	0,09	0,67	0,06
	2e	0,49	0,09	0,53	0,11	0,50	0,07
Spinoscopie	1er	0,68	0,09	0,51	0,10	0,60	0,07
	2e	0,67†	0,06	0,50†	0,06	0,59†	0,04
B-200 + examen clinique	1er	0,79	0,07	0,79	0,08	0,78	0,05
	2e	0,81	0,07	0,87	0,07	0,84	0,05
	Consensus	0,78	0,07	0,80	0,07	0,79	0,05
Spinoscopie + examen clinique	1er	0,95	0,04	0,49	0,09	0,76	0,05
	2e	0,51	0,11	0,82	0,12	0,67	0,08
	Consensus	0,82	0,06	0,76	0,08	0,79	0,05
Examen clinique + résultats des technologies*	1er	0,76	0,09	0,79	0,10	0,80	0,06
	2e	0,82	0,07	0,76	0,09	0,80	0,05
	Consensus	0,78	0,08	0,80	0,08	0,78	0,06

ET† = Erreur type

† = Indice estimé selon une approche non paramétrique⁶⁹

* = Les résultats des technologies étaient disponibles pour le B-200 et la thermographie.

Tableau 4. Évaluation de la lombalgie selon le sexe du sujet :
indices COR pour les différentes modalités et selon les lecteurs

Modalités	Lecteur	Sexe				Total	
		Hommes		Femmes			
		Indice	ET†	Indice	ET	Indice	ET
Examen clinique	1er	0,84	0,06	0,69	0,08	0,76	0,05
	2e	0,84	0,06	0,58	0,09	0,71	0,06
	Consensus	0,84	0,06	0,63	0,09	0,74	0,05
Thermographie	1er	0,39	0,10	0,32	0,09	0,35	0,07
	2e	0,40	0,11	0,21	0,10	0,31	0,07
B-200	1er	0,85	0,06	0,68	0,08	0,75	0,05
	2e	0,88	0,06	0,80	0,07	0,85	0,04
Spinoscopie	1er	0,78	0,07	0,62	0,09	0,71	0,06
	2e	0,63†	0,08†	0,80	0,13	0,79	0,06
Thermographie + examen clinique	1er	0,59	0,09	0,51	0,09	0,45	0,07
	2e	0,58	0,11	0,44	0,10	0,55	0,06
	Consensus	0,50	0,10	0,49	0,09	0,50	0,07
B-200 + examen clinique	1er	0,88	0,05	0,81	0,08	0,84	0,04
	2e	0,88	0,05	0,80	0,07	0,85	0,04
	Consensus	0,87	0,06	0,71	0,08	0,80	0,05
Spinoscopie + examen clinique	1er	0,89	0,05	0,57	0,09	0,76	0,06
	2e	0,70†	0,07†	0,56	0,12	0,63	0,08
	Consensus	0,90	0,05	0,71	0,08	0,79	0,05
Examen clinique + résultats des technologies*	1er	0,84	0,06	0,65	0,09	0,75	0,05
	2e	0,82	0,06	0,59	0,09	0,71	0,06
	Consensus	0,81	0,07	0,62	0,09	0,72	0,06

ET† = Erreur type

† = Indice estimé selon une approche non paramétrique⁶⁹

* = Les résultats des technologies étaient disponibles
pour le B-200 et la thermographie.

Tableau 5. Évaluation de la lombalgie selon le groupe d'âge des sujets :
indices COR pour les différentes modalités et selon les lecteurs

Modalités	Lecteur	Groupe d'âge				Total	
		<=30 ans		>30 ans			
		Indice	ET†	Indice	ET	Indice	ET
Examen clinique	1er	0,77	0,08	0,75	0,08	0,76	0,05
	2e	0,77	0,08	0,67	0,08	0,71	0,06
	Consensus	0,78	0,07	0,71	0,08	0,74	0,05
Thermographie	1er	0,29	0,09	0,41	0,09	0,35	0,07
	2e	0,40	0,12	0,25	0,09	0,31	0,07
B-200	1er	0,87	0,06	0,64	0,08	0,75	0,05
	2e	0,91	0,05	0,80	0,07	0,85	0,04
Spinoscopie	1er	0,73	0,08	0,70	0,08	0,71	0,06
	2e	0,78	0,11	0,79	0,08	0,79	0,06
Thermographie + examen clinique	1er	0,57	0,09	0,53	0,09	0,45	0,07
	2e	0,37	0,10	0,60	0,09	0,55	0,06
	Consensus	0,54	0,10	0,47	0,09	0,50	0,07
B-200 + examen clinique	1er	0,92	0,05	0,78	0,08	0,84	0,04
	2e	0,91	0,05	0,77	0,07	0,85	0,04
	Consensus	0,90	0,05	0,70	0,08	0,80	0,05
Spinoscopie + examen clinique	1er	0,75	0,08	0,74	0,08	0,76	0,06
	2e	0,51	0,11	0,78	0,11	0,63	0,08
	Consensus	0,76	0,08	0,83	0,06	0,79	0,05
Examen clinique + résultats des technologies*	1er	0,75	0,08	0,75	0,08	0,75	0,05
	2e	0,77	0,08	0,66	0,08	0,71	0,06
	Consensus	0,78	0,07	0,66	0,09	0,72	0,06

ET† = Erreur type

* = Les résultats des technologies étaient disponibles
pour le B-200 et la thermographie.

Tableau 6. Évaluation de la lombalgie et de l'état de simulateur/dissimulateur : comparaison des indices COR pour les différentes modalités et selon les lecteurs

Contraste	évaluation de la lombalgie				évaluation du rôle de simulateur			
	Malhonnêtes		Honnêtes		Sains		Malades	
	Diff. Indice	P	Diff. Indice	P	Diff. Indice	P	Diff. Indice	P
Exam.clinique vs thermographie								
1er L	0,00	0,96	0,61†	0,00†	N.A.	N.A.	N.A.	N.A.
2e L	-0,15	0,31	0,75	0,00†	N.A.	N.A.	N.A.	N.A.
Exam. clinique vs B-200								
1er L	-0,28	0,02†	0,13†	0,01†	0,11	0,37	0,09	0,47
2e L	-0,56	0,00†	0,07	0,14	0,36	0,00†	0,20	0,19
Exam. clinique vs spinoscopie								
1er L	-0,25	0,06	0,21†	0,00†	0,07	0,41	0,31	0,01†
2e L	-0,60	0,00†	0,25	0,02†	0,14†	0,15	0,22†	0,04†
B-200 vs spinoscopie								
1er L	0,03	0,80	0,08	0,32	-0,04	0,71	0,22	0,09
2e L	-0,04	0,81	0,18	0,12	-0,18	0,07	0,03	0,79
Thermographie + exam. clinique vs thermographie								
1er L	0,30	0,01†	0,13	0,27	N.A.	N.A.	N.A.	N.A.
2e L	0,00	0,95	0,32	0,01†	N.A.	N.A.	N.A.	N.A.
Thermographie + exam. clinique vs exam. clinique								
1er L	0,29	0,01†	-0,53†	0,00†	N.A.	N.A.	N.A.	N.A.
2e L	0,16	0,19	-0,44	0,00†	N.A.	N.A.	N.A.	N.A.
B-200 + exam. clinique vs B-200								
1er L	0,01	0,94	0,08	0,08	0,15	0,16	0,06	0,43
2e L	-0,20	0,01†	0,07	0,14	0,32	0,00†	0,34	0,01†
B-200 + exam. clinique vs exam. clinique								
1er L	0,29	0,03†	-0,02†	0,17	0,04	0,73	-0,03	0,74
2e L	0,35	0,00†	0,00	0,89	-0,04	0,48	0,14	0,12
Spinoscopie + exam. clinique vs spinoscopie								
1er L	-0,13	0,17	0,13†	0,12	0,27	0,00†	-0,02	0,88
2e L	-0,11	0,32	0,14†	0,12	-0,13†	0,07	0,31†	0,00†
Spinoscopie + exam. clinique vs exam. clinique								
1er L	0,12	0,30	-0,08†	0,03†	0,20	0,02†	-0,33	0,00†
2e L	0,49	0,00†	-0,27†	0,00†	-0,34	0,00†	0,09	0,55
B-200 + exam. clinique vs spinoscopie + exam.clinique								
1er L	0,16	0,22	0,05†	0,24	-0,16	0,03†	0,30	0,01†
2e L	-0,13	0,48	0,27†	0,00†	0,30	0,01†	0,05	0,72
Exam.clinique + résultats des technologies vs exam. clinique								
1er L	-0,03	0,57	0,00†	0,42	0,00†	0,95	-0,03	0,77
2e L	0,00†	0,97	0,00	0,83	-0,03†	0,42	0,03	0,46

L = Lecteur

N.A. = Non applicable, car l'évaluation des sujets malhonnêtes par la thermographie n'était pas utilisée.

† = Un indice estimé différent et une incertitude correspondante étaient utilisés selon une approche non paramétrique⁶⁹.

‡ = Statistique significative ($p < 0,05$).

Note : Les valeurs de p sont présentées à titre indicatif seulement, compte tenu de la multiplicité des comparaisons.

Tableau 7. Exactitude simple : aptitude à distinguer les sujets sains et malades* en rapport avec l'état réel, honnête ou malhonnête, en %

Indécis = mauvaise réponse

Modalités	Sensibilité			Spécificité			Valeur globale		
	H+M n = 41	H n = 21	M n = 20	H+M n = 46	H n = 23	M n = 23	H+M n = 87	H n = 44	M n = 43
Ex. clin.			Dias.			Simul.			
1er lecteur	68,3**	95,2	40	67,4	95,7	34,8	67,8	95,5	37,2
2e lecteur	65,9	95,2	35	54,3	87,0	21,7	59,8	90,9	27,9
Consensus	61,0	90,5	30	54,3	87,0	21,7	57,5	88,6	25,6
B-200 (1)									
1re lecture	70,7	76,2	65	65,2	87,0	43,5	67,8	81,8	53,5
2e lecture (i.c.)	78	85,7	70	76,1	95,7	56,5	77	90,9	62,8
B-200 (2)									
1re lecture	73,2	76,2	70	73,9	82,6	65,2	73,6	79,5	67,4
2e lecture (i.c.)	78,1	90,5	65	63,0	95,7	30,4	70,1	93,2	46,5
B-200 (c)	85,4	95,2	75	71,7	87,0	56,5	78,2	90,9	65,1
Spinoscopie (1)									
1re lecture	41,5	47,6	35	71,7	87,0	56,5	57,5	68,2	46,5
2e lecture (i.c.)	29,3	42,9	15	78,3	100	56,5	55,2	72,7	37,2
Spinoscopie (2)									
1re lecture	82,9	81,0	85	47,8	34,8	60,9	64,4	56,8	72,1
2e lecture (i.c.)	92,7	100	85	47,8	56,5	39,1	69,0	77,3	60,5
Spinoscopie (c)	68,3	71,4	65	69,6	78,3	60,9	69,0	75,0	62,8
Thermographie (1)									
1re lecture	36,6	33,3	40	37,0	39,1	34,8	36,8	36,4	37,2
2e lecture (i.c.)	53,7	52,4	55	45,7	39,1	52,2	49,4	45,5	53,5
Thermographie (2)									
1re lecture	75,6	66,7	85	6,5	4,3	8,7	39,0	34,1	44,2
2e lecture (i.c.)	43,9	52,4	35	10,9	13	8,7	26,4	31,8	20,9
Thermographie (c)	68,3	71,4	61,9	19,6	21,7	17,4	42,5	45,5	38,6
Ex. clin. + R.T.									
1er lecteur	68,3	95,2	40	56,5	87,0	26	62,1	90,9	32,6
2e lecteur	63,4	95,2	30	54,3	87,0	21,7	58,6	90,9	25,6
Consensus	61,0	90,5	30	54,3	87,0	21,7	57,5	88,6	25,6

i.c. = avec les informations cliniques

(c) = consensus

R.T. = résultats du B-200 + thermographie

(1) = 1^{er} lecteur

H+M = honnêtes + malhonnêtes

H = honnêtes

M = malhonnêtes

(2) = 2^e lecteur

* basée sur l'état de santé seul

** Exemple = 28 sujets malades sur 41 interprétés
correctement = 28/41 = sensibilité
de 68,2 %

Tableau 8. Exactitude simple : variations entre les lecteurs pour les sujets honnêtes et malhonnêtes, en %

Indécis = mauvaise réponse

	Honnêtes			Malhonnêtes		
	Se	Sp	V.G.	Se	Sp	V.G.
Examen clinique seul	95	87-95	95-90	40-35	34-21	37-27
consensus	90	87	88	30	21	25
consensus + R.T.	90	87	88	30	21	25
B-200 seul	76	87-82	81-79	65-70	43-65	53-67
+ i.c.	85-90	95-95	90-93	70-65	56-30	62-46
consensus	95	87	90	75	56	65
Spinoscopie seule	47-81	86-34	68-56	35-85	56-60	46-72
+ i.c.	42-100	100-56	72-77	15-85	56-39	37-60
consensus	71	78	75	65	60	62

i.c. = informations cliniques

Se = sensibilité

Sp = spécificité

V.G. = valeur globale

R.T. = résultats du B-200 + thermographie

* Pour les lignes sur lesquelles deux valeurs apparaissent, la 1^{re} correspond à la performance du 1^{er} évaluateur et la 2^e à la performance du 2^e évaluateur.

Note : décimales enlevées pour fins de présentation.

Tableau 9A. Exactitude double : aptitude à distinguer les sujets

Indécis = mauvaise réponse

Examen clinique 1re lecture			2e lecteur		Consensus	
1er lecteur		Groupes de sujets	n	%	n	%
n	%					
22/23 = 95,7		Sain honnête	20/23 = 87,0		20/23 = 87,0	
8/23 = 34,8		Sain malhonnête	5/23 = 21,7		5/23 = 21,7	
20/21 = 95,2		Malade honnête	19/21 = 90,5		19/21 = 90,5	
6/20 = 30		Malade malhonnête	4/20 = 20		5/20 = 25	
56/87 = 64,4		Tous	48/87 = 55,2		49/87 = 56,3	
42/44 = 95,5		Honnête	39/44 = 88,6		39/44 = 88,6	
14/43 = 32,6		Malhonnête	9/43 = 20,9		10/43 = 23,2	
30/46 = 65,2		Sain	25/46 = 54,3		25/46 = 54,3	
26/41 = 63,4		Malade	23/41 = 56,1		24/41 = 58,5	

Examen clinique 4e lecture (nouvelles lectures avec résultats B-200 + thermographie)						
1er lecteur		Groupes de sujets	2e lecteur		Consensus	
n	%		n	%	n	%
20/23 = 87,0		Sain honnête	20/23 = 87,0		20/23 = 87,0	
6/23 = 26,0		Sain malhonnête	5/23 = 21,7		5/23 = 21,7	
20/21 = 95,2		Malade honnête	20/21 = 95,2		19/21 = 90,5	
7/20 = 27,3		Malade malhonnête	5/20 = 25		4/20 = 20	
53/87 = 60,9		Tous	50/87 = 57,5		48/87 = 55,2	
40/44 = 90,9		Honnête	40/44 = 90,9		39/44 = 88,6	
13/43 = 30,2		Malhonnête	10/43 = 23,3		9/43 = 20,9	
26/46 = 56,5		Sain	25/46 = 54,3		25/46 = 54,3	
27/41 = 65,9		Malade	25/41 = 61,0		23/41 = 56,1	

Perf. = Performance

n = nombre

% = pourcentage

Tableau 98. Exactitude double : aptitude à distinguer les sujets sains ou malades par la thermographie

Indécis = mauvaise réponse

Thermographie 1re lecture			2e lecteur	
1er lecteur			n	%
n	%			
17/46 = 37,0		Sain	3/46 = 6,5	
N.A.		Sain	N.A.	
15/41 = 36,6		Malade	31/41 = 75,6	
N.A.		Malade	N.A.	
32/87 = 36,8		Tous les sujets	34/87 = 39,1	

Thermographie seule 2e lecture (nouvelle lecture)				Consensus
1er lecteur			2e lecteur	
n	%		n	%
21/46 = 45,7		Sain	5/46 = 10,9	9/46 = 19,6
N.A.		Sain	N.A.	N.A.
22/41 = 53,7		Malade	18/41 = 43,9	28/41 = 68,3
N.A.		Malade	N.A.	N.A.
43/87 = 49,4		Tous les sujets	23/87 = 26,4	39/87 = 44,8

N.A. = non applicable, les sujets ne pouvant simuler ou dissimuler

n = nombre

% = pourcentage

Tableau 9C. Exactitude double : aptitude à distinguer les sujets par le B-200

Indécis = mauvaise réponse

B-200 seul 1re lecture		Groupes de sujets	2e lecteur	
1er lecteur			n	%
n	%		n	%
15/23 = 65,2		Sain honnête	17/23 = 73,9	
3/23 = 13,0		Sain malhonnête	1/23 = 4,3	
13/21 = 61,9		Malade honnête	15/21 = 71,4	
4/20 = 20		Malade malhonnête	0/20 = 0	
35/87 = 40,2		Tous	33/87 = 37,9	
28/44 = 63,6		Honnête	32/44 = 72,7	
7/43 = 16,3		Malhonnête	1/43 = 2,3	
18/46 = 39,1		Sain	18/46 = 39,1	
17/41 = 41,5		Malade	15/41 = 36,6	

B-200 2e lecture (nouvelle avec les informations cliniques)					
1er lecteur		Groupes de sujets	2e lecteur		Consensus
n	%		n	%	
21/23 = 91,3		Sain honnête	21/23 = 91,3		18/23 = 78,3
11/23 = 47,8		Sain malhonnête	6/23 = 26,1		12/23 = 52,2
16/21 = 76,2		Malade honnête	19/21 = 90,5		19/21 = 90,5
11/20 = 55		Malade malhonnête	12/20 = 60		11/20 = 55
59/87 = 67,8		Tous	58/87 = 66,7		60/87 = 69,0
37/44 = 84,1		Honnête	40/44 = 90,9		37/44 = 84,0
22/43 = 51,1		Malhonnête	18/43 = 41,9		23/43 = 53,5
32/46 = 69,6		Sain	27/46 = 58,6		30/46 = 65,2
27/41 = 65,9		Malade	31/41 = 75,6		30/41 = 73,2

n = nombre
% = pourcentage

Tableau 9D. Exactitude double : aptitude à distinguer les sujets par la spinoscopie

Indécis = mauvaise réponse

Spinoscopie seule 1re lecture			2e lecteur	
1er lecteur		Groupes de sujets	n	%
n	%			
19/23 = 82,6		Sain honnête	6/23 = 26,1	
5/23 = 21,7		Sain malhonnête	7/23 = 30,4	
10/21 = 47,6		Malade honnête	17/21 = 81,0	
0/20 = 0		Malade malhonnête	3/20 = 15	
34/87 = 39		Tous	33/87 = 37,9	
29/44 = 65,9		Honnête	23/44 = 52,3	
5/43 = 11,6		Malhonnête	10/43 = 23,3	
24/46 = 52,2		Sain	13/46 = 28,3	
10/41 = 24,4		Malade	20/41 = 48,8	

Spinoscopie 2e lecture (avec les informations cliniques)					
1er lecteur		Groupes de sujets	2e lecteur		Consensus
n	%		n	%	n
21/23 = 91,3		Sain honnête	13/23 = 56,5		17/23 = 73,9
10/23 = 43,5		Sain malhonnête	8/23 = 34,8		12/23 = 52,2
9/21 = 42,9		Malade honnête	20/21 = 95,2		15/21 = 71,4
3/20 = 15		Malade malhonnête	13/20 = 65		9/20 = 45
42/87 = 48,3		Tous	54/87 = 62,1		53/87 = 60,9
29/44 = 65,9		Honnête	33/44 = 75,0		32/44 = 72,7
13/43 = 30,2		Malhonnête	21/43 = 48,8		21/43 = 48,8
30/46 = 65,2		Sain	21/46 = 45,7		29/46 = 63
12/41 = 29,3		Malade	33/41 = 80,5		24/41 = 58,5

n = nombre
% = pourcentage

Tableau 9E. Exactitude double : aptitude à distinguer les sujets,
en % (sommaire)

Indécis = mauvaise réponse

%	Valeur globale
69,0	B-200 2e lecture, consensus
67,8	B-200 2e lecture, 1er lecteur
66,7	B-200 2e lecture, 2e lecteur
64,4	Examen clinique 1re lecture, 1er lecteur
62,1	Spinoscopie 2e lecture, 2e lecteur
60,9	Spinoscopie 2e lecture, consensus
60,9	Examen clinique 4e lecture, 1er lecteur
57,5	Examen clinique 4e lecture, 2e lecteur
56,3	Examen clinique 1re lecture, consensus
55,2	Examen clinique 1re lecture, 2e lecteur
55,2	Examen clinique 4e lecture, consensus

% = pourcentage

Tableau 9F. Exactitude double : aptitude à distinguer les sujets, performance chez les sujets honnêtes ou malhonnêtes, en % (sommaire)

Indécis = mauvaise réponse

Performance	Sujets	
	Honnêtes %	Malhonnêtes %
Ex. clin. (1), 1re lecture	95,5	32,6
Ex. clin. (2), 4e lecture	90,9	23,3
Ex. clin. (1), 4e lecture	90,9	30,2
Ex. clin. (2), 1re lecture	88,6	20,9
Ex. clin. (c), 1re lecture	88,6	23,2
Ex. clin. (c), 4e lecture	88,6	20,9
B-200 (2), 2e lecture	90,9	41,9
B-200 (c), 2e lecture	84,0	53,5
B-200 (1), 2e lecture	84,1	51,1
Spinoscopie (2), 2e lecture	75,0	48,8
B-200 (2), 1re lecture	72,7	2,3
Spinoscopie (c), 2e lecture	72,7	48,8
Spinoscopie (1), 1re lecture	65,9	11,6
Spinoscopie (1), 2e lecture	65,9	30,2
B-200 (1), 1re lecture	63,6	18,6
Spinoscopie (2), 1re lecture	52,3	23,3

Ex. clin. = examen clinique

(c) = consensus

(1) = 1^{er} lecteur

(2) = 2^e lecteur

% = pourcentage

Tableau 9G. Exactitude double : aptitude à distinguer les sujets, performance chez les sujets sains ou malades, en % (sommaire)

Indécis = mauvaise réponse

Performance	Sujets	
	Sains %	Malades %
B-200 (1), 2e lecture	69,6	65,9
B-200 (c), 2e lecture	65,2	73,2
Ex. clin. (1), 1re lecture	65,2	63,4
Spinoscopie (1), 2e lecture	65,2	29,3
Spinoscopie (c), 2e lecture	63	58,5
B-200 (2), 2e lecture	56,5	75,6
Ex. clin. (1), 4e lecture	56,5	65,9
Ex. clin. (2), 1re lecture	54,3	56,1
Ex. clin. (c), 1re lecture	54,3	58,5
Ex. clin. (2), 4e lecture	54,3	61,0
Ex. clin. (c), 4e lecture	54,3	56,1
Spinoscopie (1), 1re lecture	52,2	24,4
Spinoscopie (2), 2e lecture	45,7	80,5

Ex. clin. = examen clinique

(c) = consensus

(1) = 1^{er} lecteur

(2) = 2^e lecteur

% = pourcentage

Tableau 10. Exactitude double : aptitude à distinguer les sujets*
quant à la sensibilité, la spécificité et la valeur globale
pour les sujets honnêtes ou malhonnêtes, en %

Indécis = mauvaise réponse

Modalités	Sensibilité			Spécificité			Valeur globale		
	H+M n = 41	H n = 21	M n = 20	H+M n = 46	H n = 23	M n = 23	H+M n = 87	H n = 44	M n = 43
Ex. clin.			Diss.			Simul.			
1er lecteur	63,4	95,2	30	65,2	95,7	34,8	64,4	95,5	32,6
2e lecteur	56,1	90,5	20	54,3	87,0	21,7	55,2	88,6	20,9
consensus	58,5	90,5	25	54,3	87,0	21,7	56,3	88,6	23,2
B-200 (1)									
1re lecture	41,5	61,9	20	41,3	65,2	13,0	40,2	63,6	16,3
2e lecture (i.c.)	65,9	76,2	55	69,6	89,3	47,8	67,8	84,1	51,1
B-200 (2)									
1re lecture	36,6	71,4	0	39,1	73,9	4,3	37,9	72,7	2,3
2e lecture (i.c.)	75,6	90,5	60	58,6	91,3	26,1	66,7	90,9	41,9
B-200 (c)	73,2	90,5	55	65,2	78,3	52,2	69,0	84,0	53,5
Spinoscopie (1)									
1re lecture	24,4	47,6	0	52,2	82,6	21,7	39	65,9	11,6
2e lecture (i.c.)	29,3	42,9	15	65,2	91,3	43,5	48,3	65,9	30,2
Spinoscopie (2)									
1re lecture	48,8	81,0	15	28,3	26,1	30,4	37,9	52,3	23,3
2e lecture (i.c.)	80,5	95,2	65	45,7	56,5	34,8	62,1	75,0	48,8
Spinoscopie (c)	58,5	71,4	45	63	73,9	52,2	60,9	72,7	48,8
Thermographie (1)									
1re lecture	36,6	N.A.	N.A.	37,0	N.A.	N.A.	36,8	N.A.	N.A.
2e lecture (i.c.)	53,7	N.A.	N.A.	45,7	N.A.	N.A.	49,4	N.A.	N.A.
Thermographie (2)									
1re lecture	75,6	N.A.	N.A.	6,5	N.A.	N.A.	39,1	N.A.	N.A.
2e lecture (i.c.)	43,9	N.A.	N.A.	10,9	N.A.	N.A.	26,4	N.A.	N.A.
Thermographie (c)	68,3	N.A.	N.A.	19,6	N.A.	N.A.	44,8	N.A.	N.A.
Ex. clin. + R.T.									
1er lecteur	65,9	95,2	27,3	56,5	87,0	26,0	60,9	90,9	30,2
2e lecteur	61,0	95,2	25	54,3	87,0	21,7	57,5	90,9	23,3
Consensus	56,1	90,5	20	54,3	87,0	21,7	55,2	88,6	20,9

i.c. = avec les informations cliniques
(c) = consensus
R.T. = résultats du B-200 + thermographie
(1) = 1^{er} lecteur

H+M = honnêtes + malhonnêtes
H = honnêtes
M = malhonnêtes
(2) = 2^e lecteur

* basée sur l'état de santé et le rôle

Tableau 11A. Résultats selon les méthodes d'analyse COR, d'exactitude simple et d'exactitude double dans l'évaluation de la lombalgie selon l'état réel pour les sujets honnêtes et malhonnêtes (moyenne des résultats obtenus pour chaque lecture)

Sujets	COR* Indice	Ex.S.** % V.G.+	Ex.D.*** % V.G.
Honnêtes :			
Examen clinique	0,99	93,2	92,1
Examen clinique + résult. technol.++	0,99	90,9	90,9
B-200 + examen clinique	0,99	92,0	87,5
B-200	0,91	80,7	68,1
Spinoscopie + examen clinique	0,82	75,0	70,5
Spinoscopie	0,78	62,5	59,1
Thermographie + examen clinique	0,52	37,0	N.A.
Thermographie	0,30	35,3	N.A.
Malhonnêtes :			
B-200	0,72	60,5	9,3
Spinoscopie	0,73	59,3	17,5
B-200 + examen clinique	0,63	54,7	46,5
Spinoscopie + examen clinique	0,61	48,9	39,5
Thermographie + examen clinique	0,53	37,2	N.A.
Thermographie	0,38	40,7	N.A.
Examen clinique	0,30	32,6	26,8
Examen clinique + résult. technol.++	0,29	29,1	26,8

* COR = Indice de la courbe opérationnelle des résultats (total des simulateurs^{60,69})

** Ex.S. = % Valeur globale, exactitude simple

*** Ex.D. = % Valeur globale, exactitude double

+ V.G. = Valeur globale

++ = Résultats du B-200 + thermographie

Indécis classifiés comme
mauvaises réponses

Tableau 11B. Résultats selon les méthodes d'analyse COR, d'exactitude simple et d'exactitude double dans l'évaluation de la lombalgie selon l'état réel pour l'ensemble des sujets honnêtes et malhonnêtes (moyenne des résultats obtenus pour chaque lecture)

Sujets	COR* Indice	Ex.S.** % V.G.+	Ex.D.*** % V.G.
Honnêtes + malhonnêtes :			
B-200 + examen clinique	0,85	73,6	67,3
B-200	0,80	70,7	39,1
Spinoscopie	0,75	61,0	38,5
Examen clinique	0,74	63,8	59,8
Examen clinique + résult. technol.++	0,73	60,3	58,7
Spinoscopie + examen clinique	0,69	62,1	55,1
Thermographie + examen clinique	0,50	37,9	37,9
Thermographie	0,33	37,9	38,0

* COR = Indice de la courbe opérationnelle des résultats
(total des simulateurs)

** Ex.S. = % Valeur globale, exactitude simple

*** Ex.D. = % Valeur globale, exactitude double

+ V.G. = Valeur globale

++ = Résultats du B-200 + thermographie

Indécis classifiés comme
mauvaises réponses

ANNEXE I

Résultats de la spinoscopie obtenus de Spinex

Résultats de la spinoscopie obtenus de Spinex

Des raisons d'ordres différents (logistique, communications, nature du traitement des données) ont entraîné diverses dérogations à la séquence des opérations prévues au protocole de recherche initial.

À cause de ces différences, l'équipe de recherche a convenu, en accord avec l'IRSST, que les résultats de Spinex seraient traités séparément et présentés sous forme d'annexe au rapport.

Ces différences portent notamment sur les points suivants :

- lectures faites après le regroupement de l'information sur tous les sujets, et non sujet par sujet, comme ce fut le cas pour les autres utilisateurs des technologies et les évaluateurs cliniques;
- lectures faites d'emblée avec le consensus de deux utilisateurs;
- exclusions de huit sujets pour cause de données dites insuffisantes à la 1^{re} lecture et de trois sujets (pas nécessairement les mêmes) au cours de la 2^e lecture.

Les raisons invoquées par Spinex pour justifier les principales dérogations au protocole étaient les suivantes :

1. données manquantes (examens incomplets) pour certains sujets;
2. difficultés à distinguer les sujets simulateurs et dissimulateurs occasionnées par ces circonstances.

L'équipe de recherche a souligné les éléments suivants à Spinex :

1. Données incomplètes de la spinoscopie

- La méthode et le déroulement de la recherche étaient connus de tous les participants à l'étude puisqu'ils leur avaient été présentés au cours de réunions d'information, particulièrement à la suite de multiples rencontres entre les D^{rs} Serge Gracovetsky et Nicolas Newman de Spinex et l'équipe de recherche pour tenter d'achever les lectures selon le protocole de recherche.
- Le choix des techniciens et l'appréciation de leurs compétences étaient sous la responsabilité des utilisateurs ou des promoteurs des technologies à l'étude. L'équipe s'est assurée que ces participants avaient reçu une formation adéquate, soit qu'ils possédaient une attestation de formation, soit encore, lorsque cette information n'était pas disponible, comme dans le cas de la spinoscopie, que les utilisateurs ou les promoteurs avaient sélectionné les meilleurs candidats. Spinex avait de fait désigné son meilleur technicien, soit M. Steeve Asselin, un étudiant en maîtrise auprès du D^r Gracovetsky et ayant participé aux travaux de développement du spinoscope. C'est d'ailleurs M. Asselin qui a formé les premiers techniciens en spinoscopie au Québec et aux États-Unis. Son remplaçant pour l'examen des derniers sujets de l'étude a aussi été choisi par Spinex.

- M. Asselin a fourni une attestation écrite à l'effet que les enregistrements ont été faits selon les pratiques alors en vigueur et qu'ils sont interprétables. Son cahier de bord détaillé sur le déroulement de chaque examen est disponible à l'IRSST.
 - La phase I de l'étude avec les sujets normaux avait justement pour but de familiariser tous les participants avec les difficultés pouvant surgir au cours de la phase II.
 - Afin de respecter son mandat d'assurer la non-divulgence d'informations pouvant contribuer à révéler l'état des sujets, l'équipe de recherche a demandé aux techniciens des technologies, par écrit, de respecter la confidentialité des résultats. Il était convenu par tous et stipulé dans le protocole de recherche que les enregistrements seraient scellés après chaque test et qu'ils ne seraient disponibles que pour les lectures en cas par cas.
 - Au moment de la préparation des sujets, l'équipe de recherche leur a demandé de se soumettre intégralement à tous les tests selon leur capacité physique et les a avisés qu'il était inhérent à l'étude que si une portion d'un test s'avérait trop exigeante, cela faisait partie des limites de ce test. Cette limite s'appliquait à tous les tests, incluant les examens cliniques. Il aurait été contre l'éthique d'effectuer un test qui aurait pu être préjudiciable à un malade ou même à un sujet normal, et le technicien de Spinex l'a bien compris, comme en témoignent ses notes, de même que la technicienne du B-200 et les évaluateurs cliniques.
 - Enfin, soulignons que deux sujets sains, incapables de terminer les tests selon l'exigence de la tâche, ont été éliminés (Annexe III).
2. Difficultés de distinguer entre les sujets dissimulateurs et simulateurs
- Rappelons que la participation de ces sujets était connue de tous les participants dès le début de l'étude.
 - Leur présence faisait partie intégrante de l'étude puisque les fabricants affirmaient que certaines des technologies étaient aptes à détecter de tels sujets.
 - Il était prévu de tenir compte, dans l'analyse des résultats, de ce degré de difficulté accrue pour les technologies. Pour ce faire, nous avons prévu et utilisé, en plus de l'évaluation par la technique COR, d'abord une analyse d'exactitude simple, à savoir la classification exacte de l'absence ou de la présence d'une lombalgie indépendamment de l'état de simulateur/dissimulateur, puis une analyse d'exactitude double, c'est-à-dire la classification exacte des sujets dans un des quatre groupes.

Soulignons, pour conclure ces deux points, que les évaluateurs cliniques et les utilisateurs de la thermographie et du B-200 se sont soumis à cette évaluation, même si les évaluateurs cliniques et les utilisateurs du B-200 ont connu des difficultés similaires à celles des utilisateurs de la spinoscopie.

À la suite d'un long échange de correspondance entre l'IRSST et le D^r Serge Gracovetsky, échange suscité par la décision de ce dernier de ne pas se conformer aux exigences du protocole de recherche ni d'utiliser les fiches prévues permettant une analyse par la méthode COR, nous avons quand même accepté de traiter ses résultats. Il a été convenu avec lui et avec l'IRSST que ses résultats seraient présentés en annexe de notre rapport, et que les limites inhérentes à ce traitement des données seraient explicitées, à savoir :

- lectures faites après le regroupement de l'information sur tous les sujets, et non sujet par sujet, selon la procédure suivie par les autres utilisateurs de technologies et par les évaluateurs cliniques;
- lectures faites avec le consensus de deux utilisateurs;
- exclusion de huit sujets pour cause de données dites insuffisantes à la 1^{re} lecture et de trois sujets (pas nécessairement les mêmes) au cours de la 2^e lecture.

Nous avons reçu du D^r Gracovetsky une première série de données le 23 novembre 1994 et une seconde en décembre 1994, après remise des examens cliniques. Ces données sont résumées dans les tableaux 1A et 1B, et les résultats sont présentés dans les tableaux 2, 3 et 4, suivant le même format que ceux des autres technologies et de l'examen clinique. Il faut remarquer l'exclusion de huit sujets à la 1^{re} lecture et de trois autres à la 2^e lecture pour cause de données insuffisantes. Les tableaux 3 et 4 contrastent la performance de la spinoscopie avec l'analyse de l'exactitude double avant et après l'exclusion des indécis, au nombre de 13 pour la 1^{re} lecture et de 15 pour la 2^e lecture.

Les résultats ont donc dû être analysés en tenant compte de ces dérogations au protocole initial.

Nous avons aussi procédé à une analyse selon la méthode COR (courbe de la caractérisation opérationnelle des résultats). Pour ce faire, John A. Hanley, du service d'épidémiologie et de biostatistiques de l'Université McGill, a transcrit et codifié les résultats bruts selon une échelle similaire à celle utilisée dans la recherche. Les tableaux 5A et 5B font état des résultats suivant la méthode COR, et les tableaux 5C et 5D des résultats quant à la sensibilité et la spécificité lorsque le nombre insuffisant de sujets malhonnêtes ne permettait pas de générer des courbes COR. Les tableaux 6 et 7 donnent le codage individuel de chaque sujet.

Rappel : Les 1^{res} lectures effectuées par Spinex sont présentées au tableau 6 intitulé : "True and IRSST Average Spinoscope Decision", Nov. 23, 1994, p. 1 à 16.

Les 2^{es} lectures effectuées par Spinex avec les informations cliniques sont présentées au tableau 7 intitulé : "True and IRSST Average Spinoscope Decision", Dec. 16, 1994, p. 1 à 30.

LISTE DES TABLEAUX

Tableau 1A	Transcription des données de Spinex du cahier blanc des résultats (p. 1 à 16) pour la 1 ^{re} lecture	A39
Tableau 1B	Transcription des données de Spinex du cahier blanc des résultats (p. 1 à 30) pour la 2 ^e lecture	A43
Tableau 2	Exactitude simple : aptitude à distinguer les sujets sains et malades en fonction de l'état réel, honnêtes ou malhonnêtes, en % (Spinex)	A47
Tableau 3	Spinoscopie de Spinex, exactitude double : aptitude à distinguer les sujets	A48
Tableau 4	Exactitude double : aptitude à distinguer les sujets quant à la sensibilité, la spécificité et la valeur globale pour les sujets honnêtes ou malhonnêtes, en %	A50
Tableaux 5A et 5B	Exactitude diagnostique des résultats de Spinex par la méthode COR	A51
Tableau 5A	Évaluation de la lombalgie et incertitude correspondante conditionnelle à l'état réel des sujets malhonnêtes ou honnêtes; 1 ^{re} lecture : spinoscopie seule 2 ^e lecture : spinoscopie + informations cliniques	A51
Tableau 5B	Évaluation de l'état de simulateur et incertitude correspondante conditionnelle à l'état réel des sujets sains ou malades; 1 ^{re} lecture : spinoscopie seule	A51
Tableau 5C	Fréquence, sensibilité et spécificité de la spinoscopie pour l'évaluation de l'état de simulateur (simulateur ou non); 1 ^{re} lecture <u>sans</u> les informations cliniques	A52
Tableau 5D	Fréquence, sensibilité et spécificité de la spinoscopie pour l'évaluation de l'état de simulateur (simulateur ou non); 2 ^e lecture <u>avec</u> les informations cliniques	A53
Tableau 6	True and IRSST Average Spinoscope Decision - Expert System Version 2.15 Wednesday, November 23, 1994 08:41	A54

A38

LISTE DES TABLEAUX (suite)

Tableau 7	True and IRSST Average Spinoscope Decision
	Expert System Version 2.15
	Tuesday, December 16, 1994 11:05 A70

Tableau 1A. Transcription des données de Spinex du cahier blanc
des résultats (p. 1 à 16) pour la 1^{re} lecture

Légende : H.S. = Health Status M = Mixed
 S.S. = Status Simulator B = Borderline
 C.L. = Confidence Level D = Doubtful
 - = Undecided

	H.S.	S.S.	C.L. Separating Normal from Abnormal	Comments
102536	0	0	100 %	
110139	0	0	100 %	
115603	0B	0	47 %	
120745	1M	-	100 %	Pretend N
130717	0M	-	76 %	Pretend N
134107	1M	0	100 %	
136375	0	0	100 %	
138637				Out
141132	0	0	100 %	
145162	1M	-	100 %	Pretend N
154014	1M	0	100 %	
157427	0	0	100 %	
160370	1M	0	100 %	
160714	0	0	96 %	
162132	1MB	0	50 %	
164315	0	0	100 %	
168693	0	0	100 %	
174047	1BM	0	5 %	
180713				Out
181696	1M	0	100 %	
188593	0	0	100 %	
189709	0	0	100 %	Erreur G : 409=709
192214	0	0	100 %	
193873	1BM	0	21 %	
207050	1M	0	100 %	

A40

	H.S.	S.S.	C.L. Separating Normal from Abnormal	Comments
210884				Out
212536	0	0	100 %	
219120	0B	-	15 %	
237918	1BM	0	27 %	
259131	1M	-	100 %	
301215	1M	0	100 %	
329459	0	0	76 %	
333294	0	0	100 %	
342623	1M	-	D 76 %	Pretend N
349110	0	0	100 %	
355106	0	-	100 %	
378181	0	-	86 %	Pretend Abn
383728	0	0	100 %	
388813	0	0	100 %	
392590	0B	0	42 %	
397947	1M	0	100 %	
426336	1BM	0	21 %	
432023	1BM	-	D 8 %	Pretend Abn
439445	0	0	100 %	
444669	0	0	100 %	
475165	1BM	0	D 22 %	
488239	0	0	76 %	
494968	0	0	100 %	
508668	1BM	0	6 %	
518017	1M	0	100 %	
542595	0	0	100 %	
566534	0	0	100 %	
580068	0B	0	47 %	
591589	0	0	98 %	
597283	0B	0	16 %	

	H.S.	S.S.	C.L. Separating Normal from Abnormal	Comments
600393	0	0	76 %	
606924	0B	0	43 %	
608676	0	1	100 %	Pretend Abn
629233	1BM	-	D 8 %	Pretend N
645188	1M	0	D 100 %	
649875	0B	0	52 %	
660839	1M	-	100 %	Pretend N
690108	1M	0	100 %	
702418	0	1	69 %	Pretend Abn
704921				Out
722399	0	0	100 %	
758325	0	0	100 %	
769239	1BM	0	21 %	
776326	0M	-	70 %	
800369	0	0	100 %	
803808	1M	0	100 %	
812166	0	0	100 %	
820360	0BM	0	35 %	
820726	1M	0	100 %	
823200				Out
823736	1M	0	100 %	
828686	1M	-	74 %	Pretend N
829292	0	0	100 %	
884132	1M	0	100 %	
885854	0	0	100 %	
895423	0	0	100 %	
902755				Out
910123	0	0	100 %	
911174				Out
923925				Out

A42

	H.S.	S.S.	C.L. Separating Normal from Abnormal	Comments
965191	1M	0	57 %	
966190	1M	0	100 %	

1^{re} lecture

Total de l'étude : 87	----->	[	0	sain honnête
Sujets enlevés : 8			3	sain malhonnête
Différence : 79			3	malade honnête
	2		malade malhonnête	

Tableau 1B. Transcription des données de Spinex du cahier blanc
des résultats (p. 1 à 30) pour la 2^e lecture

Légende : H.S. = Health Status B = Borderline
S.S. = Status Simulator - = Undecided

	H.S.	S.S.	Comments
102536	0	0	
110139	0	0	
115603	0B	0	
120745	0B	-	
130717	0	0	
134107	1	0	
136375	0	0	
138637	1B	1	
141132	0	0	
145162	1	-	
154014	1B	-	
157427	1B	0	
160370	1B	0	
160714	0	0	
162132	1	0	
164315	0	0	
168693	0	0	
174047	1B	0	
180713	1	0	
181696	1	0	
188593	0	-	
189709	0	0	
192214	0B	0	
193873	1B	0	
207050	1B	0	

A44

	H.S.	S.S.	Comments
210884	1	1	
212536	0	0	
219120	0	-	
237918	1B	-	
259131	1	-	
301215	1B	0	
329459	0B	0	
333294	0	0	
342623	0	0	
349110	0	0	
355106	1B	-	
378181	0	-	
383728	0B	0	
388813	0	0	
392590	0	0	
397947	1	0	
426336	1	0	
432023	1B	-	
439445	0	0	
444669	0	0	
475165	0	0	
488239	0	0	
494968	0	0	
508668	1B	0	
518017	1	0	
542595	0	0	
566534	0	0	
580068	0B	0	
591589	-	0	
597283	0	0	

	H.S.	S.S.	Comments
600393	0	0	
606924	0B	0	
608676	0	1	
629233	0B	0	
645188	1	0	
649875	0B	0	
660839	1	-	
690108	1	1	
702418	1B	-	
704921			Out
722399	0	0	
758325	0	0	
769239	0B	0	
776326	0	0	
800369	0	0	
803808	1	0	
812166	0	0	
820360	0B	0	
820726	1	0	
823200	1	-	
823736	1	0	
828686	1	0	
829292	0	0	
884132	0B	0	
885854	0	0	
895423	0	0	
902755			Out
910123	0	0	
911174	1	-	
923925			Out

A46

	H.S.	S.S.	Comments
965191	1	0	
966190	1	0	

2^e lecture

Total de l'étude : 87

Sujets enlevés : 3

Différence : 84

----->

[1 sain malhonnête
2 malade malhonnête

Tableau 2. Exactitude simple : aptitude à distinguer les sujets sains et malades*
en fonction de l'état réel, honnêtes ou malhonnêtes, en % (Spinex)

Indécis = mauvaise réponse

Exactitude simple	Sensibilité			Spécificité			Valeur globale		
	H+M	H	M	H+M	H	M	H+M	H	M
Spinoscopie			Diss.			Simul.			
	n=36	n=18	n=18	n=43	n=23	n=20	n=79	n=41	n=38
1re lecture	61,1	66,7	55,6	76,7	78,3	75	69,6	73,2	65,8
	n=39	n=21	n=18	n=45	n=23	n=22	n=84	n=44	n=40
2e lecture (i.c.)	61,5	71,4	50,0	75,6	95,7	54,5	69	84,1	52,5

i.c. = informations cliniques

n = nombre de sujets

* basée sur l'état de santé seul

H+M = honnêtes + malhonnêtes

H = honnêtes

M = malhonnêtes

Tableau 3. Spinoscopie de Spinex, exactitude double :
aptitude à distinguer les sujets

Indécis = mauvaise réponse

Indécis = enlevés

Spinoscopie 1re lecture (n = 79)	Groupes de sujets	(n = 67)	Indécis enlevés
17/23 = 74,0	Sain honnête	17/20 = 85	3
1/20 = 5	Sain malhonnête	1/16 = 6,3	4
11/18 = 61,1	Malade honnête	11/17 = 64,7	1
0/18 = 0	Malade malhonnête	0/14 = 0	4
29/79 = 37,0	Tous	29/67 = 43,3	
28/41 = 68,3	Honnête	28/37 = 75,7	
1/38 = 2,6	Malhonnête	1/30 = 3,3	
18/43 = 41,9	Sain	18/36 = 50,0	
11/36 = 30,6	Malade	11/31 = 35,5	

Total de l'étude : 87
Sujets enlevés : -8
Différence : 79

Total de l'étude : 87
Sujets enlevés : -8
Indécis enlevés : -12
Différence : 67

Tableau 3. Spinoscopie de Spinex, exactitude double :
aptitude à distinguer les sujets (suite)

Indécis = mauvaise réponse

Indécis = enlevés

Spinoscopie 2e lecture avec les informations cliniques (n = 84) Groupes de sujets (n = 69)			Indécis enlevés
21/23 = 91,3	Sain honnête	21/22 = 95,5	1
0/22 = 0	Sain malhonnête	0/13 = 0	9
13/21 = 62,0	Malade honnête	13/19 = 68,4	2
1/18 = 5,6	Malade malhonnête	1/15 = 6,7	3
35/84 = 41,7	Tous	35/69 = 50,7	
34/44 = 77,3	Honnête	34/41 = 82,9	
1/40 = 2,5	Malhonnête	1/28 = 3,6	
21/45 = 46,7	Sain	21/35 = 60	
14/39 = 35,9	Malade	14/34 = 41,2	

Total de l'étude : 87
Sujets enlevés : -3
Différence : 84

Total de l'étude : 87
Sujets enlevés : -3
Indécis enlevés : -15
Différence : 69

Tableau 4. Exactitude double : aptitude à distinguer les sujets**
quant à la sensibilité, la spécificité et la valeur globale
pour les sujets honnêtes ou malhonnêtes, en %

Exactitude double	Sensibilité			Spécificité			Valeur globale		
	H+M	H	M	H+M	H	M	H+M	H	M
1re lecture*	n=36 30,6	n=18 61,1	n=18 0	n=43 41,9	n=23 74,0	n=20 5	n=79 37,0	n=41 68,3	n=38 2,6
(indécis enlevés)	n=31 35,5	n=17 64,7	n=14 0	n=36 50,0	n=20 85	n=16 6,3	n=67 43,3	n=37 75,7	n=30 3,3
2e lecture (i.c.)*	n=39 35,9	n=21 62,0	n=18 5,6	n=45 46,7	n=23 91,3	n=22 0	n=84 41,7	n=44 77,3	n=40 2,5
(indécis enlevés)	n=34 41,2	n=19 68,4	n=15 6,7	n=35 60	n=22 95,5	n=13 0	n=69 50,7	n=41 82,9	n=28 3,6

i.c. = informations cliniques

* = Indécis = mauvaise réponse

H+M = honnêtes + malhonnêtes

H = honnêtes

M = malhonnêtes

** basée sur l'état de santé et le rôle

1re lecture	Total de l'étude :	87	----->	3 sain malhonnête 3 malade honnête 2 malade malhonnête
	Sujets enlevés :	-8		
	Différence :	79		

1re lecture (indécis enlevés) :	-12	----->	3 sain honnête 4 sain malhonnête 1 malade honnête 4 malade malhonnête
Différence :	67		

2e lecture (i.c.)	Total de l'étude :	87	---->	1 sain malhonnête 2 malade malhonnête
	Sujets enlevés :	-3		
	Différence :	84		

2e lecture (indécis enlevés)	:	-15	---->	1 sain honnête 9 sain malhonnête 2 malade honnête 3 malade malhonnête
Différence :	:	69		

Tableaux 5A et 5B. Exactitude diagnostique des résultats de Spinex
par la méthode COR

Tableau 5A. Évaluation de la lombalgie et incertitude correspondante
conditionnelle à l'état réel des sujets malhonnêtes ou honnêtes;
1^{re} lecture : spinoscopie seule,
2^e lecture : spinoscopie + informations cliniques

	État des simulateurs					
	Malhonnêtes (n = 43)		Honnêtes (n = 44)		Total (n = 87)	
	Indice	ET**	Indice	ET	Indice	ET
Spinoscopie seule	0,56	0,11	0,86	0,06	0,74	0,06
Spinoscopie + informations cliniques	0,58	0,10	0,93	0,03	0,80	0,05

Tableau 5B. Évaluation de l'état de simulateur et incertitude correspondante
conditionnelle à l'état réel des sujets sains ou malades;
1^{re} lecture : spinoscopie seule

	État des sujets					
	Sains (n = 46)		Malades (n = 41)		Total (n = 87)	
	Indice	ET	Indice	ET	Indice	ET
Spinoscopie seule	0,64*	0,06	0,65	0,17	0,63	0,13

* Un indice estimé et une incertitude correspondante étaient
utilisés selon une approche non paramétrique.

** ET = Erreur type

Note : Indice COR non fiable pour la 2^e lecture (spinoscopie + informations cliniques), car les lecteurs n'ont pas utilisé tous les niveaux de l'échelle graduée d'incertitude; seulement la sensibilité et la spécificité peuvent être données. De même, les données insuffisantes ne permettent pas de faire une analyse des sous-groupes basée sur l'état des sujets.

Tableau 5C. Fréquence, sensibilité et spécificité de la spinoscopie pour l'évaluation de l'état de simulateur (simulateur ou non); 1^{re} lecture sans les informations cliniques

a) Fréquence des résultats du test en rapport avec l'état réel

	Spinoscopie			
	Honnêtes	Indécis	Malhonnêtes	Total (n = 87)
État réel				
Honnêtes (-)	37	6	1	44
Malhonnêtes (+)	27	10	6	43

b) Sensibilité et spécificité utilisant seulement les résultats certainement positif et certainement négatif

Sensibilité : $6/33 = 0,18$

Spécificité : $37/38 = 0,97$

c) Sensibilité et spécificité à deux seuils différents (tous les résultats)

Sensibilité : $16/43 = 0,37$ $6/43 = 0,14$

Spécificité : $37/44 = 0,84$ $43/44 = 0,98$

Tableau 5D. Fréquence, sensibilité et spécificité de la spinoscopie pour l'évaluation de l'état de simulateur (simulateur ou non); 2^e lecture avec les informations cliniques

a) Fréquence des résultats du test en rapport avec l'état réel

	Spinoscopie			Total* (n = 87)
	Honnêtes	Indécis	Malhonnêtes	
État réel				
Honnêtes (-)	39	3	1	43
Malhonnêtes (+)	26	12	4	42

* Données manquantes = 2

b) Sensibilité et spécificité utilisant seulement les résultats certainement positif et certainement négatif

Sensibilité : $4/30 = 0,13$

Spécificité : $39/40 = 0,98$

c) Sensibilité et spécificité à deux seuils différents (tous les résultats)

Sensibilité : $16/42 = 0,38$ $4/42 = 0,10$

Spécificité : $39/43 = 0,91$ $42/43 = 0,98$

Note : Indice COR non fiable pour la 2^e lecture (spinoscopie + informations cliniques), car les lecteurs n'ont pas utilisé tous les niveaux de l'échelle graduée d'incertitude; seulement la sensibilité et la spécificité peuvent être données. De même, les données insuffisantes ne permettent pas de faire une analyse des sous-groupes basée sur l'état des sujets.

Tableau 6

Wednesday, November 23, 1994 08:41

Page 1 of 10

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

A54

ID #	Summary of True Spinoscope Decision Threshold set at 80%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the machine use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which lead the rules of the Expert System
102536 U 22 21/JAN/91 Load lifted Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 7.6%	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%	Thoracic Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness Left Right Relative importance of pathology 0 % Do not know Not enough data	Honest Cheater Unclear Pretend NORMAL Pretend ABNORMAL	✓
110139 F 45 17/DEC/90 Load lifted Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 5.8%	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%	Thoracic Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness Left Right Relative importance of pathology 0 % Do not know Not enough data	Honest Cheater Unclear Pretend NORMAL Pretend ABNORMAL	✓
115603 U 24 08/JAN/93 Load lifted Max 17 kg	Normal up to 17 kg. No data beyond that load. Average distance to threshold 1.6%	Normal / Borderline Valid Average ? Yes Confidence level separating normal from abnormal 47%	Thoracic Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness Left Right Relative importance of pathology 33 % Do not know Not enough data	Honest Cheater Unclear Pretend NORMAL Pretend ABNORMAL	Restrictions on the left middle and lower spine compensated by a relatively hypermobile TL junction. Borderline case because the pathology is well compensated but the clinician may think otherwise.
120745 F 30 22/APR/91 Load lifted Max 23 kg	Normal up to 3 kg, then abnormal up to 23 kg. Average distance to threshold -8.6%	Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100%	Thoracic Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness Left Right Relative importance of pathology 100 % Do not know Not enough data	Honest Cheater Unclear Pretend NORMAL Pretend ABNORMAL	Highly resistant in axial tension to the left. Consider disarticulation. Minor sprain.
130717 F 32 16/APR/91 Load lifted Max 23 kg	Normal up to 12 kg, then abnormal up to 23 kg. Average distance to threshold 2.6%	Normal Valid Average ? Yes Confidence level separating normal from abnormal 76%	Thoracic Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness Left Right Relative importance of pathology 17 % Do not know Not enough data	Honest Cheater Unclear Pretend NORMAL Pretend ABNORMAL	Mild torsional injury to the left at L5. Weak evidence of disarticulation.
134107 U 21 31/AUG/92 Load lifted Max 9 kg	Normal unloaded and abnormal thereafter Average distance to threshold -13.6%	Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100%	Thoracic Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness Left Right Relative importance of pathology 100 % Do not know Not enough data	Honest Cheater Unclear Pretend NORMAL Pretend ABNORMAL	Minor hypermobility at TL junction. Irregular response suggest mild sprain.

Wednesday, November 23, 1994 08:41

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 90%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which lead the idea of the Expert System
136375 M 30 21JAO01 Lead blood Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 0.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	Thoraco Lumbar Junction ☐ Lowest lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	✓ No reasons to stop lifting more. Probably normal.
136637 F 30 28JAO02 Lead blood Max 8 kg	Remove from population Average distance to threshold 0.0 %	Abnormal / Borderline Valid Average ? Doubtful Confidence level separating normal from abnormal 4 %	Thoraco Lumbar Junction ☐ Lowest lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	X Was downgraded because of TL restriction. The performance increases with loads (very good at 31 kg). May be apprehensive at first. Essentially normal in spite of his TL restriction. Candidate for demotion.
141132 F 30 22JAO00 Lead blood Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 3.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	Thoraco Lumbar Junction ☐ Lowest lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	✓ Was downgraded because of TL restriction. The performance increases with loads (very good at 31 kg). May be apprehensive at first. Essentially normal in spite of his TL restriction. Candidate for demotion.
145162 M 30 21JAO01 Lead blood Max 31 kg	Normal unloaded and abnormal thereafter Average distance to threshold -0.0 %	Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	Thoraco Lumbar Junction ☐ Lowest lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	✓ Quasi normal unloaded but immediate sprain pattern at 4.8 kg. There is not enough data to corroborate this finding. Return for human assessment.
154014 F 30 10JAO00 Lead blood Max 8 kg	Abnormal Average distance to threshold -0.0 %	Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	Thoraco Lumbar Junction ☐ Lowest lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	✓ Mild scoliosis.
157427 F 30 17JAO00 Lead blood Max 12 kg	Normal up to 12 kg. No data beyond that load. Average distance to threshold 3.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	Thoraco Lumbar Junction ☐ Lowest lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	✓

This list replace the printout dated Monday November 21, 1994

A55

Wednesday, November 23, 1994 08:41

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 90%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which fired the rules of this Expert System
160370 F 36 28JAN91 Load Blad Max 23 kg	Normal up to 3 kg, then abnormal up to 23 kg. Average distance to threshold -2.0 %	Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	Thoraco Lumbal Junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Mild right scoliosis. Serious drop of performance at 23 kg suggesting that this load is close to her limits. ✓
160714 F 47 19NOV90 Load Blad Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 10.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 99 %	Thoraco Lumbal Junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Very strong female. ✓
162132 M 25 05JUN93 Load Blad Max 23 kg	Normal unloaded and abnormal thereafter Average distance to threshold -1.5 %	Abnormal / Borderline Valid Average ? Yes Confidence level separating normal from abnormal 59 %	Thoraco Lumbal Junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Very mild lumbal injury coupled with a consistent upper lumbar stiffness lowered the score below the threshold. Clinicians may not detect this anomaly and see this subject as normal. ✓
164315 M 26 17JAN91 Load Blad Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 7.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	Thoraco Lumbal Junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Some muscular strategy at high loads. ✓
168693 M 45 17DEC90 Load Blad Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 7.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	Thoraco Lumbal Junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	✓
174047 M 33 21FEB91 Load Blad Max 14 kg	Normal up to 8 kg, then abnormal up to 14 kg. Average distance to threshold 0.0 %	Abnormal / Borderline Valid Average ? Yes Confidence level separating normal from abnormal 5 %	Thoraco Lumbal Junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Excellent unloaded but sprain and evidence of mild lumbal injury at higher loads. Did not complete 23 kg. ✓

Wednesday, November 23, 1994 08:41

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 90%	Average Decision This methodical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functioned Anomalies (Pathology)	Malingering	Comments Main events which feed the rules of the Expert System
180713 F 60 210C02 Load 10kg Max 0 kg	Remove from population Average distance to threshold -13.6 %	Abnormal Valid Average ? Doubtful Confidence level separating normal from abnormal 100%	Thoracic Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (radio, overweight) Potential conflict clinicians Sprain All lumbar stiffness Do not know Not enough data	Honest Cheater Unclear Pretend NORMAL Pretend ABNORMAL	Slow sprain pattern. Minor right scoliosis. Possibly honest injured but not enough data to confirm.
181696 F 20 00F002 Load 10kg Max 14 kg	Normal up to 1 kg, then abnormal up to 14 kg. Average distance to threshold -7.6 %	Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100%	Thoracic Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (radio, overweight) Potential conflict clinicians Sprain All lumbar stiffness Do not know Not enough data	Honest Cheater Unclear Pretend NORMAL Pretend ABNORMAL	Clear sprain at 14 kg.
186593 M 31 22AP01 Load 10kg Max 14 kg	Normal up to 14 kg. No data beyond that load. Average distance to threshold 0.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%	Thoracic Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (radio, overweight) Potential conflict clinicians Sprain All lumbar stiffness Do not know Not enough data	Honest Cheater Unclear Pretend NORMAL Pretend ABNORMAL	No reasons to interrupt the 14 kg lift unless the patient apprehended a flare of pain and aborted before the onset of sprain.
189409 M 20 01F001 Load 10kg Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 6.6 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%	Thoracic Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (radio, overweight) Potential conflict clinicians Sprain All lumbar stiffness Do not know Not enough data	Honest Cheater Unclear Pretend NORMAL Pretend ABNORMAL	Minor left scoliosis
192214 F 20 00F002 Load 10kg Max 14 kg	Normal up to 14 kg. No data beyond that load. Average distance to threshold 3.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%	Thoracic Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (radio, overweight) Potential conflict clinicians Sprain All lumbar stiffness Do not know Not enough data	Honest Cheater Unclear Pretend NORMAL Pretend ABNORMAL	
193673 M 41 27JA02 Load 10kg Max 6 kg	Normal up to 1 kg, then abnormal up to 5 kg. Average distance to threshold -1.0 %	Abnormal / Borderline Valid Average ? Mixed Confidence level separating normal from abnormal 2.1 %	Thoracic Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (radio, overweight) Potential conflict clinicians Sprain All lumbar stiffness Do not know Not enough data	Honest Cheater Unclear Pretend NORMAL Pretend ABNORMAL	Only one load completed. Sprain pattern detected but not enough data to confirm.

A57

This list replace the printout dated Monday November 21, 1994

Wednesday, November 23, 1994 08:41

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 80%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which feed the rules of the Expert System
207050 F 30 04J002 Load lifted Max 14 kg	Abnormal Average distance to threshold -0.8%	Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100%	Thoraco Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Possible time out at 23 kg.
210884 F 30 30J002 Load lifted Max 0 kg	Remove from population Average distance to threshold -3.9%	Abnormal Valid Average ? Doubtful Confidence level separating normal from abnormal 89%	Thoraco Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☐ Honest ☒ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	No reasons to refuse lifting. Has traces of torsional injury which cannot be confirmed because of lack of data. Must be removed from population since testing protocol was not followed and no data recovery is possible.
212536 F 30 04J002 Load lifted Max 23 kg	Normal up to 23 kg. No data beyond that load. Average distance to threshold 0.8%	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%	Thoraco Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☐ Honest ☒ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	
219120 F 32 30J002 Load lifted Max 32 kg	Normal up to 32 kg. No data beyond that load. Average distance to threshold 0.8%	Normal / Borderline Valid Average ? Yes Confidence level separating normal from abnormal 15%	Thoraco Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☐ Honest ☒ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Abd radiation anomaly downgraded an otherwise excellent and consistent performance in flexion. Very difficult anomaly to detect by clinicians. May be candidate for disqualification. Since right on the border, the decision can go either way.
237918 F 35 13J002 Load lifted Max 6 kg	Normal up to 3 kg, then abnormal up to 5 kg. Average distance to threshold -1.8%	Abnormal / Borderline Valid Average ? Yes Confidence level separating normal from abnormal 27%	Thoraco Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☐ Honest ☒ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	The trend is towards pathology as load increases. No reasons to stop lifting and not enough data collected to confirm.
259131 F 47 01J002 Load lifted Max 5 kg	Normal unloaded and abnormal thereafter Average distance to threshold -5.9%	Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100%	Thoraco Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☐ Honest ☒ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Torsional problem unconfirmed.

Wednesday, November 23, 1994 08:41

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plusminus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 99%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which lead the rules of the Expert System
301215 M 23 16NOV90 Load lifted Max 14 kg	Normal up to 9 kg, then abnormal up to 14 kg. Average distance to threshold -4.8%	Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness ☐ Left ☐ Right Relative importance of pathology 100 % ☐ Do not know ☐ Not enough data	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Good unloaded, bad loaded. Sprain. Minor right scoliosis. ✓
329459 F 26 16SEP92 Load lifted Max 14 kg	Normal up to 14 kg. No data beyond that load. Average distance to threshold 2.0%	Normal Valid Average ? Yes Confidence level separating normal from abnormal 76%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness ☐ Left ☐ Right Relative importance of pathology 17 % ☐ Do not know ☐ Not enough data	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	A bit above. ✓
333294 F 23 23JAN91 Load lifted Max 14 kg	Normal up to 14 kg. No data beyond that load. Average distance to threshold 7.8%	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness ☐ Left ☐ Right Relative importance of pathology 9 % ☐ Do not know ☐ Not enough data	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	✓
342623 M 17 30JAN92 Load lifted Max 31 kg	Normal up to 10 kg, then abnormal up to 31 kg. Average distance to threshold -2.0%	Abnormal Valid Average ? Doubtful Confidence level separating normal from abnormal 76%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness ☐ Left ☐ Right Relative importance of pathology 8.3 % ☐ Do not know ☐ Not enough data	☐ Honest ☐ Cheater ☒ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	The 31 kg lift was very poor. Subject exceeded his limits. Possible dislocation. ✓
349110 M 30 24JUN91 Load lifted Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 3.9%	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness ☐ Left ☐ Right Relative importance of pathology 0 % ☐ Do not know ☐ Not enough data	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Probably normal but atypical. Very small indication of sprain at high load. Probably not significant. ✓
355106 F 25 23JAN91 Load lifted Max 6 kg	Normal up to 5 kg. No data beyond that load. Average distance to threshold 10.8%	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness ☐ Left ☐ Right Relative importance of pathology 0 % ☐ Do not know ☐ Not enough data	☐ Honest ☐ Cheater ☒ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Time out at 14 kg. ✓

Wednesday, November 23, 1994 08:41

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 90%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which feed the rules of the Expert System
378181 F 40 16AP002 Load lifted Max 5 kg	Normal up to 5 kg. No data beyond that load. Average distance to threshold 2.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 99 %	Thoraco Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Highly inconsistent. Suspect simulation. No actual injury detected. ✓
383728 F 27 26JAO01 Load lifted Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 3.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	Thoraco Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Slow at 31 kg but otherwise normal. May have reached her limits. ✓
388813 F 26 130E000 Load lifted Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 3.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	Thoraco Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Minor upper spine stiffness. ✓
392590 Load lifted Max 35 kg	Normal up to 35 kg. No data beyond that load. Average distance to threshold 1.0 %	Normal / Borderline Valid Average ? Yes Confidence level separating normal from abnormal 42 %	Thoraco Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Minor upper spine stiffness. ✓
397947 F 27 24SE002 Load lifted Max 12 kg	Abnormal Average distance to threshold -14.0 %	Abnormal / Borderline Valid Average ? Mixed Confidence level separating normal from abnormal 100 %	Thoraco Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Sprain pattern. ✓
426336 M 40 22JDO02 Load lifted Max 23 kg	Normal unloaded and abnormal thereafter Average distance to threshold -1.0 %	Abnormal / Borderline Valid Average ? Mixed Confidence level separating normal from abnormal 21 %	Thoraco Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Half of the lateral bend data is missing and not recoverable. Consider removing this subject from the population since no information on torsional pathology is detectable. ✓

Wednesday, November 23, 1994 08:41

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 80%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which feed the rules of the Expert System
432023 F 31 22HC060 Load 14 kg Max 14 kg	Normal up to 8 kg, then abnormal up to 14 kg. Average distance to threshold -1.0 %	Abnormal / Borderline Mixed Valid Average ? Yes Confidence level separating normal from abnormal 8 %	☐ Thoracic Lumbar junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Proland NORMAL ☐ Proland ABNORMAL	Better at 14 kg than 8 kg. TL restriction is minor.
439445 M 26 04FE081 Load 14 kg Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 7.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	☐ Thoracic Lumbar junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Proland NORMAL ☐ Proland ABNORMAL	
444669 M 26 13DE060 Load 14 kg Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 7.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	☐ Thoracic Lumbar junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Proland NORMAL ☐ Proland ABNORMAL	Minor right lumbar scoliosis. 31 kg close to the limit. Minor axial anomaly.
475165 F 26 15AF062 Load 14 kg Max 23 kg	Normal up to 8 kg, then abnormal up to 23 kg. Average distance to threshold -1.0 %	Abnormal / Borderline Mixed Valid Average ? Doubtful Confidence level separating normal from abnormal 22 %	☐ Thoracic Lumbar junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Proland NORMAL ☐ Proland ABNORMAL	Too fast to have sprain. Contradictions in behavior.
488239 M 36 16AL060 Load 14 kg Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 2.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 78 %	☐ Thoracic Lumbar junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Proland NORMAL ☐ Proland ABNORMAL	Minor pathology for loads greater than 14 kg.
494966 F 41 08AC060 Load 14 kg Max 23 kg	Normal up to 23 kg. No data beyond that load. Average distance to threshold 6.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	☐ Thoracic Lumbar junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Proland NORMAL ☐ Proland ABNORMAL	23 Kg close to ten trials.

ID #	Summary of True Spinoscope Decision Threshold set at 80%	Average Decision The mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which feed the rules of the Expert System
508668 F 22 07FE01 Load Med Max 14 kg	Normal up to 8 kg, then abnormal up to 14 kg. Average distance to threshold 0.0%	Abnormal / Borderline Valid Average ? Yes Confidence level separating normal from abnormal 0.0%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	The whole lumbar spine has minor stiffness. ✓
518017 F 20 06FE02 Load Med Max 23 kg	Normal unloaded and abnormal thereafter Average distance to threshold -0.0%	Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	The lower lumbar restriction is clear. ✓
542595 F 23 17JAC01 Load Med Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 7.0%	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	✓
568534 M 34 04FE01 Load Med Max 23 kg	Normal up to 23 kg. No data beyond that load. Average distance to threshold 9.0%	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	✓
580068 M 20 16JAC02 Load Med Max 14 kg	Normal up to 14 kg. No data beyond that load. Average distance to threshold 1.0%	Normal / Borderline Valid Average ? Yes Confidence level separating normal from abnormal 47%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	✓
591589 F 34 17JAC01 Load Med Max 23 kg	Normal up to 23 kg. No data beyond that load. Average distance to threshold 3.0%	Normal Valid Average ? Yes Confidence level separating normal from abnormal 9.0%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	A bit slow at 23 kg. May have exceeded her limit. ✓

Wednesday, November 23, 1994 08:41

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 90%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which limit the rules of the Expert System
597283 M 55 11JAN93 Load lifted Max 23 kg	Normal up to 23 kg. No data beyond that load. Average distance to threshold 1.8%	Normal / Borderline Valid Average ? Yes Confidence level separating normal from abnormal 16%	☒ Thoraco Lumbar junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (redo, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness ☒ Left ☐ Right Relative importance of pathology 33 % ☒ Do not know ☒ Not enough data	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Minimal pathology around 20 kg. ✓
600393 M 40 23JUN90 Load lifted Max 14 kg	Normal up to 14 kg. No data beyond that load. Average distance to threshold 2.9%	Normal Valid Average ? Yes Confidence level separating normal from abnormal 76%	☒ Thoraco Lumbar junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (redo, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness ☒ Left ☐ Right Relative importance of pathology 17 % ☒ Do not know ☒ Not enough data	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Could be disqualification since the 9 and 14 kg lifts indicate sprain but the 4kg and 23 kg lifts data are too corrupted to confirm the pathology. ✓
606924 F 43 28JUN92 Load lifted Max 8 kg	Normal up to 8 kg. No data beyond that load. Average distance to threshold 1.9%	Normal / Borderline Valid Average ? Yes Confidence level separating normal from abnormal 43%	☒ Thoraco Lumbar junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (redo, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness ☒ Left ☐ Right Relative importance of pathology 33 % ☒ Do not know ☒ Not enough data	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Overweight - Caution in interpreting skin markers motion. ✓
608676 F 27 21FEB91 Load lifted Max 5 kg	Normal up to 5 kg. No data beyond that load. Average distance to threshold 8.8%	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%	☒ Thoraco Lumbar junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (redo, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness ☒ Left ☐ Right Relative importance of pathology 0 % ☒ Do not know ☒ Not enough data	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	✓
629233 M 18 17JUN90 Load lifted Max 32 kg	Normal up to 10 kg, then abnormal up to 32 kg. Average distance to threshold -1.8%	Abnormal / Borderline Mixed Valid Average ? Doubtful Confidence level separating normal from abnormal 8%	☒ Thoraco Lumbar junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (redo, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness ☒ Left ☐ Right Relative importance of pathology 67 % ☒ Do not know ☒ Not enough data	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	The pathology interferes with function at high loads. The 32 kg lift is laborious. Consider 32 kg to be the limit of frequent lifting. ✓
645188 M 51 22JUN92 Load lifted Max 23 kg	Normal up to 7 kg, then abnormal up to 23 kg. Average distance to threshold -3.8%	Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100%	☒ Thoraco Lumbar junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (redo, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness ☒ Left ☐ Right Relative importance of pathology 100 % ☒ Do not know ☒ Not enough data	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	The spine stiffness increases with loads. ✓

This list replace the printout dated Monday November 21, 1994

Wednesday, November 23, 1994 08:41

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 80%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which lead the rules of the Expert System
649875 M 32 01J002 Load lifted Max 23 kg	Normal up to 23 kg. No data beyond that load. Average distance to threshold 1.0 %	Normal / Borderline Valid Average ? Yes Confidence level separating normal from abnormal 52 %	Thoracic Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Minor pathology.
660839 M 21 26SE001 Load lifted Max 31 kg	Abnormal Average distance to threshold -10.8 %	Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	Thoracic Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Very consistent just below the threshold of normality. The pathology is minor but consistently there.
690108 F 22 11J003 Load lifted Max 5 kg	Abnormal Average distance to threshold -8.0 %	Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	Thoracic Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Sprain pelvis.
702418 F 33 18J000 Load lifted Max 5 kg	Normal up to 5 kg. No data beyond that load. Average distance to threshold 2.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 00 %	Thoracic Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	No reasons to stop lifting. Consider stimulation.
704921 M 43 20J001 Load lifted Max 0 kg	Remove from population Average distance to threshold 4.0 %	Normal Valid Average ? Doubtful Confidence level separating normal from abnormal 100 %	Thoracic Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	No reason for refusing to lift anything. Strongly suspect stimulation.
722399 M 21 27J002 Load lifted Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 0.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	Thoracic Lumbar junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	

Wednesday, November 23, 1994 08:41

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 90%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which feed the rules of the Expert System
758325 M 33 31/0002 Load 10kg Max 23 kg	Normal up to 23 kg. No data beyond that load. Average distance to threshold 7.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	☐ Thoraco Lumbar junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	✓
769239 F 33 19/0000 Load 10kg Max 14 kg	Normal up to 6 kg, then abnormal up to 14 kg. Average distance to threshold -1.0 %	Abnormal / Borderline Valid Average ? Yes Confidence level separating normal from abnormal 21 %	☐ Thoraco Lumbar junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Torsional injury on the right. ✓
778326 M 25 18/0002 Load 10kg Max 31 kg	Normal up to 22 kg, then abnormal up to 31 kg. Average distance to threshold 2.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 70 %	☐ Thoraco Lumbar junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Lower lumbar restriction increasing with loads. The pathology interferes with function quite late and is likely to be missed by the clinician. ✓
800369 F 47 22/0000 Load 10kg Max 23 kg	Normal up to 23 kg. No data beyond that load. Average distance to threshold 7.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	☐ Thoraco Lumbar junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Quite good for his age. ✓
803808 F 43 20/0001 Load 10kg Max 14 kg	Normal unloaded and abnormal thereafter Average distance to threshold -3.0 %	Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	☐ Thoraco Lumbar junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Overweight - Caution in interpreting skin markers motion, especially the lateral bend data on left. ✓
812166 M 33 28/0001 Load 10kg Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 0.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	☐ Thoraco Lumbar junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	✓

A65

This list replace the printout dated Monday November 21, 1994

Wednesday, November 23, 1994 08:41

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 50%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which limit the rules of the Expert System
820360 F 47 184000 Load lifted Max 31 kg	Normal up to 23 kg, then abnormal up to 31 kg. Average distance to threshold 1.0 %	Normal / Borderline Mixed Valid Average ? Yes Confidence level separating normal from abnormal 35 %	☒ Thoracic Lumbar junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (radio, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness ☒ Left ☐ Right Relative importance of pathology 33 % ☒ Do not know ☒ Not enough data	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Failed at 31 kg. Slight stiffness, possibly normal for age. ✓
820726 M 37 245001 Load lifted Max 14 kg	Abnormal Average distance to threshold -27.0 %	Abnormal Mixed Valid Average ? Yes Confidence level separating normal from abnormal 100 %	☒ Thoracic Lumbar junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (radio, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness ☒ Left ☐ Right Relative importance of pathology 100 % ☒ Do not know ☒ Not enough data	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Failed to lift 23 kg. Stiffness detected at 4/5. ✓
823200 F 28 203001 Load lifted Max 0 kg	Remove from population Average distance to threshold -14.0 %	Abnormal OUT Valid Average ? Doubtful Confidence level separating normal from abnormal 100 %	☒ Thoracic Lumbar junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (radio, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness ☒ Left ☐ Right Relative importance of pathology 100 % ☒ Do not know ☒ Not enough data	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Still pelvic detected during the unloaded recovery from flexion. The refusal to comply with the test protocol did not permit to assess the normally portion. Hence the relatively high ratio of pathology which must be rejected as possible artifact. This subject must be removed from the population. ✗
823738 F 28 084003 Load lifted Max 12 kg	Normal unloaded and abnormal thereafter Average distance to threshold -7.0 %	Abnormal Mixed Valid Average ? Yes Confidence level separating normal from abnormal 100 %	☒ Thoracic Lumbar junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (radio, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness ☒ Left ☐ Right Relative importance of pathology 100 % ☒ Do not know ☒ Not enough data	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Consistent with disimulation except for speed which is good. ✓
828686 M 23 08HC000 Load lifted Max 31 kg	Normal up to 15 kg, then abnormal up to 31 kg. Average distance to threshold -2.0 %	Abnormal Mixed Valid Average ? Doubtful Confidence level separating normal from abnormal 74 %	☒ Thoracic Lumbar junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (radio, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness ☒ Left ☐ Right Relative importance of pathology 83 % ☒ Do not know ☒ Not enough data	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Slow down at 31 kg. The EMG response is atypical. Possible attempt to fool the machine. Sprain detected at high loads. ✓
829292 F 20 07FE001 Load lifted Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 0.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	☒ Thoracic Lumbar junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (radio, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness ☒ Left ☐ Right Relative importance of pathology 0 % ☒ Do not know ☒ Not enough data	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Very supple. ✓

Wednesday, November 23, 1994 09:41

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 80%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which feed the rules of the Expert System
884132 M 30 00AP003 Load lifted Max 12 kg	Normal unloaded and abnormal thereafter Average distance to threshold -14.0 %	Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	Thoraco Lumbar function Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Reeling and shoving with loads. Sprain.
885854 M 35 24SE002 Load lifted Max 35 kg	Normal up to 35 kg. No data beyond that load. Average distance to threshold 0.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	Thoraco Lumbar function Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	
895423 M 25 21JAN01 Load lifted Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 5.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	Thoraco Lumbar function Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	
902755 M 30 22OC002 Load lifted Max 6 kg	Remove from population Average distance to threshold 3.0 %	Normal Valid Average ? Doubtful Confidence level separating normal from abnormal 100 %	Thoraco Lumbar function Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Did not lift 12 kg. Sprain at 4.8 kg but not unloaded. Not enough data to confirm.
910123 F 30 22OC002 Load lifted Max 17 kg	Normal up to 17 kg. No data beyond that load. Average distance to threshold 4.0 %	Normal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	Thoraco Lumbar function Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Anomaly in axial rotation. Probably old injury well recovered. Minor impact on function. Very supple spine.
911174 Load lifted Max 0 kg	Remove from population Average distance to threshold -19.0 %	Abnormal Valid Average ? Doubtful Confidence level separating normal from abnormal 100 %	Thoraco Lumbar function Lower lumbar restriction LB torsion injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Very likely a fraud. Not enough data to analyze the source of the inconsistency. This subject must be removed from the population since the deviation from the testing protocol is too severe to recover the missing data.

A67

This list replace the printout dated Monday November 21, 1994

Wednesday, November 23, 1994 08:41
TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15
with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 80%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which feed the rules of the Expert System
923925 M 36 10M000 Lead 0 kg Max 0 kg	Remove from population Average distance to threshold: 9.0 %	Normal Valid Average ? Doubtful Confidence level separating normal from abnormal: 100 %	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness ☐ Left ☐ Right Relative importance of pathology: 0 % ☐ Do not know ☐ Not enough data	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Normal unloaded. No reasons to refuse lifting more. The protocol was not followed. Remove from population. X
965191 M 47 01J002 Lead 0 kg Max 23 kg	Normal unloaded and abnormal thereafter Average distance to threshold: -2.9 %	Abnormal Valid Average ? Yes Confidence level separating normal from abnormal: 87 %	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness ☐ Left ☐ Right Relative importance of pathology: 93 % ☐ Do not know ☐ Not enough data	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Individual functional parameters are OK. The subject score was downgraded because of an unusual global response suggesting a torsional injury at 4/5 on the right indicating also the left side. ✓
966190 M 43 21F001 Lead 0 kg Max 23 kg	Normal unloaded and abnormal thereafter Average distance to threshold: -9.0 %	Abnormal Valid Average ? Yes Confidence level separating normal from abnormal: 100 %	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness ☐ Left ☐ Right Relative importance of pathology: 100 % ☐ Do not know ☐ Not enough data	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Although the patients lifted 23 kg, he did it using excessive muscular action. Torsional injury at 4/5 mainly on the left. ✓

Wednesday, November 23, 1994 08:41

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 96%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which lead the rules of the Expert System
------	---	--	-------------------------------------	-------------	---

Summary of Blind IRSST Study

Average Decision	Validity of Average Decision																
<table> <tr> <td>Normal</td><td>42</td></tr> <tr> <td>Normal Borderline</td><td>0</td></tr> <tr> <td>Abnormal</td><td>26</td></tr> <tr> <td>Abnormal Borderline</td><td>11</td></tr> </table>	Normal	42	Normal Borderline	0	Abnormal	26	Abnormal Borderline	11	<table> <tr> <td>Number of valid decisions</td><td>76</td></tr> <tr> <td>Number of doubtful decisions</td><td>12</td></tr> </table>	Number of valid decisions	76	Number of doubtful decisions	12				
Normal	42																
Normal Borderline	0																
Abnormal	26																
Abnormal Borderline	11																
Number of valid decisions	76																
Number of doubtful decisions	12																
<table> <tr> <td colspan="2">Miscellaneous</td></tr> <tr> <td>OUT</td><td>0</td></tr> <tr> <td>Mixed</td><td>10</td></tr> </table>	Miscellaneous		OUT	0	Mixed	10	<table> <tr> <td colspan="2">Detection of fraud</td></tr> <tr> <td>Cheer Honest</td><td>62</td></tr> <tr> <td>Cheer Cheaters</td><td>0</td></tr> <tr> <td>Pretend Normal</td><td>7</td></tr> <tr> <td>Pretend Abnormal</td><td>10</td></tr> </table>	Detection of fraud		Cheer Honest	62	Cheer Cheaters	0	Pretend Normal	7	Pretend Abnormal	10
Miscellaneous																	
OUT	0																
Mixed	10																
Detection of fraud																	
Cheer Honest	62																
Cheer Cheaters	0																
Pretend Normal	7																
Pretend Abnormal	10																

Definitions

Summary of TRUE Spinoscope Decision: Very few injured individuals are unable to move in a near normal manner when they are not under stress (i.e. loaded). In that case, the Machine specifies that an injured individual can lift "X" kg in a near normal manner. For loads beyond "X" kg, the injury interferes with an otherwise near normal function and the pattern of coordination shifts significantly towards an abnormal status. The decision threshold is taken to be 90%.

A clinician may express the same idea by deciding that a subject is suitable for "light duties". What is meant is that the subject is presumably capable of lifting light loads in a near normal manner, and is presumably incapable of lifting heavier loads safely. The machine description is far more precise in that regard. The TRUE Decision which follows a rigid interpretation protocol is the only method recommended in operating the Spinoscope.

The AVERAGE Distance to Threshold. Measures the distance between the Average probability of normality across all loads and the arbitrary threshold of 90%. In other words, the probability of normality is the Average distance augmented by 90%.

The AVERAGE Decision: Is the average probability of normality for all completed movements. Such an average response (Normal/Abnormal) is a mathematical curiosity which may have no clinical meaning in some situations. This would be the case of subject 130717 who has a near normal response from zero to 12 kg lift and then an abnormal response from 13 to 23 kg lift. The average decision would then be neither normal nor abnormal but borderline, although the TRUE Decision has clearly defined two mutually exclusive zones for normality and abnormality. Hence, this average decision does not characterize what the subject can do and SpineX does not recommend the use of such averages in those particular situations.

Below the AVERAGE Decision, a qualifier (being either MIXED or OUT) is given. MIXED means that there exists a normal and abnormal range of movement. OUT means that the data is not suitable for machine processing.

The Confidence level. Characterizes the confidence in the calculated mathematical Average Decision when the average of the probability of normality is too close to the Expert system's decision threshold of 90%. At the limit, when the Average distance is zero (probability of normality is 90%), it is impossible to determine if the subject is normal or abnormal. The Confidence level is then zero. A Confidence level of one expresses that the Average is sufficiently far from the decision threshold to be acceptable (outside the $\pm 2\%$ zone around the 90% threshold).

Functional Anomalies. Intended to express the nature of pathology from a functional and not an anatomical point of view. The confidence level expresses how strong the pathology appears to the expert system. Even normal individuals may have traces of functional pathology which are deemed sufficiently minor to be unimportant for the range of loads tested. Hence, being normal does not exclude the presence of pathology. It is only when the pathology crosses a certain threshold that the subject is classified abnormal.

The Index "Relative Importance of pathology" expresses that the pathological score is sufficiently large to force the probability of normality to be lowered below the 90% threshold. It must be noted that being classified as "normal" does not imply the absence of pathology. Being normal simply means that the pathology, if present, is contained. This is the case of subject 130717 in whom the pathology is contained up to 12 kg and then outside the normal coordination for heavier loads lifted.

Malingering. The subjects are classified into three groups: honest, cheater, and unclear. A subject deemed cheater or classified as unclear is furthermore assigned a label "pretend normal" if the subject is in fact abnormal but tries to dissimulate his condition, and "Pretend abnormal" if the subject is in fact normal and tries to simulate.

Comments. To assist the clinician in understanding the process, a short description extracted from the expert system reasoning process is given.

Tableau 7

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 90%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Make events which find the sides of the Expert System
102536 M 22 21 JAN 91 Lead 17 kg Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold: 7.8 %	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal: 100 %	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar affixes	☐ Left ☐ Right ☐ Relative importance of pathology 0 % ☐ Do not know ☐ Not enough data	☐ Honest ☐ Cheater ☐ Unclear ☐ Prolong NORMAL ☐ Prolong ABNORMAL
First reading	Clinical Examination Consensus: NN - HH Consistency of conclusions: Yes ☐ No ☐ Partial ☐ Do not know ☐	True Decision UNCHANGED MATCH: 1 4 5 Clinician's examination: Invalid Input selected for ROC calculations: Clinician	Second Reading of Average Decision ☐ Normal ☐ Normal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure Test complete ? Yes		
Second reading					
110139 F 45 17 DEC 90 Lead 17 kg Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold: 5.0 %	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal: 100 %	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar affixes	☐ Left ☐ Right ☐ Relative importance of pathology 0 % ☐ Do not know ☐ Not enough data	☐ Honest ☐ Cheater ☐ Unclear ☐ Prolong NORMAL ☐ Prolong ABNORMAL
First reading	Clinical Examination Consensus: U - H Consistency of conclusions: Yes ☐ No ☐ Partial ☐ Do not know ☐	True Decision UNCHANGED MATCH: 1 4 5 Clinician's examination: Invalid Input selected for ROC calculations: Spine	Second Reading of Average Decision ☐ Normal ☐ Normal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure Test complete ? Yes		
Second reading					
115803 M 21 05 MAR 90 Lead 17 kg Max 17 kg	Normal up to 17 kg. No data beyond that load. Average distance to threshold: 1.0 %	First Reading Normal / Borderline Valid Average ? Yes Confidence level separating normal from abnormal: 47 %	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar affixes	☐ Left ☐ Right ☐ Relative importance of pathology 33 % ☐ Do not know ☐ Not enough data	☐ Honest ☐ Cheater ☐ Unclear ☐ Prolong NORMAL ☐ Prolong ABNORMAL
First reading	Clinical Examination Consensus: U - U Consistency of conclusions: Yes ☐ No ☐ Partial ☐ Do not know ☐	True Decision UNCHANGED MATCH: 1 4 5 Clinician's examination: Invalid Input selected for ROC calculations: Spine	Second Reading of Average Decision ☐ Normal ☐ Normal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure Test complete ? Yes		Restrictions on the left middle and lower spine compensated by a relatively hypermobile T1 junction. Borderline cases because the pathology is well compensated but the clinician may think otherwise.
Second reading					

MATCH CODES

- 1 Full agreement.
- 2 Definitive since clinicians are "indecisive".
- 3 Definitive since clinicians are divided.
- 4 True Spinoscope decision first reading low weights.
- 5 Definitive since at least one clinician flipped.

ERROR CODES

- 1 Correctable error
- 2 Serious disagreement

Definitions

- For the physical examination
- ND Dynamic exam Normal
 17 No clinical data recorded

- For the individual conclusions
- HH Certainly Normal
 N Probably Normal
 U Undecided
 A(b) Abnormal borderline
 A Probably Abnormal
 AA Certainly Abnormal

There is a bias in the dynamic portion of the official examination: Dr Fortin finds more pathological cases than Dr Beaulieu. This bias is somewhat compensated by history and pain in the official conclusions. In general, the expert system agrees with Dr Beaulieu's decision. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not physicians. The Spinoscope exceeding function without pain or history, is therefore bound to agree more with Dr Beaulieu.

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 80%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which lead the rules of the Expert System
120745 F 20 22APR01 Lead 20 kg	Normal up to 3 kg, then abnormal up to 23 kg. Average distance to threshold -8.9%	First Reading: Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100%	Left ☐ Right ☐ Relative importance of pathology 100 % ☐ Do not know ☐ Not enough data Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Spinal ☐ All lumbar afflicts	☐ Honest ☐ Cheater ☐ Unclear ☐ Prolong NORMAL ☐ Prolong ABNORMAL	Highly inconsistent in axial torsion to the left. Consider disorientation. Minor sprain.
First reading					
Second reading					
130717 F 32 18APR01 Lead 20 kg	Normal up to 12 kg, then abnormal up to 23 kg. Average distance to threshold 2.0%	First Reading: Normal Valid Average ? Yes Confidence level separating normal from abnormal 78%	Left ☐ Right ☐ Relative importance of pathology 17 % ☐ Do not know ☐ Not enough data Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Spinal ☐ All lumbar afflicts	☐ Honest ☐ Cheater ☐ Unclear ☐ Prolong NORMAL ☐ Prolong ABNORMAL	Phylogeny is effective on function at a level of stress corresponding to loads exceeding 8 kg which is close to the conditions of the clinical examination. Mild torsional injury to the left at 4/5. Weak evidence of disorientation.
First reading					
Second reading					
134107 M 21 31APR01 Lead 10 kg	Normal unloaded and abnormal thereafter Average distance to threshold -12.6%	First Reading: Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100%	Left ☐ Right ☐ Relative importance of pathology 100 % ☐ Do not know ☐ Not enough data Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Spinal ☐ All lumbar afflicts	☐ Honest ☐ Cheater ☐ Unclear ☐ Prolong NORMAL ☐ Prolong ABNORMAL	Minor hypermobility at T1 junction. Irregular response suggested mild sprain.
First reading					
Second reading					

For the individual conclusions

MM	Certainly Normal	MM	Honest
U	Probably Normal	MM	Certainly honest
U	Undecided		
Ab	Abnormal borderline		
A	Probably Abnormal		
AA	Certainly Abnormal		

For the physical examination

ND	Dynamic exam Normal
??	No clinical data recorded

Definitions

ERROR CODES

1	Correctable error
2	Serious disagreement

MATCH CODES

1	Full agreement
2	Default since clinicians are "indecisive"
3	Default since clinicians are divided
4	True Spinoscope decision first reading low weights
5	Default since at least one clinician flipped

There is a bias in the dynamic portion of the clinical examination: Dr Fertis finds more pathological cases than Dr Bover. This bias is somewhat compensated by history and pain in the clinical examination. In general, the expert system agrees with Dr Bover's decision. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not physicians. The Spinoscope exceeding function without pain or history, is therefore based to agree more with Dr Bover.

Tuesday, December 13, 1994 11:06

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of probability 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 90%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which find the sides of the Expert System
136375 M 20 21JAN91 Lead 11kg Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 6.6%	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%	☒ Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (radio, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness	☒ Left ☒ Right Relative importance of pathology 0 % ☒ Do not know ☒ Not enough data	☒ Honest ☒ Cheater ☒ Unsure ☒ Pretend NORMAL ☒ Pretend ABNORMAL
First reading					
Second reading					
136637 F 20 21JAN92 Lead 11kg Max 2 kg	Remove from population Average distance to threshold 9.6%	First Reading Abnormal / Borderline OUT Valid Average ? Doubtful Confidence level separating normal from abnormal 4%	☒ Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (radio, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness	☒ Left ☒ Right Relative importance of pathology 0 % ☒ Do not know ☒ Not enough data	☒ Honest ☒ Cheater ☒ Unsure ☒ Pretend NORMAL ☒ Pretend ABNORMAL
First reading					No reasons to stop lifting more. Probably normal.
Second reading					
141132 F 30 21NOV90 Lead 11kg Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 9.6%	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%	☒ Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (radio, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness	☒ Left ☒ Right Relative importance of pathology 0 % ☒ Do not know ☒ Not enough data	☒ Honest ☒ Cheater ☒ Unsure ☒ Pretend NORMAL ☒ Pretend ABNORMAL
First reading					
Second reading					

MATCH CODES	
1	Full agreement.
2	Default since clinicians are "Indecisive".
3	Default since clinicians are divided.
4	True Spinoscope decision first reading low weight.
5	Default since at least one clinician flipped.

ERROR CODES	
1	Correctable error
2	Serious disagreement

Definitions	
For the physical examination	ND Dynamic exam Normal 77 No clinical data recorded

For the individual conclusions	
HH	Certainly Normal
N	Probably Normal
U	Unclassified
Ab	Abnormal borderline
A	Probably Abnormal
AA	Certainly Abnormal

There is a bias in the dynamic portion of the clinical examination: Dr Fortin finds more pathological cases than Dr Boerter. This bias is somewhat compensated by history and pain in the clinical conclusions. In general, the expert system agrees with Dr Boerter's decision. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not physicians. The Spinoscope exceeding function without pain or history, is therefore based to agree more with Dr Boerter.

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 80%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which lead the value of the Expert System
145162 U 25 21JAO1 Lead 10kg Max 31 kg	Normal unloaded and abnormal thereafter Average distance to threshold -0.9 %	First Reading Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	Thoracic Lumbar Junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (radio, overweight) Potential conflict clinicians Spinal All lumbar stiffness	Left Right Relative importance of pathology 100 % Do not know Not enough data	Was downgraded because of TL restriction. The performance increases with loads (very good at 31 kg). May be apprehensive at first. Essentially normal in spite of the TL restriction. Candidate for ablation.
First reading	Clinical Examination Individual Dexter AA A(b) - U Sinister AA A(b) - U Consensus A(b) - U	Consistency of conclusions Yes No Diagnose No Flip Partial Do not know	True Decision UNCHANGED MATCH 1 4 2 5 3	Second Reading of Average Decision Normal Borderline Abnormal Borderline Abnormal Remove from population Unsure Test complete ? Yes	A trunk velocity of 61% when lifting 31 kg is a good sign of collaboration.
Second reading	For the individual conclusions NH Certainly Normal H Honest N Probably Normal HH Certainly honest U Undecided	For the physical examination ND Dynamic exam Normal T? No clinical data recorded	ERROR 1 4 2 5 3	ERROR CODES 1 Correctable error 2 Serious disagreement	
154014 F 30 17MA000 Lead 10kg Max 8 kg	Abnormal Average distance to threshold -0.9 %	First Reading Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	Thoracic Lumbar Junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (radio, overweight) Potential conflict clinicians Spinal All lumbar stiffness	Left Right Relative importance of pathology 100 % Do not know Not enough data	Quiet normal unloaded but immediate spin pain at 4.6 kg. There is not enough data to corroborate this finding. Problem for human assessment.
First reading	Clinical Examination Individual Dexter AA N - G Sinister AA N - G Consensus NN - GG	Consistency of conclusions Yes No Diagnose No Flip Partial Do not know	True Decision UNCHANGED MATCH 1 4 2 5 3	Second Reading of Average Decision Normal Borderline Abnormal Borderline Abnormal Remove from population Unsure Test complete ? No	Conclusion reverses the clinical observations since the subject "Took OK". Spinoscope agrees with both Fortin and Bourlier initial assessment.
Second reading	For the individual conclusions NH Certainly Normal H Honest N Probably Normal HH Certainly honest U Undecided	For the physical examination ND Dynamic exam Normal T? No clinical data recorded	ERROR 1 4 2 5 3	ERROR CODES 1 Correctable error 2 Serious disagreement	
167427 F 30 17MA000 Lead 10kg Max 12 kg	Normal up to 12 kg. No data beyond that load. Average distance to threshold 3.9 %	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	Thoracic Lumbar Junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (radio, overweight) Potential conflict clinicians Spinal All lumbar stiffness	Left Right Relative importance of pathology 0 % Do not know Not enough data	The Spinoscope detected lateral bending anomaly was not strong enough to force the expert system to declare the subject abnormal.
First reading	Clinical Examination Individual Dexter AA A(b) - H Sinister A A(b) - HH Consensus A(b) - HH	Consistency of conclusions Yes No Diagnose No Flip Partial Do not know	True Decision UNCHANGED MATCH 1 4 2 5 3	Second Reading of Average Decision Normal Borderline Abnormal Borderline Abnormal Remove from population Unsure Test complete ? Half	
Second reading	For the individual conclusions NH Certainly Normal H Honest N Probably Normal HH Certainly honest U Undecided	For the physical examination ND Dynamic exam Normal T? No clinical data recorded	ERROR 1 4 2 5 3	ERROR CODES 1 Correctable error 2 Serious disagreement	

MATCH CODES

- 1 Full agreement.
- 2 Default since clinicians are "indecisive".
- 3 Default since clinicians are divided.
- 4 True Spinoscope decision first reading low weights.
- 5 Default since at least one clinician flipped.

ERROR CODES

- 1 Correctable error
- 2 Serious disagreement

Definitions

There is a bias in the dynamic portion of the clinical examination: Dr Fortin finds more pathological cases than Dr Bourlier. This bias is somewhat compensated by history and pain in the clinical conclusions. In general, the expert system agrees with Dr Bourlier's decision. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not physiotherapists. The Spinoscope assessing function without pain or history, is therefore based to agree more with Dr Bourlier.

For the individual conclusions
NH Certainly Normal H Honest
N Probably Normal HH Certainly honest
U Undecided

For the physical examination
ND Dynamic exam Normal
T? No clinical data recorded

Tuesday, December 13, 1994 11:05

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 99%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which led the sides of the Expert System
160370 F 30 28/10/91 Load lifted Max 23 kg	Normal up to 3 kg, then abnormal up to 23 kg. Average distance to threshold -2.9%	First Reading Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent fixation ☐ Poor data (yello, overweight) ☐ Potential corralled clinicians ☐ Sprain ☐ All lumbar stiffness	☒ Honest ☐ Cheater ☐ Unclear ☐ Preferred NORMAL ☐ Preferred ABNORMAL	Minor right scoliosis. Serious drop of performance at 23 kg suggesting that this load is close to her limits.
First reading	True Decision UNCHANGED	Second Reading of Average Decision	☐ Left ☐ Right Relative importance of pathology 95 % ☐ Do not know ☐ Not enough data	☐ Honest ☐ Cheater ☐ Unclear ☐ Preferred NORMAL ☐ Preferred ABNORMAL	Clinician detected a minor pathology @ 45%. The True Spinoscope decision specifies that the pathology is effective for loads greater than 3 kg. Spinoscope and clinicians agree on the outcome for the conditions of the physical examination (Le zero touch). Excellent illustration of the inappropriateness of compressing the data to get a normal/abnormal average which is not representative of the clinical situation.
Second reading	Clinical Examination	Consensus	☐ Individual ☐ Consensus N - U	☐ Left ☐ Right Relative importance of pathology 95 % ☐ Do not know ☐ Not enough data	☐ Honest ☐ Cheater ☐ Unclear ☐ Preferred NORMAL ☐ Preferred ABNORMAL
160714 F 47 18/02/90 Load lifted Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 10.9%	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal 99%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent fixation ☐ Poor data (yello, overweight) ☐ Potential corralled clinicians ☐ Sprain ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Preferred NORMAL ☐ Preferred ABNORMAL	Very strong female.
First reading	True Decision UNCHANGED	Second Reading of Average Decision	☐ Left ☐ Right Relative importance of pathology 8 % ☐ Do not know ☐ Not enough data	☐ Honest ☐ Cheater ☐ Unclear ☐ Preferred NORMAL ☐ Preferred ABNORMAL	Clinician detected a minor pathology @ 45%. The True Spinoscope decision specifies that the pathology is effective for loads greater than 3 kg. Spinoscope and clinicians agree on the outcome for the conditions of the physical examination (Le zero touch). Excellent illustration of the inappropriateness of compressing the data to get a normal/abnormal average which is not representative of the clinical situation.
Second reading	Clinical Examination	Consensus	☐ Individual ☐ Consensus NN - HH	☐ Left ☐ Right Relative importance of pathology 8 % ☐ Do not know ☐ Not enough data	☐ Honest ☐ Cheater ☐ Unclear ☐ Preferred NORMAL ☐ Preferred ABNORMAL
162132 M 25 05/04/93 Load lifted Max 23 kg	Normal unloaded and abnormal thereafter Average distance to threshold -1.5%	First Reading Abnormal / Borderline Valid Average ? Yes Confidence level separating normal from abnormal 80%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent fixation ☐ Poor data (yello, overweight) ☐ Potential corralled clinicians ☐ Sprain ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Preferred NORMAL ☐ Preferred ABNORMAL	Very mild torsional injury coupled with a consistent upper lumbar stiffness lowered the scores below the threshold. Clinicians may not detect this anomaly and see this subject as normal.
First reading	True Decision UNCHANGED	Second Reading of Average Decision	☐ Left ☐ Right Relative importance of pathology 75 % ☐ Do not know ☐ Not enough data	☐ Honest ☐ Cheater ☐ Unclear ☐ Preferred NORMAL ☐ Preferred ABNORMAL	Clinician detected a minor pathology @ 45%. The True Spinoscope decision specifies that the pathology is effective for loads greater than 3 kg. Spinoscope and clinicians agree on the outcome for the conditions of the physical examination (Le zero touch). Excellent illustration of the inappropriateness of compressing the data to get a normal/abnormal average which is not representative of the clinical situation.
Second reading	Clinical Examination	Consensus	☐ Individual ☐ Consensus AA - HH	☐ Left ☐ Right Relative importance of pathology 75 % ☐ Do not know ☐ Not enough data	☐ Honest ☐ Cheater ☐ Unclear ☐ Preferred NORMAL ☐ Preferred ABNORMAL

Definitions

For the physical examination

ND Dynamic exam Normal
 ?? No clinical data recorded

For the individual conclusions

NN	Certainly Normal	N	Honest
N	Probably Normal	HH	Certainly honest
U	Unclassified		
Ab	Abnormal borderline		
A	Probably Abnormal		
AA	Certainly Abnormal		

ERROR CODES

- 1 Correctable error
- 2 Serious disagreement

MATCH CODES

- 1 Full agreement.
- 2 Default since clinicians are "Indecisive".
- 3 Default since clinicians are divided.
- 4 True Spinoscope decision first reading low weights.
- 5 Default since at least one clinician flipped.

There is a bias in the dynamic portion of the clinical examination: Dr Fortin finds more pathological cases than Dr Beaulieu. This bias is somewhat compensated by history and pain in the clinical conclusions. In general, the expert system agrees with Dr Beaulieu's decision. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not physicians. The Spinoscope exceeding function without pain or history, is therefore

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 90%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which lead the value of the Expert System
164315 M 25 17JUN91 Load 31 kg Max 31 kg	Normal up to 31 kg. No data beyond that load.	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%	Thoracic Lumbar Junction Lower lumbar restriction LB torsion injury Excellent function Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	Honest Cheater Unclear Pretend NORMAL Pretend ABNORMAL	Some muscular strategy at high loads.
First reading	Clinical Examination Individual Exam Felt Boulder Consensus AA - HH A(b) - H	Consistency of conclusions Yes No Diagnose Fp Do not know Period	True Decision UNCHANGED ERROR 1 4 2 5 3 Valid Isp selected for ROC calculations Clinician	Second Reading of Average Decision Honest Cheater Unclear Test complete ? Yes	A trunk velocity of 100% when lifting 31 kg is a good sign of normality, although the reduction in ROM and lateral bending data indicate that the subject is approaching some limits for loads above 31 kg, thereby demonstrating the presence of a subtle pathology.
Second reading	Clinical Examination Individual Exam Felt Boulder Consensus AA - HH A(b) - H	Consistency of conclusions Yes No Diagnose Fp Do not know Period	True Decision UNCHANGED ERROR 1 4 2 5 3 Valid Isp selected for ROC calculations Clinician	Second Reading of Average Decision Honest Cheater Unclear Test complete ? Yes	
168693 M 45 17DEC90 Load 31 kg Max 31 kg	Normal up to 31 kg. No data beyond that load.	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%	Thoracic Lumbar Junction Lower lumbar restriction LB torsion injury Excellent function Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	Honest Cheater Unclear Pretend NORMAL Pretend ABNORMAL	
First reading	Clinical Examination Individual Exam Felt Boulder Consensus NN - HH NN - HH	Consistency of conclusions Yes No Diagnose Fp Do not know Period	True Decision UNCHANGED MATCH 1 4 2 5 3 Valid Isp selected for ROC calculations Clinician	Second Reading of Average Decision Honest Cheater Unclear Test complete ? Yes	
Second reading	Clinical Examination Individual Exam Felt Boulder Consensus NN - HH NN - HH	Consistency of conclusions Yes No Diagnose Fp Do not know Period	True Decision UNCHANGED MATCH 1 4 2 5 3 Valid Isp selected for ROC calculations Clinician	Second Reading of Average Decision Honest Cheater Unclear Test complete ? Yes	
174047 M 33 21FEB91 Load 14 kg Max 14 kg	Normal up to 6 kg, then abnormal up to 14 kg.	First Reading Abnormal / Borderline Valid Average ? Yes Confidence level separating normal from abnormal 5%	Thoracic Lumbar Junction Lower lumbar restriction LB torsion injury Excellent function Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	Honest Cheater Unclear Pretend NORMAL Pretend ABNORMAL	Excellent unloaded but sprain and evidence of mild lateral injury at higher loads. Did not complete 23 kg.
First reading	Clinical Examination Individual Exam Felt Boulder Consensus AA - HH AA - HH	Consistency of conclusions Yes No Diagnose Fp Do not know Period	True Decision UNCHANGED MATCH 1 4 2 5 3 Valid Isp selected for ROC calculations Clinician	Second Reading of Average Decision Honest Cheater Unclear Test complete ? Yes	
Second reading	Clinical Examination Individual Exam Felt Boulder Consensus AA - HH AA - HH	Consistency of conclusions Yes No Diagnose Fp Do not know Period	True Decision UNCHANGED MATCH 1 4 2 5 3 Valid Isp selected for ROC calculations Clinician	Second Reading of Average Decision Honest Cheater Unclear Test complete ? Yes	

For the individual conclusions

HH	Certainly Normal	HH	Honest
N	Probably Normal	HH	Certainly
U	Uncertain	HH	honest

A(b) Abnormal borderlines
A Probably abnormal
AA Certainly abnormal

For the physical examination

ND Dynamic exam Normal
TT No clinical data recorded

There is a bias in the dynamic portion of the clinical examination: Dr Fortin finds more pathological cases than Dr Bourcier. This bias is somewhat compensated by history and pain in the clinical conclusions. In general, the expert system agrees with Dr Bourcier's decision. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not physiotherapists. The Spinoscope assessing function without pain or history, is therefore bound to agree more with Dr Bourcier.

Definitions

ERROR CODES

1	Correctable error
2	Serious disagreement

MATCH CODES

1	Full agreement.
2	Default since clinicians are "wedge".
3	Default since clinicians are divided.
4	True Spinoscope decision first reading low weights.
5	Default since at least one clinician stopped.

Tuesday, December 13, 1994 11:05

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 80%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which lead to the value of the Expert System
180713 F 50 210002 Load 10kg Max 0 kg	Remove from population Valid Average ? Doubtful Confidence level separating normal from abnormal: 100%	First Reading: Abnormal Valid Average ? Doubtful Confidence level separating normal from abnormal: 100%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unsure ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Slow sprain pattern. Minor right scoliosis. Possibly honest injured but not enough data to confirm.
First reading					
Second reading					
181696 F 20 007002 Load 10kg Max 14 kg	Normal up to 1 kg, then abnormal up to 14 kg. Average distance to threshold: -7.0%	First Reading: Abnormal Valid Average ? Yes Confidence level separating normal from abnormal: 100%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unsure ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Clear sprain at 14 kg.
First reading					
Second reading					
188593 M 31 22AP01 Load 10kg Max 14 kg	Normal up to 14 kg. No data beyond that load. Average distance to threshold: 0.0%	First Reading: Normal Valid Average ? Yes Confidence level separating normal from abnormal: 100%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unsure ☐ Pretend NORMAL ☐ Pretend ABNORMAL	No reasons to interrupt the 14 kg RTI unless the patient apprehended a flare of pain and aborted before the onset of sprain.
First reading					
Second reading					

Definitions

For the physical examination

ND Dynamic exam Normal
T? No clinical data recorded

For the individual conclusions

NN Certainly Normal
N Probably Normal
U Undecided

A(b) Abnormal Borderline
A Probably Abnormal
AA Certainly Abnormal

ERROR CODES

1 Correctable error
2 Serious disagreement

MATCH CODES

1 Full agreement
2 Default since clinicians are "pessimistic"
3 Default since clinicians are divided
4 True Spinoscope decision first reaching low weights
5 Default since at least one clinician stopped.

There is a bias in the dynamic portion of the clinical examination: Dr Ferlin finds more pathological cases than Dr Boverat. This bias is somewhat compensated by history and pain in the official conclusions. In general, the expert system agrees with Dr Boverat's decision. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not physiatrists. The Spinoscope assessing function without pain or history, is therefore

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 90%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which lead the sides of the Expert System
189409 M 20 OFF001 Load 10kg Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold: 6.6%	First Reading: Normal Valid Average: Yes Confidence level separating normal from abnormal: 100%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (ratio, overweight) ☐ Potential conflict children ☐ Spinal ☐ All lumbar affords ☐ Left ☐ Right Relative importance of pathology: 0 % ☐ Do not know ☐ Not enough data	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Minor left scabule
First reading					
Second reading					
192214 F 20 OFF002 Load 14kg Max 14 kg	Normal up to 14 kg. No data beyond that load. Average distance to threshold: 2.0%	First Reading: Normal Valid Average: Yes Confidence level separating normal from abnormal: 100%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (ratio, overweight) ☐ Potential conflict children ☐ Spinal ☐ All lumbar affords ☐ Left ☐ Right Relative importance of pathology: 0 % ☐ Do not know ☐ Not enough data	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	
First reading					
Second reading					
193673 M 41 27JAN92 Load 8kg Max 8 kg	Normal up to 7 kg, then abnormal up to 8 kg. Average distance to threshold: -1.6%	First Reading: Abnormal / Borderline Valid Average: Mixed Confidence level separating normal from abnormal: 21%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (ratio, overweight) ☐ Potential conflict children ☐ Spinal ☐ All lumbar affords ☐ Left ☐ Right Relative importance of pathology: 6.7 % ☐ Do not know ☐ Not enough data	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Only one hand completed. Spinal pattern detected but no enough data to confirm.
First reading					
Second reading					

For the individual conclusions
 NN Certainly Normal
 N Probably Normal
 U Undecided
 A(b) Abnormal borderline
 A Probably Abnormal
 AA Certainly Abnormal

For the physical examination
 ND Dynamic exam Normal
 ?? No clinical data recorded

Definitions
 There is a bias in the dynamic portion of the clinical examination: Dr Forth finds more pathological cases than Dr Bourcier. This bias is somewhat compensated by history and pain in the clinical conclusions. In general, the expert system agrees with Dr Bourcier's decision. This may be due to the fact that the expert system has been trained by orthopedic engineers and not physiatrists. The Spinoscope assessing function without pain or history, is therefore bound to agree more with Dr Bourcier.

ERROR CODES
 1 Correctable error
 2 Serious disagreement

MATCH CODES
 1 Full agreement
 2 Default since children are "Indecisive".
 3 Default since children are divided.
 4 True Spinoscope decision first reading low weights.
 5 Default since at least one children slipped.

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

A78

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 90%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments
207050 F 30 OALJ002 Load 10kg Max 14 kg	Abnormal	First Reading: Mixed Valid Average ? Yes Confidence level depending on normal from abnormal: 100%	Thoracic Lumbar Junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (radio, overweight) Potential conflict clinicians Spinal All lumbar stiffness	Honest Cheater Unclear Pretest NORMAL Pretest ABNORMAL	Possible time out at 23 kg.
First reading	Clinical Examination Consensus: NN - HH Consistency of conclusions: Yes Average distance to threshold: -0.9%	True Decision CHIEF'S examination Valid Input selected for ROC calculations Clinician	Second Reading of Average Decision Normal Normal Borderline Abnormal Borderline Abnormal Remove from population Unclear	Honest Cheater Unclear Test complete ? No	The Spinoscope detected TL pathology was possibly given too much importance by the expert system. The clinicians are probably right.
Second reading	Individual Decision Doctor: ND Perf: ND Backup: ND	Remove from population Valid Average ? Doubtful Confidence level depending on normal from abnormal: 99%	Thoracic Lumbar Junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (radio, overweight) Potential conflict clinicians Spinal All lumbar stiffness	Honest Cheater Unclear Pretest NORMAL Pretest ABNORMAL	No reasons to refuse lifting. His traces of functional injury which cannot be confirmed because of lack of data. Must be removed from population since testing protocol was not followed and no data recovery is possible.
210884 F 30 OALJ002 Load 10kg Max 0 kg	Remove from population	First Reading: OUT Valid Average ? Doubtful Confidence level depending on normal from abnormal: 99%	Thoracic Lumbar Junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (radio, overweight) Potential conflict clinicians Spinal All lumbar stiffness	Honest Cheater Unclear Pretest NORMAL Pretest ABNORMAL	Subject did not collaborate at all and hence was classified as cheater by the Spinoscope. The evidence presented by the clinicians indicate that this lack of collaboration may be due to pathology. Not sure.
First reading	Clinical Examination Consensus: AA - HH Consistency of conclusions: Yes Average distance to threshold: -3.9%	True Decision CHIEF'S examination Valid Input selected for ROC calculations Clinician	Second Reading of Average Decision Normal Normal Borderline Abnormal Borderline Abnormal Remove from population Unclear	Honest Cheater Unclear Test complete ? No	
Second reading	Individual Decision Doctor: AA Perf: AA Backup: AA	Match 1 2 3 4 5	Thoracic Lumbar Junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (radio, overweight) Potential conflict clinicians Spinal All lumbar stiffness	Honest Cheater Unclear Pretest NORMAL Pretest ABNORMAL	
212536 F 30 OALJ002 Load 10kg Max 23 kg	Normal up to 23 kg. No data beyond that load.	First Reading: Yes Valid Average ? Yes Confidence level depending on normal from abnormal: 100%	Thoracic Lumbar Junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (radio, overweight) Potential conflict clinicians Spinal All lumbar stiffness	Honest Cheater Unclear Pretest NORMAL Pretest ABNORMAL	
First reading	Clinical Examination Consensus: NN - HH Consistency of conclusions: Yes Average distance to threshold: 9.9%	True Decision CHIEF'S examination Valid Input selected for ROC calculations Clinician	Second Reading of Average Decision Normal Normal Borderline Abnormal Borderline Abnormal Remove from population Unclear	Honest Cheater Unclear Test complete ? Yes	
Second reading	Individual Decision Doctor: ND Perf: ND Backup: ND	Match 1 2 3 4 5	Thoracic Lumbar Junction Lower lumbar restriction LB torsion injury Excellent flexion Poor data (radio, overweight) Potential conflict clinicians Spinal All lumbar stiffness	Honest Cheater Unclear Pretest NORMAL Pretest ABNORMAL	

For the individual conclusions

NN	Certainly Normal	H	Honest
N	Probably Normal	HH	Certainly
U	Unclassified		

A(b) Abnormal borderline
A Probably Abnormal
AA Certainly Abnormal

For the physical examination

ND Dynamic exam Normal
?? No clinical data recorded

There is a bias in the dynamic portion of the clinical examination: Dr Fortin finds more pathological cases than Dr Boerler. This bias is accounted compensated by history and pain in the clinical examinations. In general, the expert system agrees with Dr Boerler's decision. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not physiotherapists. The Spinoscope assessing function without pain or history, is therefore bound to agree more with Dr Boerler.

Definitions

1 Correctable error
2 Serious disagreement

ERROR CODES

1 Full agreement
2 Default since clinicians are "Indecisive".
3 Default since clinicians are divided.
4 True Spinoscope decision first reading low weights.
5 Default since at least one clinician flipped.

MATCH CODES

1 Full agreement
2 Default since clinicians are "Indecisive".
3 Default since clinicians are divided.
4 True Spinoscope decision first reading low weights.
5 Default since at least one clinician flipped.

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 96%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments
219120 F 31 30A000 Lead 1000 Max 20 kg	Normal up to 32 kg. No data beyond that load. Valid Average ? Yes Confidence level separating normal from abnormal 15 %	First Reading Normal / Borderline Valid Average ? Yes Confidence level separating normal from abnormal 15 %	☒ Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (redo, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar effusions ☐ Left ☐ Right Relative importance of pathology 40 % ☐ Do not know ☐ Not enough data ☐ Normal ☐ Normal Borderline ☐ Abnormal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure Test complete ? Yes	☐ Honest ☐ Cheater ☐ Unsure ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Atrial relation extremely degraded on otherwise excellent and consistent performance in fiction. Very difficult anomaly to detect by clinicians. May be candidate for declassification. Since right on the border, the decision can go either way.
3 Second reading					The clinicians themselves disagree. Trend to normality. A trunk velocity of 110°/s when lifting 32 kg is a good sign of normality. This will be consistent with both physical examinations which were degraded by the clinicians.
237918 F 35 19A000 Lead 1000 Max 5 kg	Normal up to 3 kg, then abnormal up to 5 kg. Valid Average ? Yes Confidence level separating normal from abnormal 27 %	First Reading Abnormal / Borderline Valid Average ? Yes Confidence level separating normal from abnormal 27 %	☒ Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (redo, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar effusions ☐ Left ☐ Right Relative importance of pathology 67 % ☐ Do not know ☐ Not enough data ☐ Normal ☐ Normal Borderline ☐ Abnormal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure Test complete ? No	☐ Honest ☐ Cheater ☐ Unsure ☐ Pretend NORMAL ☐ Pretend ABNORMAL	The trend is towards pathology as load increases. No reasons to stop lifting and not enough data collected to confirm.
3 Second reading					
259131 F 47 01A000 Lead 1000 Max 5 kg	Normal unloaded and abnormal thereafter Valid Average ? Yes Confidence level separating normal from abnormal 100 %	First Reading Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	☒ Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (redo, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar effusions ☐ Left ☐ Right Relative importance of pathology 100 % ☐ Do not know ☐ Not enough data ☐ Normal ☐ Normal Borderline ☐ Abnormal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure Test complete ? No	☐ Honest ☐ Cheater ☐ Unsure ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Torsional problem unconfirmed.
3 Second reading					

For the individual conclusions

certainty	Normal	H	Honest
M	Probably Normal	H	Certainly
U	Undecided	H	honest

For the physical examination

ND	Dynamic exam Normal
17	No clinical data recorded

Definitions

There is a bias in the dynamic portion of the clinical examination: Dr Fortin finds more pathological cases than Dr Bourcier. This bias is somewhat compensated by history and pain in the clinical conclusions. In general, the expert system agrees with Dr Bourcier's decision. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not physiotherapists. The Spinoscope assessing function without pain or history, is therefore bound to agree more with Dr Bourcier.

ERROR CODES

1	Correctable error
2	Severe disagreement

MATCH CODES

1	Full agreement
2	Default since clinicians are "indecisive"
3	Default since clinicians are divided
4	True Spinoscope decision first reading low weights
5	Default since at least one clinician stopped

Tuesday, December 13, 1994 11:05

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 95%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which lead the rules of the Expert System
301215 M 23 14NOV90 Lead 14 kg	Normal up to 3 kg, then abnormal up to 14 kg. Average distance to threshold -4.8%	First Reading Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100%	Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unsure ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Good unloaded, bad loaded. Sprain. Minor right scoliosis.
5 First reading					
5 Second reading					
329459 F 28 14NOV90 Lead 14 kg	Normal up to 14 kg. No data beyond that load. Average distance to threshold 2.6%	True Decision UNCHANGED ERROR 1 4 2 5 3 Clinician's examination Invalid Input selected for ROC calculations Spino	Second Reading of Average Decision ☐ Normal ☐ Normal Borderline ☐ Abnormal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure	☐ Honest ☐ Cheater ☐ Unsure ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Borderline found minor ROM problem but ignored it. Spinoscope established that Borderline is right and that the problem he detected affects function for loads above 3 kg.
5 First reading					
5 Second reading					
333294 F 23 23JAN91 Lead 14 kg	Normal up to 14 kg. No data beyond that load. Average distance to threshold 7.9%	True Decision UNCHANGED ERROR 1 4 2 5 3 Clinician's examination Valid Input selected for ROC calculations Clinician	Second Reading of Average Decision ☐ Normal ☐ Normal Borderline ☐ Abnormal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure	☐ Honest ☐ Cheater ☐ Unsure ☐ Pretend NORMAL ☐ Pretend ABNORMAL	The functionality decreases with loads and at 14 kg is close to borderline (one per cent above the decision threshold of 80%). There is indeed a minor sprain detected interfering with function as early as the 5 kg RL. Yet it was not enough to have the expert system downgraded it to abnormal status. The clinicians are probably right.
5 First reading					
5 Second reading					

For the individual conclusions	Definitions	ERROR CODES	MATCH CODES
NN Certainly Normal N Probably Normal U Undecided Ab Normal borderline A Probably Abnormal AA Certainly Abnormal	For the physical examination ND Dynamic exam Normal ?? No clinical data recorded There is a bias in the dynamic portion of the clinical examination: Dr Fortin finds more pathological cases than Dr Bourcier. This bias is somewhat compensated by history and pain in the clinical examinations. In general, the expert system agrees with Dr Bourcier's decision. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not physiatrists. The Spinoscope assessing function without pain or history, is therefore bound to agree more with Dr Bourcier.	1 Correctable error 2 Serious disagreement	1 Full agreement 2 Default since clinicians are "Indecisive" 3 Default since clinicians are divided 4 True Spinoscope decision first reading low weights. 5 Default since at least one clinician flipped.

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 95%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which led the idea of the Expert System
342623	Normal up to 10 kg, then abnormal up to 31 kg. Average distance to threshold: -2.0 %	First Reading Abnormal Valid Average ? Doubtful Confidence level separating normal from abnormal: 76 %	☒ Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (radio, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar affrays	☐ Honest ☐ Cheater ☒ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	The 31 kg IM was very poor. Subject exceeded his limits. Possible demeritization.
First reading					
Second reading					
349110	Normal up to 31 kg. No data beyond that load. Average distance to threshold: 3.0 %	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal: 100 %	☒ Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (radio, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar affrays	☐ Honest ☐ Cheater ☒ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Probably normal but atypical. Very small indication of sprain at high load. Probably not significant.
First reading					
Second reading					
355106	Normal up to 5 kg. No data beyond that load. Average distance to threshold: 10.0 %	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal: 100 %	☒ Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (radio, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar affrays	☐ Honest ☐ Cheater ☒ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Time out at 14 kg.
First reading					
Second reading					

For the individual conclusions

WH	Certainty	Normal	H	Honest
N	Probably	Normal	HH	Certainty
U	Undecided			

A(b) Abnormal borderline
A Probably Abnormal
AA Certainly Abnormal

For the physical examination

ND Dynamic exam Normal
77 No clinical data recorded

There is a bias in the dynamic portion of the clinical examination: Dr Fortin finds more pathologic areas than Dr Boeris. This bias is somewhat compensated by history and pain in the clinical examination. In general, the expert system agrees with Dr Boeris's decision. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not physicians. The Spinoscope assessing function without pain or history, is therefore bound to agree more with Dr Boeris.

Definitions

ERROR CODES

1	Correctable error
2	Serious disagreement

MATCH CODES

1	Full agreement
2	Default since clinicians are "Indeter"
3	Default since clinicians are divided
4	True Spinoscope decision first reading low weights
5	Default since at least one clinician flipped

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with bordering zone of plus-minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 95%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the RROST	Functional Anomalies (Pathology)	Malingering	Comments
378181 F 40 VAL002 Lead 100g Max 8 kg	Normal up to 8 kg. No data beyond that load. Average distance to threshold 2.0%	First Reading Normal Valid Average ? Yes Confidence level exceeding normal item threshold 99%	Thoracic Lumbar Junction ☐ Normal ☐ Lower lumbar restriction ☐ LB lumbar injury ☐ Excellent flexion ☐ Poor data (buds, overweight) ☐ Potential conflict difficulties ☐ Sprain ☐ All lumbar effusions Left ☐ Right ☐ Relative importance of pathology 17 % ☐ Do not know ☐ Not enough data	☐ Honest ☐ Cheater ☐ Unsure ☐ Prolonged NORMAL ☐ Prolonged ABNORMAL	Highly inconsistent. Suspect simulation. No solid injury detected.
383728 F 27 SLA001 Lead 100g Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 3.0%	First Reading Normal Valid Average ? Yes Confidence level exceeding normal item threshold 100%	Thoracic Lumbar Junction ☐ Normal ☐ Lower lumbar restriction ☐ LB lumbar injury ☐ Excellent flexion ☐ Poor data (buds, overweight) ☐ Potential conflict difficulties ☐ Sprain ☐ All lumbar effusions Left ☐ Right ☐ Relative importance of pathology 0 % ☐ Do not know ☐ Not enough data	☐ Honest ☐ Cheater ☐ Unsure ☐ Prolonged NORMAL ☐ Prolonged ABNORMAL	Show at 31 kg but otherwise normal. May have reach her limits.
388513 F 29 100380 Lead 100g Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 3.0%	First Reading Normal Valid Average ? Yes Confidence level exceeding normal item threshold 100%	Thoracic Lumbar Junction ☐ Normal ☐ Lower lumbar restriction ☐ LB lumbar injury ☐ Excellent flexion ☐ Poor data (buds, overweight) ☐ Potential conflict difficulties ☐ Sprain ☐ All lumbar effusions Left ☐ Right ☐ Relative importance of pathology 0 % ☐ Do not know ☐ Not enough data	☐ Honest ☐ Cheater ☐ Unsure ☐ Prolonged NORMAL ☐ Prolonged ABNORMAL	A trunk velocity of 487% when lifting 31 kg is average. Any pathology detected cannot be that significant.

Definitions

For the physical examination	
NO	Dynamic exam Normal
??	No ethical data recorded

There is a bias in the dynamic portion of the clinical examination: Dr Pavitt finds more pathological cases than Dr Cropper. This bias is counterbalanced by history and pain in the clinical conditions. In general, the expert opinions agree with Dr Bessie's decision. This may be due to the fact that the expert opinion has been influenced by orthopaedic surgeons and not physicians. The Spinalscope assessing function without pain or history, is therefore bound to agree more with Dr Bessie.

For the individual conclusions

MM	Certainly Probably Probably Undecided	Normal Normal Normal Normal	MM	Need Certainty Need Need
MM				
A				
AA				

ERROR CODES

1	Correctable error
2	Perious disagreement

MATCH CODES

- 1 Full agreement.
- 2
- 3 Default stress clitics are "subsets".
- 4
- 5 Default stress clitics are divided.
- 6
- 7 True Spinocephale decision that reaching low weights.
- 8
- 9 Default stress at least one clitic misused.

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 80%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which lead the value of the Expert System
392590	Normal up to 35 Kg. No data beyond that load.	First Reading Normal / Borderline Valid Average ? Yes Confidence level separating normal from abnormal 4.2 %	☒ Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB tendon injury ☒ Excellent flexion ☒ Poor data (redo, overweight) ☒ Potential corded clinicians ☒ Sprain ☒ All lumbar effusions	☒ Honest ☒ Cheater ☒ Unsure ☒ Pretend NORMAL ☒ Pretend ABNORMAL	Minor upper spine effusions.
First reading					
Second reading					
397947	Abnormal	First Reading Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 10.0 %	☒ Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB tendon injury ☒ Excellent flexion ☒ Poor data (redo, overweight) ☒ Potential corded clinicians ☒ Sprain ☒ All lumbar effusions	☒ Honest ☒ Cheater ☒ Unsure ☒ Pretend NORMAL ☒ Pretend ABNORMAL	Sprain pattern.
First reading					
Second reading					
426336	Normal unloaded and abnormal thereafter	First Reading Abnormal / Borderline Valid Average ? Yes Confidence level separating normal from abnormal 2.1 %	☒ Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB tendon injury ☒ Excellent flexion ☒ Poor data (redo, overweight) ☒ Potential corded clinicians ☒ Sprain ☒ All lumbar effusions	☒ Honest ☒ Cheater ☒ Unsure ☒ Pretend NORMAL ☒ Pretend ABNORMAL	Half of the lateral bend data is missing and not recoverable. Consider removing this subject from the population since no information on functional pathology is detectable.
First reading					
Second reading					

For the individual conclusions

HH	Certainly Normal	N	Normal
HN	Probably Normal	HH	Certainly
U	Unloaded	HH	Normal

At(h) Abnormal borderline
A Probably abnormal
AA Certainly abnormal

For the physical examination

ND	Dynamic exam Normal
T?	No clinical data recorded

There is a bias in the dynamic portion of the clinical examination: Dr Forth reads more pathological cases than Dr Benier. This bias is somewhat compensated by history and pain in the clinical conclusions. In general, the expert system agrees with Dr Benier's decision. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not physicians. The Spinoscope ascending function without pain or history, is therefore based to agree more with Dr Benier.

Definitions

1	Correctable error
2	Serious disagreement

ERROR CODES

1	Full agreement.
2	Default since clinicians are "Indecisive".
3	Default since clinicians are divided.
4	True Spinoscope decision first reading low weights.
5	Default since at least one clinician flipped.

MATCH CODES

1	Full agreement.
2	Default since clinicians are "Indecisive".
3	Default since clinicians are divided.
4	True Spinoscope decision first reading low weights.
5	Default since at least one clinician flipped.

A84

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 90%	Average Decision This methodical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which lead the value of the Expert System
432023 F 21 22HO00 Load lifted Max 14 kg	Normal up to 6 kg, then abnormal up to 14 kg. Average distance to threshold -1.0 %	First Reading Abnormal / Borderline Mixed Valid Average ? Yes Confidence level separating normal from abnormal 9 %	☐ Thoracic Lumbar junction ☐ Lower lumbar realignment ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Spinal ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Border of 14 kg then 6 kg. TL realignment is correct.
3	Clinical Examination Individual Doctor: [blank] Exam: [blank] Consensus: A(b)-HH A(b)-HH	True Decision UNCHANGED MATCH ☐ 1 ☐ 4 ☐ 2 ☐ 5 ☐ 3	Second Reading of Average Decision ☐ Normal ☐ Normal Borderline ☐ Abnormal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure	☐ Honest ☐ Cheater ☐ Unclear ☐ Not enough data Test complete ? Half	✓
439445 M 28 04CE01 Load lifted Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 7.9 %	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	☐ Thoracic Lumbar junction ☐ Lower lumbar realignment ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Spinal ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	The clinicians themselves disagree. A trunk velocity of 76% when lifting 31 kg is a good sign of normality.
5	Clinical Examination Individual Doctor: [blank] Exam: [blank] Consensus: U - U U - U	True Decision UNCHANGED MATCH ☐ 1 ☐ 4 ☐ 2 ☐ 5 ☐ 3	Second Reading of Average Decision ☐ Normal ☐ Normal Borderline ☐ Abnormal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure	☐ Honest ☐ Cheater ☐ Unclear ☐ Not enough data Test complete ? Yes	✓
444669 M 20 12CE00 Load lifted Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 7.0 %	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	☐ Thoracic Lumbar junction ☐ Lower lumbar realignment ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Spinal ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Minor right lumbar sections, 31 kg close to the limit. Minor axial anomaly.
7	Clinical Examination Individual Doctor: [blank] Exam: [blank] Consensus: NN - HH NN - HH	True Decision UNCHANGED MATCH ☐ 1 ☐ 4 ☐ 2 ☐ 5 ☐ 3	Second Reading of Average Decision ☐ Normal ☐ Normal Borderline ☐ Abnormal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure	☐ Honest ☐ Cheater ☐ Unclear ☐ Not enough data Test complete ? Yes	✓

For the individual conclusions

NN	Certainly Normal	N	Normal
N	Probably Normal	HH	Certainly honest
U	Undecided		
Ab)	Abnormal borderline		
A	Probably Abnormal		
AA	Certainly Abnormal		

For the physical examination

ND	Dynamic exam Normal
??	No critical data recorded

Definitions

ERROR CODES

1	Correctable error
2	Serious disagreement

MATCH CODES

1	Full agreement.
2	Default since clinicians are "borderline".
3	Default since clinicians are divided.
4	True Spinoscopic decision first reading low weights.
5	Default since at least one clinician slipped.

There is a bias in the dynamic portion of the official examination: Dr Fortin finds more pathological axes than Dr Bevilacqua. This bias is somewhat compensated by history and pain in the clinical examination. In general, the expert system agrees with Dr Bevilacqua's decision. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not physiotherapists. The Spinoscope assessing function without pain or history, is therefore

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 90%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Make marks which lead the notes of the Expert System
475165	Normal up to 9 kg, then abnormal up to 23 kg.	First Reading: Abnormal / Borderline Mixed Valid Average ? Doubtful Confidence level separating normal from abnormal: 2.2 %	☐ Thoracic Lumbar junction ☐ Lower lumbar retraction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness ☐ Left ☐ Right Relative importance of pathology: 0.7 % ☐ Do not know ☐ Not enough data	☐ Honest ☐ Cheater ☐ Unclear	Too fast to have sprain. Contradictions in behavior.
First reading					
Second reading					
488239	Normal up to 31 kg. No data beyond that load.	First Reading: Normal Valid Average ? Yes Confidence level separating normal from abnormal: 7.9 %	☐ Thoracic Lumbar junction ☐ Lower lumbar retraction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness ☐ Left ☐ Right Relative importance of pathology: 1.7 % ☐ Do not know ☐ Not enough data	☐ Honest ☐ Cheater ☐ Unclear	Minor pathology for loads greater than 14 kg.
First reading					
Second reading					
494968	Normal up to 23 kg. No data beyond that load.	First Reading: Normal Valid Average ? Yes Confidence level separating normal from abnormal: 6.6 %	☐ Thoracic Lumbar junction ☐ Lower lumbar retraction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness ☐ Left ☐ Right Relative importance of pathology: 0 % ☐ Do not know ☐ Not enough data	☐ Honest ☐ Cheater ☐ Unclear	23 Kg close to her limits.
First reading					
Second reading					

For the individual conclusions

NN Certainly Normal
 N Probably Normal
 U Undecided
 A(b) Abnormal borderline
 A Probably Abnormal
 AA Certainly Abnormal

For the physical examination

ND Dynamic exam Normal
 ? No clinical data recorded

Definitions

1 Correctable error
 2 Baroque disagreement

ERROR CODES

1 Full agreement
 2 Default since clinicians are "indeed"
 3 Default since clinicians are divided
 4 True Spinoscope decision first reading low weights
 5 Default since at least one clinician slipped

MATCH CODES

1 Full agreement
 2 Default since clinicians are "indeed"
 3 Default since clinicians are divided
 4 True Spinoscope decision first reading low weights
 5 Default since at least one clinician slipped

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

A86

ID #	Summary of True Spinoscope Decision Threshold set at 90%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the RSST	Functional Anomalies (Pathology)	Malingering	Comments
508668 F 22 OFFER01 Lead 14 kg Max 14 kg	Normal up to 6 kg, then abnormal up to 14 kg. Average distance to threshold 9.0%	First Reading Abnormal / Borderline Valid Average ? Yes Confidence level separating normal from abnormal 8%	☒ Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB tension injury ☒ Excellent flexion ☒ Poor data (radio, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness	☒ Honest ☒ Cheater ☒ Unsure ☒ Pretend NORMAL ☒ Pretend ABNORMAL	The whole lumbar spine has minor stiffness.
First reading					
Second reading					
518017 F 20 OFFER02 Lead 29 kg Max 29 kg	Normal unloaded and abnormal thereafter Average distance to threshold 9.0%	First Reading Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100%	☒ Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB tension injury ☒ Excellent flexion ☒ Poor data (radio, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness	☒ Honest ☒ Cheater ☒ Unsure ☒ Pretend NORMAL ☒ Pretend ABNORMAL	The lower lumbar restriction is clear.
First reading					
Second reading					
542595 F 23 17JAN01 Lead 31 kg Max 31 kg	Normal up to 31 kg, No data beyond that load. Average distance to threshold 7.0%	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%	☒ Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB tension injury ☒ Excellent flexion ☒ Poor data (radio, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness	☒ Honest ☒ Cheater ☒ Unsure ☒ Pretend NORMAL ☒ Pretend ABNORMAL	
First reading					
Second reading					

For the individual conclusions	For the physical examination	Definitions	ERROR CODES	MATCH CODES
NN Certainly Normal N Probably Normal U Undecided A(N) Abnormal borderline A Probably Abnormal AA Certainly Abnormal	MD Dynamic exam Normal 77 No clinical data recorded There is a bias in the dynamic portion of the clinical examination: Dr Ferin finds more pathological cases than Dr Boettler. This bias is somewhat compensated by history and pain in the clinical examinations. In general, the expert system agrees with Dr Boettler's decision. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not physiatrists. The Spinoscope assessing function without pain or history, is therefore	1 Correctable error 2 Surface disagreement	1 Full agreement. 2 Default since clinicians are "Indecisive". 3 Default since clinicians are divided. 4 True Spinoscope decision first reading low weights. 5 Default since at least one clinician stopped.	

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of probability 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 99%	Average Decision	Functional Anomalies (pathology)	Malingering	Comments
566534 M 34 OFFD01 Load 20 kg Max 20 kg	Normal up to 23 kg. No data beyond that load. Average distance to threshold 9.9%	First Reading Normal Valid Average ? Yes Confidence level exceeding normal from abnormal 100%	☒ Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (radio, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness	☒ Honest ☒ Cheater ☒ Unclear ☒ Pretend NORMAL ☒ Pretend ABNORMAL	Bourier dynamic physical examination and conclusion differ. Spinoscope agrees with Bourier initial assessment. A trunk velocity of 165% when lifting 20 kg with a range of motion of 117 degrees is a quite spectacular and definitely not consistent with pathology.
First reading	Clinical Examination Consensus A(b)-HH A(b) - H Consistency of assessment ☒ Yes ☒ Disagree ☒ No ☒ Fp ☒ Partial ☒ Do not know	True Decision UNCHANGED MATCH ☒ 01 ☒ 04 ☒ 02 ☒ 05 ☒ 03	Second Reading of Average Decision ☒ Normal ☒ Normal Borderline ☒ Abnormal Borderline ☒ Abnormal ☒ Remove from population ☒ Unsure Test complete ? Yes		
590068 M 23 16JUN92 Load 14 kg Max 14 kg	Normal up to 14 kg. No data beyond that load. Average distance to threshold 1.0%	First Reading Normal / Borderline Valid Average ? Yes Confidence level exceeding normal from abnormal 47%	☒ Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (radio, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness	☒ Honest ☒ Cheater ☒ Unclear ☒ Pretend NORMAL ☒ Pretend ABNORMAL	Lateral bending anomaly not given enough importance by the expert system. At 14 kg the index is 80% which is right on the border. This demonstrates that the pathology will interfere with function for loads exceeding 14 kg.
First reading	Clinical Examination Consensus AA - HH AA - HH Consistency of assessment ☒ Yes ☒ Disagree ☒ No ☒ Fp ☒ Partial ☒ Do not know	True Decision UNCHANGED ERROR ☒ 1 ☒ 04 ☒ 02 ☒ 05 ☒ 03	Second Reading of Average Decision ☒ Normal ☒ Normal Borderline ☒ Abnormal Borderline ☒ Abnormal ☒ Remove from population ☒ Unsure Test complete ? No		
591589 F 34 17JUN91 Load 20 kg Max 20 kg	Normal up to 23 kg. No data beyond that load. Average distance to threshold 2.0%	First Reading Normal Valid Average ? Yes Confidence level exceeding normal from abnormal 98%	☒ Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (radio, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness	☒ Honest ☒ Cheater ☒ Unclear ☒ Pretend NORMAL ☒ Pretend ABNORMAL	The detected thoracic lumbar dysfunction was not given enough importance by the expert system.
First reading	Clinical Examination Consensus A(b)-H A(b) - H Consistency of assessment ☒ Yes ☒ Disagree ☒ No ☒ Fp ☒ Partial ☒ Do not know	True Decision UNCHANGED ERROR ☒ 1 ☒ 04 ☒ 02 ☒ 05 ☒ 03	Second Reading of Average Decision ☒ Normal ☒ Normal Borderline ☒ Abnormal Borderline ☒ Abnormal ☒ Remove from population ☒ Unsure Test complete ? Yes		

For the individual conclusions

HH	Certainly Normal	H	Honest
N	Probably Normal	NN	Certainly
U	Uncertain		
Ab(b)	Abnormal borderline		
A	Probably Abnormal		
AA	Certainly Abnormal		

For the physical examination

ND	Dynamic exam Normal
Y?	No clinical data recorded

There is a bias in the dynamic portion of the clinical examination: Dr Ferlin finds more pathological cases than Dr Bourier. This bias is somewhat compensated by history and pain in the clinical conclusion. In general, the expert system agrees with Dr Bourier's decision. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not physiotherapists. The Spinoscope assessing function without pain or history, is therefore bound to agree more with Dr Bourier.

Definitions

1	Correctable error
2	Severe disagreement

ERROR CODES

1	Full agreement.
2	Default since clinicians are "borderline".
3	Default since clinicians are divided.
4	True Spinoscope decision first reading low weights.
5	Default since at least one clinician flipped.

MATCH CODES

1	Full agreement.
2	Default since clinicians are "borderline".
3	Default since clinicians are divided.
4	True Spinoscope decision first reading low weights.
5	Default since at least one clinician flipped.

TRUE and IRSST AVERAGE Splnoscope Decslon - Expert System Version 2.15

with borderline zone of plus-minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 99%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the ERST	Functional Anomalies (Pathology)	Malingering	Comments Main events which lead to the value of the Expert System
597283 M 48 11/1/03 Load 10kg Max 20 kg	Normal up to 23 kg. No data beyond that load.	First Reading Normal / Borderline Valid Average ? Yes Confidence level separating normal from abnormal 1.0 %	Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB tension injury ☒ Excellent flexion ☒ Poor data (redo, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness Second Reading of Average Decision ☒ Normal ☒ Normal Borderline ☒ Abnormal ☒ Remove from population ☒ Unsure MATCH ☒ 1 ☒ 4 ☒ 5 ☒ 2 ☒ 3 True Decision UNCHANGED Clinician's examination Invalid Input selected for ROC calculations spine Consistency of conclusions ☒ Yes ☒ No ☒ Disagree ☒ Flip ☒ Perfid ☒ Do not know	☒ Honest ☒ Cheater ☒ Unsure ☒ Pretend NORMAL ☒ Pretend ABNORMAL	Minimal pathology around 20 kg.
First reading					
Second reading					
600393 M 48 22/1/00 Load 10kg Max 14 kg	Normal up to 14 kg. No data beyond that load.	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal 2.0 %	Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB tension injury ☒ Excellent flexion ☒ Poor data (redo, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness Second Reading of Average Decision ☒ Normal ☒ Normal Borderline ☒ Abnormal ☒ Remove from population ☒ Unsure MATCH ☒ 1 ☒ 4 ☒ 5 ☒ 2 ☒ 3 True Decision UNCHANGED Clinician's examination Invalid Input selected for ROC calculations spine Consistency of conclusions ☒ Yes ☒ No ☒ Disagree ☒ Flip ☒ Perfid ☒ Do not know	☒ Honest ☒ Cheater ☒ Unsure ☒ Pretend NORMAL ☒ Pretend ABNORMAL	The clinicians themselves disagree. A trunk velocity of 110% when lifting 22 kg is a good sign of normality. Could be disqualification since the 8 and 14 kg lifts indicate sprain but the 4kg and 23 kg lifts data are too corrupted to confirm the pathology.
First reading					
Second reading					
606924 F 43 22/1/02 Load 10kg Max 9 kg	Normal up to 9 kg. No data beyond that load.	First Reading Normal / Borderline Valid Average ? Yes Confidence level separating normal from abnormal 1.0 %	Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB tension injury ☒ Excellent flexion ☒ Poor data (redo, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar stiffness Second Reading of Average Decision ☒ Normal ☒ Normal Borderline ☒ Abnormal ☒ Remove from population ☒ Unsure MATCH ☒ 1 ☒ 4 ☒ 5 ☒ 2 ☒ 3 True Decision UNCHANGED Clinician's examination Invalid Input selected for ROC calculations spine Consistency of conclusions ☒ Yes ☒ No ☒ Disagree ☒ Flip ☒ Perfid ☒ Do not know	☒ Honest ☒ Cheater ☒ Unsure ☒ Pretend NORMAL ☒ Pretend ABNORMAL	Overweight - Caution in interpreting skin markers motion.
First reading					
Second reading					

MATCH CODES

- 1 Full agreement.
- 2 Default since clinicians are "indecisive".
- 3 Default since clinicians are divided.
- 4 True Spinoscopic decision first reaching low weights.
- 5 Default since at least one clinician flipped.

ERROR CODES

- 1 Correctable error**
2 Serious disagreement!

Definitions

There is a bias in the dynamic portion of the clinical examination: Dr Farlin finds more pathological cases than Dr Boover. This bias is somewhat compensated by history and pain in the clinical conclusions. In general, the expert system agrees with Dr Boover's decisions. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not physiatrists. The Spinecoo assessing function without pain or history, is therefore

For the physical examination

ND	Dynamics exams Normal
??	No clinical data recorded

For the Individual conclusions

HH	Certainly Normal	H	Honest
NH	Probably Normal	HH	Certainly honest
U	Undecided		

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 80%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Make events which lead the rules of the Expert System
608676 F 27 21F001 Load lifted Max 6 kg	Normal up to 8 kg. No data beyond that load. Average distance to threshold: 0.0%	First Reading: Normal Valid Average? Yes Confidence level separating normal from abnormal: 100%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB Junction Injury ☐ Excellent flexion ☐ Poor data (reds, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Pseudo NORMAL ☐ Pseudo ABNORMAL	Boerler dynamic physical examination and conclusion differ. Spinoscope agrees with Boerler dynamic assessment. There is a drop in performance at 8 kg but not enough to have the subject qualify as abnormal.
Second reading	Clinical Examination Individual: ☐ Doubtful Consensus: A(b)-H A(b)-H	True Decision UNCHANGED MATCH: 1 4 5 2 3 Clinician's examination: Invalid Type selected for ROC calculations: Spine	Second Reading of Average Decision ☐ Normal ☐ Normal Borderline ☐ Abnormal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure Test complete? No		
629233 M 18 17M000 Load lifted Max 32 kg	Normal up to 10 kg, then abnormal up to 32 kg. Average distance to threshold: -1.0%	First Reading: Abnormal / Borderline Mixed Valid Average? Doubtful Confidence level separating normal from abnormal: 0%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB Junction Injury ☐ Excellent flexion ☐ Poor data (reds, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Pseudo NORMAL ☐ Pseudo ABNORMAL	The pathology interferes with function at high loads. The 32 kg lift is laborious. Consider 32 kg to be the limit of frequent lifting.
Second reading	Clinical Examination Individual: ☐ Doubtful Consensus: NN-HH NN-HH	True Decision UNCHANGED MATCH: 1 4 5 2 3 Clinician's examination: Valid Type selected for ROC calculations: Clinician	Second Reading of Average Decision ☐ Normal ☐ Normal Borderline ☐ Abnormal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure Test complete? Yes		
645188 M 31 22L006 Load lifted Max 23 kg	Normal up to 7 kg, then abnormal up to 23 kg. Average distance to threshold: -3.3%	First Reading: Abnormal Valid Average? Yes Confidence level separating normal from abnormal: 100%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB Junction Injury ☐ Excellent flexion ☐ Poor data (reds, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Pseudo NORMAL ☐ Pseudo ABNORMAL	The spine stiffness increases with loads.
Second reading	Clinical Examination Individual: ☐ Doubtful Consensus: A(b)-HH A(b)-HH	True Decision UNCHANGED MATCH: 1 4 5 2 3 Clinician's examination: Valid Type selected for ROC calculations: Clinician	Second Reading of Average Decision ☐ Normal ☐ Normal Borderline ☐ Abnormal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure Test complete? Yes		

For the individual conclusions

NN	Certainly Normal	HH	Honest
NN	Probably Normal	NN	Certainly
UU	Unclear	NN	honest
A(b)	Abnormal borderline		
A	Probably Abnormal		
AA	Certainly Abnormal		

For the physical examination

ND	Dynamic exam Normal
??	No clinical data recorded

There is a bias in the dynamic portion of the clinical examination: Dr Fortin finds more pathological cases than Dr Boerler. This bias is somewhat compensated by history and pain in the clinical conclusions. In general, the expert system agrees with Dr Boerler's decision. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not physicians. The Spinoscope assessing function without pain or history, is therefore bound to agree more with Dr Boerler.

Definitions

1	Correctable error
2	Serious disagreement

ERROR CODES

1	Full agreement.
2	Default since clinicians are "Indecisive".
3	Default since clinicians are divided.
4	True Spinoscope decision first reading low weights.
5	Default since at least one clinician slipped.

MATCH CODES

1	Full agreement.
2	Default since clinicians are "Indecisive".
3	Default since clinicians are divided.
4	True Spinoscope decision first reading low weights.
5	Default since at least one clinician slipped.

Tuesday, December 13, 1994 11:05

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 95%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the archive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which fixed the rules of the Expert System
649875	Normal up to 23 kg. No data beyond that load.	First Reading Normal / Borderline Valid Average ? Yes Confidence level separating normal from abnormal 63%	☒ Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (radio, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar affixes	☒ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Minor pathology.
First reading					
Second reading					
660839	Abnormal	First Reading Mixed Valid Average ? Yes Confidence level separating normal from abnormal 100%	☒ Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (radio, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar affixes	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Very consistent just below the threshold of normality. The pathology is minor but considerably there.
First reading					
Second reading					
690108	Abnormal	First Reading Mixed Valid Average ? Yes Confidence level separating normal from abnormal 100%	☒ Thoracic Lumbar Junction ☒ Lower lumbar restriction ☒ LB torsion injury ☒ Excellent flexion ☒ Poor data (radio, overweight) ☒ Potential conflict clinicians ☒ Sprain ☒ All lumbar affixes	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Sprain patient.
First reading					
Second reading					

For the individual conclusions		For the physical examination		Definitions		ERROR CODES		MATCH CODES	
NH	Certainly Normal	ND	Dynamic exam Normal	1	Correctable error	1	Full agreement	1	Full agreement
N	Probably Normal	T?	No clinical data recorded	2	Serious disagreement	2	Defectable error	2	Defectable error
U	Unclassified								
A(b)	Abnormal borderline								
A	Probably Abnormal								
AA	Certainly Abnormal								

There is a bias in the dynamic portion of the clinical examination: Dr Fertin finds more pathological cases than Dr Boerier. This bias is somewhat compensated by history and pain in the clinical conclusions. In general, the expert system agrees with Dr Boerier's decision. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not physiotherapists. The Spinoscope assesses function without pain or history. In therefore

Tuesday, December 13, 1994 11:05

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 90%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which lead the value of the Expert System
702418 F 33 18/03/94 Load 10kg Max 5 kg	Normal up to 5 kg. No data beyond that load. Average distance to threshold 2.0%	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal 99%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ L5/S1 Junction ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar effusions	☐ Honest ☐ Cheater ☐ Unsure ☐ Not enough data Test complete ?	No reasons to stop lifting. Consider stimulation.
3 Second reading	Clinical Examination Consensus AA - HH Average distance to threshold 4.0%	True Decision UNCHANGED ERROR 1 2 3 Clinician's examination Valid Type selected for ROC calculations Clinician	☐ Normal ☐ Normal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure	☐ Honest ☐ Cheater ☐ Unsure ☐ Not enough data Test complete ?	Data insufficient. Minor pathology on the right which does not interfere significantly with function for loads below 5 kg. <i>sample data</i>
704921 M 43 20/01/91 Load 10kg Max 0 kg	Remove from population Average distance to threshold 4.0%	First Reading Normal Valid Average ? Doubtful Confidence level separating normal from abnormal 100%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ L5/S1 Junction ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar effusions	☐ Honest ☐ Cheater ☐ Unsure ☐ Not enough data Test complete ?	No reason for refusing to lift anything. Strongly suspect stimulation.
5 Second reading	Clinical Examination Consensus A - U Average distance to threshold 0.0%	True Decision UNCHANGED ERROR 1 2 3 Clinician's examination Valid Type selected for ROC calculations Clinician	☐ Normal ☐ Normal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure	☐ Honest ☐ Cheater ☐ Unsure ☐ Not enough data Test complete ?	Disagreement between clinicians on the status of the cheater. Not enough data to decide beyond that the subject unloaded is has a normal function. The pathology, if any, will be detected at higher loads. The expert system cannot decide since the Spinoscope testing protocol was not followed.
722399 M 31 27/04/92 Load 10kg Max 21 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 0.0%	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ L5/S1 Junction ☐ Excellent flexion ☐ Poor data (redo, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar effusions	☐ Honest ☐ Cheater ☐ Unsure ☐ Not enough data Test complete ?	
5 Second reading	Clinical Examination Consensus NN - HH Average distance to threshold 0.0%	True Decision UNCHANGED MATCH 1 2 3 Clinician's examination Valid Type selected for ROC calculations Clinician	☐ Normal ☐ Normal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure	☐ Honest ☐ Cheater ☐ Unsure ☐ Not enough data Test complete ?	

For the individual conclusions

☐ Certainly Normal
☐ Probably Normal
☐ Undecided
☐ Abnormal Borderline
☐ Probably Abnormal
☐ Certainly Abnormal

For the physical examination

☐ Dynamic exam Normal
☐ No clinical data recorded

Definitions

There is a bias in the dynamic portion of the elicited examination: Dr Fortin finds more pathological cases than Dr Boerter. This bias is somewhat compensated by history and pain in the clinical conclusions. In general, the expert system agrees with Dr Boerter's decision. This may be due to the fact that the expert system has been built by orthopedic surgeons and not physicians. The Spinoscope assessing function without pain or history, is therefore bound to agree more with Dr Boerter.

ERROR CODES

1 Correctable error
 2 Serious disagreement

MATCH CODES

1 Full agreement
 2 Default since clinicians are "Indecisive"
 3 Default since clinicians are divided
 4 True Spinoscope decision first reading low weights
 5 Default since at least one clinician flipped

Tuesday, December 13, 1994 11:05

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of probability 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 90%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which feed the rules of the Expert System
758325 M 33 31/10/92 Lead 100g Max 23 kg	Normal up to 23 kg. No data beyond that load. Average distance to threshold 7.9%	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Presumed NORMAL ☐ Presumed ABNORMAL	
First reading	Clinical Examination Individual: Doctor, Exam, Decision AA A(b) - H TT A(b) - U Consensus: A(b) - U	True Decision UNCHANGED ERROR ☐ 1 ☐ 4 ☐ 2 ☐ 5 ☐ 3 ☐ 0	Second Reading of Average Decision ☐ Normal ☐ Normal Borderline ☐ Abnormal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure	☐ Honest ☐ Cheater ☐ Unclear ☐ Presumed NORMAL ☐ Presumed ABNORMAL	Disagreement on the cheater - Bourcier provides no data except "viable" 77 in support of his conclusion. It seems that he saw very little. In the event that the physical examination of Bourcier was "normal", then the Spinoscope agrees and the conclusion with the clinicians should be enriched from ERROR to MATCH. A trust velocity of 84 degrees per second at 23 kg and a ROM of 70 degrees cannot be significant.
Second reading	Clinical Examination Individual: Doctor, Exam, Decision AA A(b) - H TT A(b) - U Consensus: A(b) - U	True Decision UNCHANGED ERROR ☐ 1 ☐ 4 ☐ 2 ☐ 5 ☐ 3 ☐ 0	Second Reading of Average Decision ☐ Normal ☐ Normal Borderline ☐ Abnormal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure	☐ Honest ☐ Cheater ☐ Unclear ☐ Presumed NORMAL ☐ Presumed ABNORMAL	
769239 F 33 19/02/93 Lead 100g Max 14 kg	Normal up to 6 kg, then abnormal up to 14 kg. Average distance to threshold -3.9%	First Reading Abnormal / Borderline Mixed Valid Average ? Yes Confidence level separating normal from abnormal 21%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Presumed NORMAL ☐ Presumed ABNORMAL	Torsional injury on the right.
First reading	Clinical Examination Individual: Doctor, Exam, Decision AA A(b) - H TT A(b) - U Consensus: A(b) - U	True Decision UNCHANGED MATCH ☐ 1 ☐ 4 ☐ 2 ☐ 5 ☐ 3	Second Reading of Average Decision ☐ Normal ☐ Normal Borderline ☐ Abnormal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure	☐ Honest ☐ Cheater ☐ Unclear ☐ Presumed NORMAL ☐ Presumed ABNORMAL	
Second reading	Clinical Examination Individual: Doctor, Exam, Decision AA A(b) - H TT A(b) - U Consensus: A(b) - U	True Decision UNCHANGED MATCH ☐ 1 ☐ 4 ☐ 2 ☐ 5 ☐ 3	Second Reading of Average Decision ☐ Normal ☐ Normal Borderline ☐ Abnormal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure	☐ Honest ☐ Cheater ☐ Unclear ☐ Presumed NORMAL ☐ Presumed ABNORMAL	
778326 M 25 14/02/92 Lead 100g Max 81 kg	Normal up to 22 kg, then abnormal up to 31 kg. Average distance to threshold 3.9%	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal 79%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB torsion injury ☐ Excellent flexion ☐ Poor data (radio, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unclear ☐ Presumed NORMAL ☐ Presumed ABNORMAL	Lower lumbar restriction increasing with loads. The pathology interferes with function quite late and is likely to be released by the clinician.
First reading	Clinical Examination Individual: Doctor, Exam, Decision AA A(b) - H TT A(b) - U Consensus: A(b) - U	True Decision UNCHANGED MATCH ☐ 1 ☐ 4 ☐ 2 ☐ 5 ☐ 3	Second Reading of Average Decision ☐ Normal ☐ Normal Borderline ☐ Abnormal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure	☐ Honest ☐ Cheater ☐ Unclear ☐ Presumed NORMAL ☐ Presumed ABNORMAL	
Second reading	Clinical Examination Individual: Doctor, Exam, Decision AA A(b) - H TT A(b) - U Consensus: A(b) - U	True Decision UNCHANGED MATCH ☐ 1 ☐ 4 ☐ 2 ☐ 5 ☐ 3	Second Reading of Average Decision ☐ Normal ☐ Normal Borderline ☐ Abnormal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure	☐ Honest ☐ Cheater ☐ Unclear ☐ Presumed NORMAL ☐ Presumed ABNORMAL	

For the individual conclusions

NN Certainly Normal
 N Probably Normal
 U Unclassified
 A(b) Abnormal borderline
 A Probably Abnormal
 AA Certainly Abnormal

For the physical examination

ND Dynamic exam Normal
 TT No clinical data recorded

Definitions:

ERROR CODES

1 Correctable error
 2 Serious disagreement

MATCH CODES

- 1 Full agreement.
 2 Default since clinicians are "Indecisive".
 3 Default since clinicians are divided.
 4 True Spinoscope decision first reading low weight.
 5 Default since at least one clinician flipped.

There is a bias in the dynamic portion of the clinical examination: Dr Perin finds more pathological cases than Dr Bourcier. This bias is accentuated by history and pain in the clinical conclusions. In general, the expert system agrees with Dr Bourcier's decision. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not physiatrists. The Spinoscope assessing function without pain or history. In therefore

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 90%	Average Decision This mathematical average is not just of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which lead the value of the Expert System
800369 F 47 ZHONG Lead 100 kg Max 20 kg	Normal up to 23 kg. No data beyond that load. Average distance to threshold 7.8 %	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB lumbal injury ☐ Exquisite flexion ☐ Poor data (flex, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unsure ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Cautious for his age. ✓
First reading					
Second reading					
803808 F 43 ZALTON Lead 100 kg Max 14 kg	Normal unloaded and abnormal thereafter Average distance to threshold -9.8 %	First Reading Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB lumbal injury ☐ Exquisite flexion ☐ Poor data (flex, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unsure ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Overweight - Caution in interpreting skin markers motion, especially the lateral bend data on left. ✓
First reading					
Second reading					
812166 M 33 SALMON Lead 100 kg Max 31 kg	Normal up to 31 kg. No data beyond that load. Average distance to threshold 9.9 %	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal 100 %	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB lumbal injury ☐ Exquisite flexion ☐ Poor data (flex, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar stiffness	☐ Honest ☐ Cheater ☐ Unsure ☐ Pretend NORMAL ☐ Pretend ABNORMAL	A trunk velocity of 100% when lifting 31 kg is a good sign of normality and full collaboration. Hence, even if the subject was paid to cheat, he behaved like a honest man. The Spinoscope measures what was done and not what was intended. ✓
First reading					
Second reading					

MATCH CODES	
1	Full agreement.
2	Default since clinicians are "borderline".
3	Default since clinicians are divided.
4	True Spinoscope decision first reading low weights.
5	Default since at least one clinician slipped.

ERROR CODES	
1	Correctable error
2	Serious disagreement

Definitions	
ND	Dynamic exam Normal
T?	No clinical data recorded

There is a bias in the dynamic portion of the clinical examination: Dr Ferlie finds more pathological cases than Dr Bowler. This bias is somewhat compensated by history and pain in the clinical conclusions. In general, the expert system agrees with Dr Bowler's decision. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not physicians. The Spinoscope assessing function without pain or history, is therefore based to agree more with Dr Bowler.

For the individual conclusions	
HN	Certainly Normal
N	Probably Normal
U	Uncertain
AB	Abnormal borderline
A	Probably Abnormal
AA	Certainly Abnormal

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 99%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments
820360 F 47 1640000 Lead 1000 Max 31 kg	Normal up to 23 kg, then abnormal up to 31 kg. Average distance to threshold 1.9%	First Reading: Normal / Borderline Mixed Valid Average? Yes Confidence level separating normal from abnormal 35%	Thoracic Lumbar Junction Lower lumbar restriction LB lumbar injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	Honest Cheater Unclear Pretest NORMAL Pretest ABNORMAL	Failed at 31 kg. Slight stiffness, possibly normal for age.
First reading	Clinical Examination Consensus: A(b) - H A(b) - H	True Decision: UNCHANGED MATCH: 1 4 2 5 3 Clinician's examination: Invalid Typed selected for ROC calculations: Spine	Second Reading of Average Decision Normal Borderline Abnormal Borderline Abnormal Remove from population Unclear		Border dynamic physical examination and conclusion differ. Spinoscope agrees with Bourlier's initial assessment. A trunk velocity of 85% when lifting 23 kg is a reasonable sign that any pathology cannot be significant.
Second reading					
820726 M 87 1660001 Lead 1000 Max 14 kg	Abnormal Average distance to threshold -27.6%	First Reading: Abnormal Valid Average? Yes Confidence level separating normal from abnormal 100%	Thoracic Lumbar Junction Lower lumbar restriction LB lumbar injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	Honest Cheater Unclear Pretest NORMAL Pretest ABNORMAL	Failed to lift 23 kg. Stiffness detected at 4/5.
First reading	Clinical Examination Consensus: A - C A - C	True Decision: UNCHANGED MATCH: 1 4 2 5 3 Clinician's examination: Valid Typed selected for ROC calculations: Clinician	Second Reading of Average Decision Normal Borderline Abnormal Borderline Abnormal Remove from population Unclear		Clinician disagrees on the status of the cheater.
Second reading					
823200 F 23 1640000 Lead 1000 Max 0 kg	Remove from population Average distance to threshold -14.6%	First Reading: Abnormal Valid Average? Doubtful Confidence level separating normal from abnormal 100%	Thoracic Lumbar Junction Lower lumbar restriction LB lumbar injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	Honest Cheater Unclear Pretest NORMAL Pretest ABNORMAL	Suit pants detected during the extended recovery from flexion. The refusal to comply with the test protocol did not permit to assess the normally portion. Hence the relatively high ratio of pathology which must be rejected as possible artifact. This subject must be removed from the population.
First reading	Clinical Examination Consensus: AA - HH AA - HH	True Decision: UNCHANGED MATCH: 1 4 2 5 3 Clinician's examination: Valid Typed selected for ROC calculations: Clinician	Second Reading of Average Decision Normal Borderline Abnormal Borderline Abnormal Remove from population Unclear		
Second reading					

For the individual conclusions

NN	Certainly Normal	H	Honest
N	Probably Normal	HH	Certainly Honest
U	Unclear		
A(b)	Abnormal borderline		
A	Probably Abnormal		
AA	Certainly Abnormal		

For the physical examination

ND	Dynamic exam Normal
77	No clinical data recorded

There is a bias in the dynamic portion of the clinical examination: Dr Forth finds more pathological cases than Dr Bourlier. This bias is somewhat compensated by history and pain in the clinical conclusions. In general, the expert system agrees with Dr Bourlier's decision. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not osteopaths. The Spinoscope examines function without pain or stiffness is therefore

Definitions

1	Correctable error
2	Serious disagreement

ERROR CODES

1	Full agreement.
2	Default since clinicians are "Indecisive".
3	Default since clinicians are divided.
4	True Spinoscope decision first reading low weights.
5	Default since at least one clinician flipped.

MATCH CODES

1	Full agreement.
2	Default since clinicians are "Indecisive".
3	Default since clinicians are divided.
4	True Spinoscope decision first reading low weights.
5	Default since at least one clinician flipped.

Tuesday, December 13, 1994 11:06

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of physicians 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 80%	Average Decision This methodological average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments
823736 F 20 OARPOZ Lead third Max 18 kg	Normal unloaded and abnormal thereafter	First Reading: Mixed Valid Average: 7 Confidence level separating normal from abnormal: 100%	Thoracic Lumbar Junction Lower lumbar restriction LB tension injury Excellent flexion Poor data (flex, overweight) Potential conflict clinicians Sprain All lumbar effusions	☒ Honest ☐ Cheater ☐ Unsure ☐ Pretest NORMAL ☐ Pretest ABNORMAL	Consistent with disinclination except for speed which is good.
First reading					
Second reading					
828686 M 23 GSHOPO Lead third Max 31 kg	Normal up to 15 kg, then abnormal up to 31 kg.	First Reading: Mixed Valid Average: 7 Confidence level separating normal from abnormal: 74%	Thoracic Lumbar Junction Lower lumbar restriction LB tension injury Excellent flexion Poor data (flex, overweight) Potential conflict clinicians Sprain All lumbar effusions	☐ Honest ☐ Cheater ☐ Unsure ☐ Pretest NORMAL ☐ Pretest ABNORMAL	Slow down at 31 kg. The EMG response is atypical. Possible attempt to load the machine. Sprain detected in high back.
First reading					
Second reading					
829292 F 20 OFFERD Lead third Max 31 kg	Normal up to 31 kg. No data beyond that load.	First Reading: Yes Valid Average: 7 Confidence level separating normal from abnormal: 100%	Thoracic Lumbar Junction Lower lumbar restriction LB tension injury Excellent flexion Poor data (flex, overweight) Potential conflict clinicians Sprain All lumbar effusions	☐ Honest ☐ Cheater ☐ Unsure ☐ Pretest NORMAL ☐ Pretest ABNORMAL	Very supple.
First reading					
Second reading					

For the individual conclusions

NN	Certainly Normal	H	Honest
N	Probably Normal	NN	Certainly honest
U	Unclassified		
A(s)	Abnormal borderline		
A	Probably Abnormal		
AA	Certainly Abnormal		

For the physical examination

ND	Dynamic exam Normal
??	No clinical data recorded

Definitions

There is a bias in the dynamic portion of the clinical examination: Dr Ferlin finds more pathological cases than Dr Bourier. This bias is somewhat compensated by history and pain in the clinical examination. In general, the expert system agrees with Dr Bourier's decision. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not physiatrists. The Spinoscope recording function without pain or history, is therefore based to agree more with Dr Bourier.

ERROR CODES

1	Correctable error
2	Serious disagreement

MATCH CODES

1	Full agreement
2	Default since clinicians are "indecisive".
3	Default since clinicians are divided.
4	True Spinoscope decision first reading low weight.
5	Default since at least one clinician slipped.

Tuesday, December 13, 1994 11:05

Page 27 of 30

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 90%	Average Decision This automated average is not part of the regular use of Spinoscope and has been developed for the machine use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which lead the value of the Expert System
884132 M 50 08AP002 Lead 10kg Max 12 kg	Normal unloaded and abnormal thereafter	First Reading Abnormal Valid Average ? Yes Confidence level separating normal from abnormal 100%	Thoracic Lumbar Junction Lower lumbar restriction LB tension injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	Honest Cheater Unclear Pretend NORMAL Pretend ABNORMAL	Reading and shaking with leads. Sprain.
First reading	Clinical Examination Individual Decision Consensus N - H	Consistency of conclusions Yes ☐ No ☐ Diagono ☐ Flip ☐ Partial ☐ Do not know ☐	True Decision UNCHANGED ERROR 1 ☐ 4 ☐ 5 ☐ 2 ☐ 3 ☐	Second Reading of Average Decision Honest ☐ Cheater ☐ Unclear ☐ Not enough data ☐ Test complete ? Yes	
Second reading	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%		Thoracic Lumbar Junction Lower lumbar restriction LB tension injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL ☐	
885854 M 35 S4E002 Lead 10kg Max 26 kg	Normal up to 35 kg. No data beyond that load.	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%	Thoracic Lumbar Junction Lower lumbar restriction LB tension injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL ☐	
First reading	Clinical Examination Individual Decision Consensus NN - HH	Consistency of conclusions Yes ☐ No ☐ Diagono ☐ Flip ☐ Partial ☐ Do not know ☐	True Decision UNCHANGED MATCH 1 ☐ 4 ☐ 5 ☐ 2 ☐ 3 ☐	Second Reading of Average Decision Honest ☐ Cheater ☐ Unclear ☐ Not enough data ☐ Test complete ? Yes	
Second reading	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%		Thoracic Lumbar Junction Lower lumbar restriction LB tension injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL ☐	
895423 M 25 S1J001 Lead 10kg Max 21 kg	Normal up to 31 kg. No data beyond that load.	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%	Thoracic Lumbar Junction Lower lumbar restriction LB tension injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL ☐	
First reading	Clinical Examination Individual Decision Consensus NN - HH	Consistency of conclusions Yes ☐ No ☐ Diagono ☐ Flip ☐ Partial ☐ Do not know ☐	True Decision UNCHANGED MATCH 1 ☐ 4 ☐ 5 ☐ 2 ☐ 3 ☐	Second Reading of Average Decision Honest ☐ Cheater ☐ Unclear ☐ Not enough data ☐ Test complete ? Yes	
Second reading	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal 100%		Thoracic Lumbar Junction Lower lumbar restriction LB tension injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar stiffness	Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL ☐	

Match Codes

1 Full agreement.
2 Default since clinicians are "borders".
3 Default since clinicians are divided.
4 True Spinoscope decision first reading low weights.
5 Default since at least one clinician flipped.

Error Codes

1 Correctable error
2 Serious disagreement

Definitions

For the physical examination
ND Dynamic exam Normal
?? No clinical data recorded

There is a bias in the dynamic portion of the clinical examination: Dr Fortin finds more pathological cases than Dr Beaulieu. This bias is accounted compensated by history and pain in the clinical conclusions. In general, the expert system agrees with Dr Beaulieu's decision. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not physicians. The Spinoscope assessing function without pain or history. In this case, the Spinoscope is more accurate than the physician.

For the individual conclusions

HH Certainly Normal
H Probably Normal
N Uncertain
A Abnormal Borderline
AA Probably Abnormal
AAA Certainly Abnormal

Dr Fortin

TRUE and IRST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plasmotus 2 % around the threshold

ID #	Summary of True Sphynoscope Decision Threshold set at 90%	Average Decision This mathematical average is not part of the regular use of Sphynoscope and has been developed for the exclusive use of the RRSST	Functional Anomalies (Pathology)	Malingering	Comments
902755	Remove from population	First Reading Normal Valid Average ? Doubtful Confidence level separating normal from abnormal 3.0%	Thoracic Lumbar Junction Lower lumbar restriction LB tension injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar effuses	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Did not lift 12 kg. Sprain at 4.8 kg but not unloaded. Not enough data to confirm.
First reading					
910123	Normal up to 17 kg. No data beyond that load.	First Reading Normal Valid Average ? Yes Confidence level separating normal from abnormal 4.9%	Thoracic Lumbar Junction Lower lumbar restriction LB tension injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar effuses	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Anomaly in axial rotation. Probably old injury well recovered. Minor impact on function. Very supple spine.
First reading					
911174	Remove from population	First Reading Abnormal Valid Average ? Doubtful Confidence level separating normal from abnormal 10.0%	Thoracic Lumbar Junction Lower lumbar restriction LB tension injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar effuses	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Very likely a fraud. Not enough data to analyze the source of the biomechanics. This subject must be removed from the population since the deviation from the testing protocol are too severe to recover the missing data.
First reading					
910202	Remove from population	First Reading Abnormal Valid Average ? Doubtful Confidence level separating normal from abnormal 10.0%	Thoracic Lumbar Junction Lower lumbar restriction LB tension injury Excellent flexion Poor data (redo, overweight) Potential conflict clinicians Sprain All lumbar effuses	☐ Honest ☐ Cheater ☐ Unclear ☐ Pretend NORMAL ☐ Pretend ABNORMAL	Very likely a fraud. Not enough data to analyze the source of the biomechanics. This subject must be removed from the population since the deviation from the testing protocol are too severe to recover the missing data.
Second reading					

MATCH CODES	
1	Full agreement.
2	Default since clinicians are "Indecisive".
3	Default since clinicians are divided.
4	True Spinocope decision first reading for weights.
5	Default since at least one clinician slipped.

ERROR CODES	
1	Correctable error
2	Serious disagreement

Definitions

For the physical examination	
140	Dynamic exam Normal
77	No clinical data recorded

MM	Certainty	Normal	H	Honest
M	Probably	Normal	HH	Certainly honest
V	Undecided			

A97

First reading November 23, 1994 · Second reading December 9, 1994

826 4

Tuesday, December 13, 1994 11:05

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus-minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 99%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments
923925 N 39 1044000 Lead 1044 Max 0 kg	Remove from population Average distance to threshold 9.9%	First Reading: Normal Valid Average ? Doubtful Confidence level regarding normal from abnormal 100%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB lumbar injury ☐ Excellent lesion ☐ Poor data (reds, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar affixes	☐ Honest ☐ Cheater ☐ Unsure ☐ Pretest NORMAL ☐ Pretest ABNORMAL	Normal unloaded. No reason to refuse lifting more. The protocol was not followed. Remove from population. The clinicians themselves severely disagree. The testing protocol was not followed and the Spinoscope assessment is not confirmed and hence unreliable.
First reading	True Decision UNCHANGED Clinicians' examination: Invalid Input selected for ROC calculations: Spine	Second Reading of Average Decision ☐ Normal ☐ Normal Borderline ☐ Abnormal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure Test complete ? No	☐ Honest ☐ Cheater ☐ Unsure ☐ Pretest NORMAL ☐ Pretest ABNORMAL		
965191 U 47 064000 Lead 1044 Max 23 kg	Normal and abnormal thereafter Average distance to threshold -2.9%	First Reading: Abnormal Valid Average ? Yes Confidence level regarding normal from abnormal 87%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB lumbar injury ☐ Excellent lesion ☐ Poor data (reds, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar affixes	☐ Honest ☐ Cheater ☐ Unsure ☐ Pretest NORMAL ☐ Pretest ABNORMAL	Individual functional parameters are OK. The subject score was downgraded because of an unusual global response suggesting a functional injury at L5 on the right implicating also the left side.
First reading	True Decision UNCHANGED Clinicians' examination: Valid Input selected for ROC calculations: Clinician	Second Reading of Average Decision ☐ Normal ☐ Normal Borderline ☐ Abnormal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure Test complete ? Yes	☐ Honest ☐ Cheater ☐ Unsure ☐ Pretest NORMAL ☐ Pretest ABNORMAL		
966190 N 43 212001 Lead 1044 Max 23 kg	Normal unloaded and abnormal thereafter Average distance to threshold -3.5%	First Reading: Abnormal Valid Average ? Yes Confidence level regarding normal from abnormal 100%	☐ Thoracic Lumbar Junction ☐ Lower lumbar restriction ☐ LB lumbar injury ☐ Excellent lesion ☐ Poor data (reds, overweight) ☐ Potential conflict clinicians ☐ Sprain ☐ All lumbar affixes	☐ Honest ☐ Cheater ☐ Unsure ☐ Pretest NORMAL ☐ Pretest ABNORMAL	Although the patient lifted 23 kg, he did it using excessive muscular action. Functional injury at L5 mainly on the left.
First reading	True Decision UNCHANGED Clinicians' examination: Valid Input selected for ROC calculations: Clinician	Second Reading of Average Decision ☐ Normal ☐ Normal Borderline ☐ Abnormal Borderline ☐ Abnormal ☐ Remove from population ☐ Unsure Test complete ? Yes	☐ Honest ☐ Cheater ☐ Unsure ☐ Pretest NORMAL ☐ Pretest ABNORMAL		

For the individual conclusions

IN Certainly Normal
 N Probably Normal
 U Undecided
 Ab) Abnormal borderline
 A Probably Abnormal

For the physical examination

NO Dynamic exam Normal
 ? No clinical data recorded

There is a bias in the dynamic portion of the clinical examination: Dr Fortin finds more pathological cases than Dr Bourcier. This bias is somewhat compensated by history and pain in the clinical conclusions. In general, the expert system agrees with Dr Bourcier's decision. This may be due to the fact that the expert system has been trained for

Definitions

1 Correctable error
 2 Serious disagreement

ERROR CODES

1 Full agreement
 2 Default ethnic clinicians are "Indecisive"
 3 Default ethnic clinicians are "Indecisive"
 4 True Spinoscope decision first reading low weights
 5 Default ethnic at least one clinician flipped

MATCH CODES

1 Full agreement
 2 Default ethnic clinicians are "Indecisive"
 3 Default ethnic clinicians are "Indecisive"
 4 True Spinoscope decision first reading low weights
 5 Default ethnic at least one clinician flipped

TRUE and IRSST AVERAGE Spinoscope Decision - Expert System Version 2.15

with borderline zone of plus/minus 2 % around the threshold

ID #	Summary of True Spinoscope Decision Threshold set at 90%	Average Decision This mathematical average is not part of the regular use of Spinoscope and has been developed for the exclusive use of the IRSST	Functional Anomalies (Pathology)	Malingering	Comments Main events which lead the rules of the Expert System
------	---	--	----------------------------------	-------------	---

Summary of Blind IRSST Study

Average Decision	Validity of Average Decision														
<table> <tr> <td>Normal</td><td>42</td></tr> <tr> <td>Normal Borderline</td><td>0</td></tr> <tr> <td>Abnormal</td><td>26</td></tr> <tr> <td>Abnormal Borderline</td><td>11</td></tr> </table>	Normal	42	Normal Borderline	0	Abnormal	26	Abnormal Borderline	11	<table> <tr> <td>Number of valid decisions</td><td>76</td></tr> <tr> <td>Number of doubtful decisions</td><td>12</td></tr> </table>	Number of valid decisions	76	Number of doubtful decisions	12		
Normal	42														
Normal Borderline	0														
Abnormal	26														
Abnormal Borderline	11														
Number of valid decisions	76														
Number of doubtful decisions	12														
<table> <tr> <td>OUT</td><td>0</td></tr> <tr> <td>Mixed</td><td>19</td></tr> </table>	OUT	0	Mixed	19	<table> <tr> <td colspan="2">Detection of fraud</td></tr> <tr> <td>Clear Honest</td><td>62</td></tr> <tr> <td>Other Cheaters</td><td>6</td></tr> <tr> <td>Pretend Normal</td><td>7</td></tr> <tr> <td>Pretend Abnormal</td><td>10</td></tr> </table>	Detection of fraud		Clear Honest	62	Other Cheaters	6	Pretend Normal	7	Pretend Abnormal	10
OUT	0														
Mixed	19														
Detection of fraud															
Clear Honest	62														
Other Cheaters	6														
Pretend Normal	7														
Pretend Abnormal	10														
Miscellaneous															

Definitions

Summary of TRUE Spinoscope Decision: Very few injured individuals are unable to move in a near normal manner when they are not under stress (i.e. loaded). In that case, the Machine specifies that an injured individual can lift "X" kg in a near normal manner. For loads beyond "X" kg, the injury interferes with an otherwise near normal function and the pattern of coordination shifts significantly towards an abnormal status. The decision threshold is taken to be 90%.

A clinician may express the same idea by deciding that a subject is suitable for "light duties". What is meant is that the subject is presumably capable of lifting light loads in a near normal manner, and is presumably incapable of lifting heavier loads safely. The machine description is far more precise in that regard. The TRUE Decision which follows a rigid interpretation protocol is the only method recommended in operating the Spinoscope.

The AVERAGE Distance to Threshold. Measures the distance between the Average probability of normality across all loads and the arbitrary threshold of 90%. In other words, the probability of normality is the Average distance augmented by 90%.

The AVERAGE Decision: Is the average probability of normality for all completed movements. Such an average response (Normal/Abnormal) is a mathematical curiosity which may have no clinical meaning in some situations. This would be the case of subject 130717 who has a near normal response from zero to 12 kg lift and then an abnormal response from 13 to 23 kg lift. The average decision would then be neither normal nor abnormal but borderline, although the TRUE Decision has clearly defined two mutually exclusive zones for normality and abnormality. Hence, this average decision does not characterize what the subject can do and SpineX does not recommend the use of such averages in those particular situations.

Below the AVERAGE Decision, a qualifier (being either MIXED or OUT) is given. MIXED means that there exists a normal and abnormal range of movement. OUT means that the data is not suitable for machine processing.

The Confidence level. Characterizes the confidence in the calculated mathematical Average Decision when the average of the probability of normality is too close to the Expert system's decision threshold of 90%. At the limit, when the Average distance is zero (probability of normality is 90%), it is impossible to determine if the subject is normal or abnormal. The Confidence level is then zero. A Confidence level of one expresses that the Average is sufficiently far from the decision threshold to be acceptable (outside the + 2% zone around the 90% threshold).

Functional Anomalies. Intended to express the nature of pathology from a functional and not an anatomical point of view. The confidence level expresses how strong the pathology appears to the expert system. Even normal individuals may have traces of functional pathology which are deemed sufficiently minor to be unimportant for the range of loads tested. Hence, being normal does not exclude the presence of pathology. It is only when the pathology crosses a certain threshold that the subject is classified abnormal.

The Index "Relative Importance of pathology" expresses that the pathological score is sufficiently large to force the probability of normality to be lowered below the 90% threshold. It must be noted that being classified as "normal" does not imply the absence of pathology. Being normal simply means that the pathology, if present, is contained. This is the case of subject 130717 in whom the pathology is contained up to 12 kg and then over the normal coordination for heavier loads lifted.

Malingering. The subjects are classified into three groups: honest, cheater, and unclear. A subject deemed cheater or classified as unclear is furthermore assigned a label "pretend normal", if the subject is in fact abnormal but tries to dissimulate his condition, and "Pretend abnormal" if the subject is in fact normal and tries to simulate.

Comments. To assist the clinician in understanding the process, a short description extracted from the expert system reasoning process is given.

For the individual conclusions			
HN	Certainty Normal	H	Honest
N	Probably Normal	HN	Certainty Normal
B	Uncertain		
A(b)			
A	Abnormal borderline		
AA	Probably Abnormal		
AAA	Certainty Abnormal		

For the physical examination
ND Dynamic exam Normal
TT No clinical data recorded

Definitions

ERROR CODES	
1	Correctable error
2	Serious misassignment

MATCH CODES	
1	Full agreement.
2	Default since clinicians are "indecisive".
3	Default since clinicians are divided.
4	True Spinoscopic decision first reading low weights.
5	Default since at least one clinician slipped.

There is a bias in the dynamic portion of the clinical examination: Dr Ferlin finds more pathological cases than Dr Bourlier. This bias is somewhat compensated by history and pain in the clinical conclusions. In general, the expert system agrees with Dr Bourlier's decision. This may be due to the fact that the expert system has been trained by orthopedic surgeons and not physicians. The Spinoscope assessing function without pain or history, is therefore bound to agree more with Dr Bourlier.

ANNEXE II

Références additionnelles

Références additionnelles**B-200**

1. Aghazadeh F. Simulated dynamic lifting strength models for manual lifting. [Ph.D. dissertation]. Lubbock (TX): Texas Tech University, 1982.
2. Aghazadeh F, Ayoub MM. A comparison of dynamic and static strength models for prediction of lifting capacity. *Ergonomics* 1985;28(10):1409-17.
3. Andersson GBJ. Epidemiologic aspects on low back pain in industry. *Spine* 1981;6:53-60.
4. Ayoub MM. Psychophysical techniques. Paper presented at AIHA Conference, June 1982.
5. Ayoub MM, Bethea NJ, Deivanayagam S, Asfour SS, Bakken GM, Liles D, Mital A, et al. Determination and modeling of lifting capacity. Final Report. HEW (NIOSH) Grant No. 5R01OH-00545-02, 1978.
6. Biering-Sorensen F. A one-year prospective study of low back trouble in a general population: the prognostic value of low back history and physical measurement. *Dan Med Bull* 1984;31:362-75.
7. Chaffin DB. Human strength capability and low-back pain. *J Occup Med* 1974;16:248-54.
8. Chaffin DB, Herrin GD, Keyserling WM. Preemployment strength testing: an updated position. *J Occup Med* 1978;20:403-8.
9. Chaffin DB, Park KS. A longitudinal study of low back pain as associated with occupational weight lifting factors. *Am Ind Hyg Assoc J* 1973;34:513-25.
10. Deutsch SD, Litchman HM. A comprehensive system to evaluate back function. Presented at the Isotechnologies Users Application Seminar. Dallas (TX), June 1988.
11. Dryden RD. A predictive model for the maximum permissible weight of lift from knuckle to shoulder height [doctoral dissertation]. Lubbock (TX): Texas Tech University, 1973.
12. Dumas GA, Poulin MJ, Roy B, Gagnon M, Javanonic M. Quantitative anatomy of trunk muscles. *Proceedings of North American Congress on Biomechanics* 1986;81-2.
13. Foreman TK, Baxter CE, Troup JDG. Ratings of acceptable load and maximal isometric lifting strengths: the effects of repetition. *Ergonomics* 1984;27:1283-8.

14. Freivalds A, Chaffin DB, Garg A, Lee KS. A dynamic biomechanical evaluation of lifting maximum acceptable loads. J Biomech 1984;17:251-62.
15. Garg A. An evaluation of NIOSH guidelines for manual lifting. Proceedings of the 23rd Annual Conference of the Ergonomics Society of Australia and New Zealand. Perth, Western Australia, 1986.
16. Garg A, Sharma D, Chaffin DB, Schmidler JM. Biomechanical stresses as related to motion trajectory of lifting. Hum Factors 1983;25:527-39.
17. Garg AA, Mital A, Asfour SS. A comparison of isometric strength and dynamic lifting capability. Ergonomics 1980;23:13-27.
18. Hasue M, Fujiwara M, Kikuchi S. A new method of quantitative measurement of abdominal and back muscle strength. Spine 1980;5:163-8.
19. Hirsch T. The billion dollar backache. National Safety News, 1977.
20. Holbrook TL, Grazier K, Kelsey JL, Stauffer RN. The frequency of occurrence, impact and cost of selected musculoskeletal conditions in the United States. Chicago (IL): Am Acad Orthop Surg, 1984.
21. Kamon E, Kiser D, Pytel JL. Dynamic and static lifting capacity and muscular strength of steelmill workers. Am Ind Hyg Assoc J 1982;43:853-7.
22. Keyserling WM, Herrin GD, Chaffin DB. Isometric strength testing as a means of controlling medical incidents on strenuous jobs. J Occup Med 1980;22:332-6.
23. Khalil TM, Asfour SS, Waly SM, Genaidy AM. A comparative study of static and dynamic lifting tasks. Proceeding of the 28th Annual Meeting of the Human Factors Society 1984;595-9.
24. Kishino ND, Mayer TG, Gatchel RJ, Parrish MM, Anderson C, Gustin L, et al. Quantification of lumbar function. Part 4: Isometric and isokinetic lifting simulation in normal subjects and low-back dysfunction patients. Spine 1985;10:921-7.
25. Kroemer KHE. An isoinertial technique to assess individual lifting capacity. Hum Factors 1983;25:493-506.
26. Langrana N, Lee C, Alexander H, et al. Quantitative assessment of back strength using isokinetic testing. Spine 1984;9:287-90.
27. Liles DH, Deivanayagam S, Ayoub MM, et al. A job severity index for the evaluation and control of lifting injury. Hum Factors 1984;26:683-93.
28. Marras WS, King AI, Joynt RL. Measurements of loads on the lumbar spine under isometric and isokinetic conditions. Spine 1984;9:176-87.

29. Marras WS, Ranganajulu SL. Trunk force development during static and dynamic lifts. *Hum Factors* 1987;29:19-29.
30. Marras WS, Reilly CH. Networks of internal trunk loading activities under controlled trunk motion conditions. *Spine* 1988;13:661-7.
31. Marras WS, Wongsam PE. Flexibility and velocity of the normal and impaired lumbar spine. *Arch Phys Med Rehabil* 1986;67:213-7.
32. Mayer TG, Smith SS, Keeley J, Mooney V. Quantification of lumbar function. Part 2: Sagittal plane trunk strength in chronic low-back pain patients. *Spine* 1985;10:765-72.
33. Mayer TG, Smith SS, Kondrask G, Gatchel RJ, Carmichael TW, Mooney V. Quantification of lumbar function. Part 3: Preliminary data on torso rotation testing with myoelectric spectral analysis in normal and low-back pain subjects. *Spine* 1985;10:911-20.
34. McDaniel JW. Prediction of acceptable lift capability [doctoral dissertation]. Lubbock (TX): Texas Tech University, 1972.
35. McGill SM, Norman RW. Dynamically and statically determined low back moments during lifting. *J Biomech* 1985;18:877-85.
36. Mital A, Ayoub MM, Asfour SS, Beatha NJ. Relationship between lifting capacity and injury in occupations requiring lifting. Proceedings of the 22nd Annual Meeting of the Human Factors Society, 1978.
37. Mital AW, Karwowski W, Mazouz A, Orsair E. Prediction of maximum acceptable weight of lift in the horizontal and vertical planes using simulated job dynamic strengths. *Am Ind Hyg Assoc J* 1986;47:288-92.
38. National Institute for Occupational Safety and Health. Work practices guide for manual lifting DHHS (NIOSH). Publication No. 81-122, March 1981.
39. National Safety Council. Accident Facts. Chicago (IL), 1976.
40. Nicolaisen T, Jorgensen K. Trunk strength, back muscle endurance and low back trouble. *Scand J Rehab Med* 1985;17:121-7.
41. Nordin M, Kahanovitz N, Verderame R, Parnianpour M, Santiago Y, Viola K et al. Normal trunk muscle strength and endurance in women and the effect of exercises and electrical stimulation. *Spine* 1987;12:105-11.
42. Parnianpour M, Buchalter D, Nordin M, Kahanovitz N. A non-invasive in-vivo measurement technique for assessment of the six degrees of freedom of trunk motion. In: Lantz SA, King JA, editors. *Advances in Bioengineering*. New York: Am Soc Mech Eng 1986;28-9.

43. Parnianpour M, Nordin M, Frankel V, Kahanovitz N. The triaxial coupling of torque generation of trunk muscles during isometric exertions and the effect of fatiguing isoinertial movements on the motor output and movement patterns. *Spine* 1988;13(9):982-92.
44. Plott CC. Alternative strength testing methods of employee screening [master's thesis]. Lubbock (TX): Texas Tech University, 1983. In press.
45. Poulsen, E. Prediction of maximum loads in lifting from measurement of back muscle strength. *Progressive Therapy* 1970;1:146-9.
46. Rodgers SH. The NIOSH work practices guide to manual lifting. The physiological basis. Paper presented at AIHA Conference, June 6, 1982.
47. Seeds RH, Levene J, Goldberg HM. Normative data for IsoStation B100. *J Orthop Sports Phys Ther* 1987;9(4):141-55.
48. Seeds RH, Levene JA, Goldberg HM. Abnormal patient data for the IsoStation B100. *J Orthop Sports Phys Ther* 1988;10(4):121-33.
49. Smidt G, Herring T, Amundsen L, Roger M, Russel A, Lehman T. Assessment of abdominal and back extensor function. A quantitative approach and results for chronic low-back patients. *Spine* 1983;8:211-9.
50. Smith JL, Ayoub MM. Interim modification of the x-factor test: Bolling AFB (DC): Air Force Office of Scientific Research, 1982. Final Report Contract No. F33625-78-D-0629-0037.
51. Smith JL, Smith LA, McLaughlin TM. A biomechanical analysis of industrial manual material handlers. *Ergonomics* 1982;25:299-308.
52. Smith SS, Mayer TG, Gatchel RJ, Becker TJ. Quantification of lumbar function. Part 1: Isometric and multispeed isokinetic trunk strength measures in sagittal and axial planes in normal subjects. *Spine* 1985;10:757-64.
53. Thorstensson A, Nilsson J. Trunk muscle strength during constant velocity movements. *Scand J Rehab Med* 1982;14:61-8.
54. Troup JDG. Relation of lumbar spine disorders to heavy manual work and lifting. *Lancet* 1965;1:857-61.
55. Wood GA, Hayes KC. A kinematic model of intervertebral stress during lifting. *Br J Sports Med* 1974;8:74-9.

Références additionnelles

THERMOGRAPHIE

1. Abram SE, Asiddao CB, Reynolds AC. Increased skin temperature during transcutaneous electrical stimulation. *Anesth Analg* 1980;59:22-5.
2. Adams WJ, Lloyd JT. Empirical evaluation of the chronic pain diagnosis. Academy of Neuro-Muscular Thermography First Annual Meeting, May 1985; *Postgraduate Medicine* 1986; Special edition:86-9.
3. Addatto KA, Phillips SI, Manale BL, et al. CT & thermogram, a comparison of 91 patients. *Ortho Trans* 1985;9:215.
4. Albert SM, Glickman M, Kalish M. Thermography in orthopedics. *Ann NY Acad Sci* 1964;121:157-70.
5. Ash CJ, Shealy CN, Young PA, et al. Thermography and the sensory dermatome. *Skeletal Radiology* 1986;15:40-6.
6. Barnes RB. The early history of thermography. In: Uematsu S, editor. *Medical thermography: theory and clinical applications*. Los Angeles: Brentwood Publishing Co., 1976:1-14.
7. Bayliss WM. On the origin from the spinal cord of the vasodilator fibres of the hind limb and on the nature of these fibres. *J Physiol* 1901;26:173-209.
8. Beck L. Histamine as the potential mediator or active reflex dilation. *Fed Proc* 1965;24:1298-310.
9. Bell C, Auton J. Vasodilator neurons supplying skin and skeletal muscles of limbs. *J Auton Nerv Syst* 1938;7:257-62.
10. Berstholdt H, Brand P. Thermography: An aid in the management of insensitive feet and stumps. *Arch Phys Med Rehab* 1975;56:205-9.
11. Brelsford KL, Uematsu S. Thermographic presentation of cutaneous sensory and vasomotor activity in the injured peripheral nerve. *J Neurosurg* 1985;62:711-5.
12. Brown RXJ, Bassett LW, Wexler CE, et al. Thermography as a screening modality for nerve fiber irritation in patients with low back pain. Academy of Neuro-Muscular Thermography Second Annual Meeting, September 1986; *J Physiol*. In press.
13. Celander O, Fulkow B. A comparison of the sympathetic vasomotor fibre control of the vessels within the skin and the muscles. *Acta Physiol Scand* 1953;29:241-50.

14. Chafetz N, Wexler CE, Kaiser JA. Neuromuscular thermography of the lumbar spine with CT correlation. *Radiology* 1985;157:78.
15. Chang L, Abernathy M, O'Rourke D, et al. The evaluation of posterior thoracic temperature by telethermography, thermocouple, thermistor and liquid crystal thermography. *Thermology* 1985;1:95-101.
16. Ching C, Wexler CE. Peripheral thermographic manifestations of lumbar disc disease. *Applied Radiology* 1978;100:53-8.
17. Clark RP. Human skin temperature and its relevance in physiology and clinical assessment. In: Francis E, Ring J, Phillips B, editors. *Recent advances in medical thermology*. New York: Plenum Press, 1984:5-15.
18. Cosh JA. Infra-red detection in the assessment of rheumatoid arthritis. *Proc Royal Society Med* 1966;59 Suppl:88-93.
19. Devereaux MD, Graham R, Lachmann SM, et al. The diagnosis of stress fractures in athletes. *JAMA* 1984;252:531-3.
20. Devereaux MD, Parr GR, Lachmann SM, et al. Thermographic diagnosis in athletes with patellofemoral arthralgia. *J Bone Joint Surg* 1986;68:42-4.
21. Doupe J, Cullen CH, Macauley LJ. Studies in denervation. Circulation in skin and proximal parts of limbs. *J Neurol Psychiat* 1943;6:129-41.
22. Duensing F, Becker P, Rittmeyer K. Thermographic findings in lumbar disc protrusions. *Arch Psychiatr Nervenkr* 1973;217:53-70.
23. Eddy HC, Taylor HP. Experiences with the dermatherm (tycos) in relation to peripheral vascular disease. *Am Hart J* 1931;683-9.
24. Edeiken J, Wallace JD, Curley RF, et al. Thermography and herniated lumbar discs. *Am J Roentgenol* 1968;102:790-6.
25. Editorial: Thermography as a diagnostic aid in sciatica: A commentary on experimental methods, data interpretation and conclusions. *Thermology* 1985;1:55-8.
26. Fischer A. Diagnosis and management of chronic pain in physical medicine and rehabilitation. In: Ruskin AP, editor. *Current therapy in physiatry*. Philadelphia: WB Saunders, 1984:123-45.
27. Fischer A, Rim A, Chang C. Correlation between thermographic findings and somatosensory cortical evoked potentials in lumbosacral radiculopathies. *Thermology* 1986;2:29-33.
28. Fischer AA. The present status of neuromuscular thermography. *Academy of Neuro-Muscular Thermography First Annual Meeting, May 1985; Postgraduate Medicine* 1986; Special edition:26-33.

29. Fischer AA, Chang CH. Temperature and pressure threshold measurements in trigger points. *J Thermology* 1986;1:212-5.
30. Fischer AA, Chang CH, Kuo JC. The value of thermography in the diagnosis of radiculopathy as compared with electrodiagnosis. *Arch Phys Med Rehabil* 1983;64:526.
31. Foerster O. The dermatomes in man. *Brain* 1933;56:1-39.
32. Fox RH, Edholm OG. Nervous control of the cutaneous circulation. *Brit Med Bull* 1963;19:110-4.
33. Fraser RG, Pare JAP. Perception in chest roentgenology. In: *Diagnosis of diseases of the chest*. 2nd ed. Philadelphia: WB Saunders, 1977;1:179-83.
34. Fulkow B. Nervous control of blood vessels. *Acta Physiol Scand* 1955;35:629-63.
35. Gandhavadi B, Rosen JS, Addison RG. Autonomic pain. Features and methods of assessment. *Postgrad Med* 1982;71:85-90.
36. Gelfand DW, Ott DJ. Methodologic considerations in comparing imaging methods. *AM J Roentgenol* 1985;144:1117-21.
37. Gillstrom P. Thermography in low back pain and sciatica. *Arch Ortho Traum Surg* 1985;104:31-6.
38. Goldberg G. Thermography and magnetic resonance imaging correlated in 31 cases. *Academy of Neuro-Muscular Thermography First Annual Meeting*, May 1985; *Postgraduate Medicine* 1986; Special edition:54-8.
39. Gross D. Pain and autonomic nervous system. *Adv Neurol* 1974;4:93-103.
40. Guttmann NL. Topographic studies of disturbances of sweat secretion after complete lesions of peripheral nerves. *J Neurol Psychiatry* 1940;8:197-210.
41. Haines SJ, Haines G. Exploring the normal myelogram. *Minn Med* 1985; 68:529-32.
42. Hamilton B. An overview of proposed mechanisms underlying thermal dysfunction, In: Abernathy M, Uematsu S, editors. *Medical thermography*. Washington (DC): Am Acad of Thermog 1986;6-18.
43. Harway R. Thermographic intra-operative monitoring of a lumbar sympathectomy. A difficult pain management case. *Academy of Neuro-Muscular Thermography Second Annual Meeting*, September 1986; *J Physiol*. In press.
44. Harway RA. A comparison of the results of thermography, CT scanning and myelogram evaluations in 66 patients with lumbar pain. *Academy of Neuro-Muscular Thermography First Annual Meeting*, May 1985; *Postgraduate Medicine* 1986; Special edition:73.

45. Head H. On disturbances of sensation with a special reference to the pain of visceral disease. *Brain* 1893;16:1-133.
46. Head H, Campbell AW. The pathway of herpes zoster and its bearing on sensory localization. *Brain* 1900;23:353-523.
47. Henderson H, Hackett M. The value of thermography in peripheral vascular disease. *Angiology* 1978;29:65-75.
48. Henderson HP, Hackett MEJ. Thermographic examination of the legs and the fibrinogen uptake test in the diagnosis of deep-vein thrombosis: A comparison. *Acta Thermographica* 1979;4:104-7.
49. Hendler N, Uematsu S, Long D. Thermographic validation of physical complaints in "psychogenic pain" patients. *Psychosomatics* 1982;23:183-7.
50. Herschel J. On the distribution of the calorific rays in the solar system. *Phil Trans Roy Soc* 1840;131:51-2.
51. Herschel W. Investigations of the powers of the prismatic colors to heat and illuminate objects. *Phil Trans Roy Soc* 1800;90:255-92.
52. Hobbins WB. Thermography and pain. In: *Biomedical Thermology*. New York: Alan R. Liss Inc., 1982:361-75.
53. Howe JF, Loeser JD, Calvin WH. Mechanosensitivity of dorsal root ganglia and chronically injured axons: A physiological basis for the radicular pain of nerve root compression. *Pain* 1977;3:25-41.
54. Hubbard J, Maultsby J, Wexler CE. Lumbar and cervical thermography for nerve fiber impingement: A critical review. *Clinical J Pain* 1986;2:131-7.
55. Hubbard JE. Statistical review of thermography in neurology practice, pain evaluation. *Academy of Neuro-Muscular Thermography First Annual Meeting; May 1985: Postgraduate Medicine* 1986: Special edition:65-72.
56. Hubbard JE, Hoyt C. Pain evaluation by electronic infrared thermography: Correlations with symptoms, EMG, myelogram and CT scan. *Thermology* 1985;1:26-35.
57. Hubbard JE, Hoyt C. Pain evaluation in 805 studies by infra-red imaging. *Thermology* 1985;1:161-6.
58. Ishigaki T, Sakuma S, Okae S, et al. Infra-red imaging in syringomyelia. Correlation with the laterality of the syrinx. *Thermology* 1986;1:135-41.
59. Jaffe CC. Why an image? *Invest Radiol* 1984;19:248-9.
60. Keegan JJ, Garrett FD. The segmental distribution of the cutaneous nerves in the limbs of man. *Anatomical Record* 1948;102:409-37.

61. Kenins P. Identification of the unmyelinated sensory nerves which evoke plasma extravasation in response to autonomic stimulation. *Neurosci Lett* 1981;25:137-41.
62. Kiss F. Sympathetic elements in cranial and spinal ganglia. *J Anat* 1932;66:488-98.
63. Kundel HL. Disease prevalence and radiological decision making. *Invest Radiol* 1982;17:107-9.
64. Kure K, Nitta Y, Tuzi M, et al. Demonstration of special parasympathetic nerve fibers in the dorsal or posterior roots of the lumbar region of the spine. *Quart J Exp Physiol* 1928;18:333-44.
65. Kure K, Saegusa G, Kawaguchi K, et al. On the parasympathetic (spinal parasympathetic) fibers in the dorsal roots and their cells of origin in the spinal cord. *Quart J Exp Physiol* 1930;20:51-66.
66. Lawson RN. Implications of surface temperatures in the diagnosis of breast cancer. *Can Med Assoc J* 1956;75:309.
67. Lee B. Hemodynamic evaluation in selection of amputation level. *Bull Prosthet Res* 1984;Fall:85-104.
68. Lee B, Trainor F, Kavner D, et al. Noninvasive hemodynamic evaluation in selection of amputation level. *Surg Gynecol Obstet* 1979;149:242-4.
69. LeRoy P. Intraoperative thermography. In: Abernathy M, Uematsu S, editors. *Medical thermography*. Hanover (PA): Sheridan Press, 1986:293-5.
70. LeRoy PL, Christian CR, Filasky R. Diagnostic thermography in low back pain syndromes. *Clin J Pain* 1985;1:4-13.
71. Lukeschitsch H, Plattner R, Signumd R, et al. Thermography as a means for quantitative assessment of stump and phantom pains. *Prosthet Ortho Intl* 1984;8:76-81.
72. Mahoney L, McCulloch J, Csima A. Thermography in back pain. *Thermology* 1985;1:43-50.
73. Matles AL. Reflex sympathetic dystrophy in a child: A case report. *Bull Hosp Joint Dis* 1971;32:193.
74. Maultsby JA, Deaton PC, Tannenbaum SI, et al. Electronic thermography: A valuable aid in evaluating and treating symptomatic spinal pathology. Experience with 800 thermographic evaluations. In: Abernathy M, Uematsu S, editors. *Medical thermography*. Hanover (PA): Sheridan Press, 1986:165-9.
75. Maultsby JA, Meek JB, Routon J, et al. Thermography: Its correlation with the pain drawing. *Academy of Neuro-Muscular Thermography First Annual Meeting, May 1985; Postgraduate Medicine* 1986; Special edition:90-2.

76. McLachlan EM, Janig W. The cell bodies of origin of sympathetic and sensory axons in some skin and muscle nerves of the cat hindlimb. *J Comp Neurol* 1983;214:115-30.
77. Melzack R, Wall PD. Pain mechanisms. A new theory. *Science* 1965;150:971-9.
78. Mitchell GAG. Anatomy of the autonomic nervous system. London: E and S Livingstone Ltd; 1985.
79. Mountcastle VB. Medical physiology. 13th ed. St. Louis: CV Mosby, 1974.
80. Nakano K. Liquid crystal contact thermography (LCT) in the clinic evaluation of traumatic low back pain (LBP). *J Neurological Ortho Med Surg* 1984;5:3-6.
81. Nakano K. Liquid crystal contact thermography (LCT) in the evaluation of patients with upper limb entrapment neuropathies. *Neuro Ortho J Med Surg* 1984;5:97-102.
82. Newman R, Seres J, Miller E. Liquid crystal thermography in the evaluation of chronic back pain. *Pain* 1984;20:293-305.
83. Owens S, Atkinson ER, Lees DE. Thermographic evidence of reduced sympathetic tone with transcutaneous nerve stimulation. *Anest* 1979; 50:62-5.
84. Pedersen H, Blunck C, Garner E. Anatomy of lumbosacral posterior rami and meningeal branches of spinal nerves. *Journal BJS* 1956;38A:377-90.
85. Perelman RB, Adler D, Humphreys M. Comparison of lumbosacral thermograms with CT scans. In: Abernathy M, Uematsu S, editors. Medical thermography. Hanover (PA): Sheridan Press, 1986:127-33.
86. Pochaczewsky R, Pillari G, Feldman F. Liquid crystal contact thermography of deep venous thrombosis. *American J Radiology* 1982;138:717-23.
87. Pochaczewsky R, Wexler CE, Meyers PH, et al. Liquid crystal thermography of the spine and extremities. Its value in the diagnostic of spinal root syndromes. *J Neurosurg* 1982;56:386-95.
88. Price J, Mudge AW. A subpopulation of rat dorsal root ganglion neurons is catecholaminergic. *Nature* 1983;301:241-3.
89. Pulst M, Haller P. Thermographic assessment of impaired sympathetic function in peripheral nerve injuries. *J Neurol* 1981;226:35-42.
90. Randall WC, Cox JW, Alexander WF, et al. Direct examination of sympathetic outflow in man. *J Appl Physiol* 1955;7:688-98.
91. Raskin MR, Martin-Lopez M, Sheldon JJ. Lumbar thermography in discogenic diseases. *Radiology* 1976;119:149-52.

92. Richter CP, Woodruff BG. Lumbar sympathetic dermatomes in man determined by the electrical skin resistance method. *J Neurophysiol* 1945;8:323-38.
93. Ring EFJ, Collins AJ, Bacon PA, et al. Quantitation of thermography in arthritis using multi-isothermal analysis. *Am Rheum Dis* 1974;33:353-6.
94. Roberts WJ, Elardo SM. Sympathetic activation of A-delta nociceptors. *Somatosens Res* 1985;33-44.
95. Robicsek R, Masters T, Daugherty H, et al. The value of thermography in the early diagnosis of postoperative sternal wound infections. *Thorac Cardio Surgeon* 1984;32:260-5.
96. Saidman J. *Diagnostic et traitement des maladies de la colonne vertébrale*. Paris: G. Doin, 1949.
97. Sato A, Schmidt RF. Somatosympathetic reflexes: Afferent fibers, central pathways, discharge characteristics. *Physiol Ref* 1973;53:916-47.
98. Schmitt M, Guillot Y. Thermography and muscular injuries in sports medicine. In: Ring EFJ, Phillips B, editors. *Recent advances in medical thermology*. New York: Plenum Press, 1984:439-45.
99. Seguin E. *Medical thermometry and human temperature*. New York: William Wood and Co., 1876.
100. Sherrington CS. Experiments in examination of the peripheral distribution of the fibres of the posterior roots of some spinal nerves I. *Trans Roy Soc Lond* 1892;184:641-763.
101. Sherrington CS. Experiments in examination of the peripheral distribution of the fibres of the posterior roots of some spinal nerves II. *Trans Roy Soc Lond* 1898;190:45-187.
102. Spence V, McCollum P, Walker W, et al. Assessment of tissue viability in relation to the selection of amputation level. *Prosthet Ortho Intl* 1984;8:67-76.
103. Standards for neuro-muscular thermography. Academy of Neuro-Muscular Thermography First Annual Meeting, May 1985; *Postgraduate Medicine* 1986; Special edition:6-7.
104. Stary O. The pathogenesis of discogenic disease. Review of Czechoslovak Medicine 1956;2:1-16.
105. Stary O, Figar S. Die reflexwirkungen nozizeptiver reize im bewegungssegment. Manuelle medizin und ihre wissenschaftlichen grundlagen. In: Wolf-Trier HD. *Verlag fur Psycikalische Medizine*. Heidelberg: 1970: 84-9.

106. Stevens RT. Catecholamine varicosities in cat dorsal root ganglion and spinal ventral roots. *Brain Res* 1983;261:151-4.
107. Stoll AM, Hardy JD. Direct experimental comparison of several sources temperature measuring devices. *Rev Scient Instru* 1949;20:679-86.
108. Stoll AM, Hardy JD. Study of thermocouples as skin thermometers. *J Applied Physiol* 1950;2:531-43.
109. Tichauer ER. The objective collaboration of basic pain through thermography. *J Occ Med* 1977;19:727-31.
110. Uematsu S. Computerized infrared thermographic imaging in evaluation of disorders in the peripheral and central nervous systems. In: Youmans JR, editor. *Neurosurgery*. 3rd ed. Philadelphia: WB Saunders. In press.
111. Uematsu S. Symmetry of skin temperature comparing one side of the body to the other. *Thermology* 1985;1:4-7.
112. Uematsu S. Thermographic imaging of cutaneous sensory segment in patients with peripheral nerve injury. Skin-temperature stability between sides of the body. *J Neurosurg* 1985;62:716-20.
113. Uematsu S, Hendler N, Hungerford D, et al. Thermography and electromyography in the differential diagnosis of chronic pain syndromes and reflex sympathetic dystrophy. *Electromyogr Clin Neurophysiol* 1981; 21:165-82.
114. Uematsu S, Long DM. Thermography in chronic pain. In: Uematsu S, editor. *Medical thermography: theory and clinical applications*. Los Angeles: Brentwood Publishing Co., 1976;52-68.
115. Uricchio J. The clinical use of thermography in orthopedic practice. *Academy of Neuro-Muscular Thermography First Annual Meeting, May 1985; Postgraduate Medicine* 1986; Special edition:62-4.
116. Uricchio J, Walbroel C. Blinded reading of electronic thermography. *Academy of Neuro-Muscular Thermography First Annual Meeting, May 1985; Postgraduate Medicine* 1986; Special edition:47-53.
117. Uricchio JV. Electronic thermography. *J Florida Med Assn* 1983;70:889-94.
118. Wall P, Devor M. Sensory afferent impulses originate from dorsal root ganglion as well as from the periphery in normal and nerve injured rats. *Pain* 1983;17:321-39.
119. Weinstein S, Weinstein G. A clinical comparison of cervical thermography with EMG, CT scanning, myelography and surgical procedures in 500 patients. *Academy of Neuro-Muscular Thermography First Annual Meeting, May 1985; Postgraduate Medicine* 1986; Special edition:44-6.

120. Weinstein SA, Weinstein G. A review of 500 patients with low back complaints: Comparison of five clinically accepted diagnostic modalities. Academy of Neuro-Muscular Thermography First Annual Meeting, May 1985; Postgraduate Medicine 1986; Special edition:40-3.
121. Weinstein SA, Weinstein G. Clinical perspectives in patients with back and neck pain: A review of 2,000 patients. Academy of Neuro-Muscular Thermography Second Annual Meeting, September 1986; J Physiol. In press.
122. Wexler CE. An overview of liquid crystal and electronic lumbar, thoracic, and cervical thermography. Tarzana (CA): Thermographic Services Inc., 1981.
123. Wexler CE. Artifact detection in neuromuscular thermography: Rationale for repeating the neuromuscular thermogram three times at 15-minute intervals as described in the standards of the Academy of Neuromuscular Thermography. Academy of Neuro-Muscular Thermography First Annual Meeting, May 1985; Postgraduate Medicine 1986; Special edition:34-9.
124. Wexler CE. Atlas of thermographic lumbar patterns. Tarzana (CA): Thermographic Services Inc., 1983.
125. Wexler CE. Neuromuscular thermography: Its past, present status and future prospects. Academy of Neuro-Muscular Thermography First Annual Meeting, May 1985; Postgraduate Medicine 1986; Special edition:9-20.
126. Wexler CE. Thermographic evaluation of trauma (Spine). Acta Thermographica 1980;5:3-10.
127. White JC, Sweet WH. Pain and the neurosurgeon: A 40-year experience. Springfield (IL): Charles C Thomas, 1969;87-98.
128. Wiesel SW, Tsourmas N, Feffer H, et al. A study of computer-assisted tomography, the incidence of positive CAT scans in an asymptomatic group of patients. Spine 1984;9:549-51.
129. Williams PL, Warwick R. Functional neuroanatomy of man. Philadelphia: WB Saunders, 1975.
130. Wirth FP, Rutherford RD. A civilian experience with causalgia. Arch Surg 1970;100:633.

ANNEXE III

Sujets refusés

SUJETS REFUSÉS

MOTIFS DE REFUS D'ADMISSION :

Sujets malades après une première évaluation :

- lombalgie antérieure	54
- signes cliniques insuffisants	39
- patients retournés au travail	23
- radiographie anormale	22
- dorsalgie associée	16
- hernie discale lombaire	11
- douleur aux membres inférieurs	11
- spondylolisthesis	8
- chirurgie aux membres inférieurs	3
- accident de travail > 3 mois	3
- grossesse	3
- obésité	2
- non intéressé	1
- âge	1
- autres raisons nuisant à l'étude	39
TOTAL :	236

Sujets malades après une deuxième évaluation, c.-à-d. ne faisant pas le consensus des deux médecins sélectionneurs sur l'état des lombalgiques admissibles selon les critères de l'étude :

- signes cliniques insuffisants	2
- spondylolisthesis	2
- néoplasie	2
- arthrose lombaire trop marquée	1
- coxarthrose	1
- lombalgie antérieure	1
- radiculopathie clinique	1
- lombalgie exagérée	1
TOTAL :	11

Sujets sains refusés après les examens médicaux des médecins sélectionneurs :

- spondylolisthesis	2
- scoliose	1
- sacralisation lombaire	1
- chirurgie aux membres inférieurs	1
- antécédents de lombalgie avec >1 jour absence du travail	1
TOTAL :	6

Sujets sains refusés en raison de tests non terminés :	2
TOTAL :	2

ANNEXE IV

Cahier des formulaires utilisés

(en dépôt à l'IRSST)

LISTE DES FORMULAIRES UTILISÉS

Fiche 1	Consentement
Fiche 2	Échelle subjective d'incapacité fonctionnelle de la lombalgie d'Oswestry
Fiche 3	Échelle subjective d'incapacité de la lombalgie de Roland-Morris
Fiche 4	Échelle visuelle analogue de la douleur de Scott
Fiche 5	Classification des travailleurs(euses) selon l'emploi par les médecins sélectionneurs
Fiche 6	Questionnaire auto-administré aux participants
Fiche 7	Questionnaire de Spinex (Annulé avec l'accord de Spinex)
Fiche 29	Classification selon le patient
Fiche 31	Rapport de radiographie

Médecins sélectionneurs

Fiche 8	Classification des sujets (consensus)
Fiche 8A	Critères d'inclusion et d'exclusion des sujets atteints de lombalgie et des sujets sains (D ^r Roland Proulx)
Fiche 8B	Critères d'inclusion et d'exclusion des sujets atteints de lombalgie et des sujets sains (D ^r Lesly Germain)

Médecins experts

Fiche 9	Classification des sujets - Examen clinique seulement : physiatre (D ^r Luc Fortin)
Fiche 9A	Fiche d'informations - Examen clinique du physiatre (D ^r Luc Fortin)
Fiche 10	Classification des sujets - Examen clinique seulement : neurochirurgien (D ^r Guy Bouvier)
Fiche 10A	Fiche d'informations - Informations du neurochirurgien (D ^r Guy Bouvier)

A124

Fiche 11 Classification des sujets -
Examen clinique seulement : consensus
Physiatre et neurochirurgien

1^{re} lecture : technologie, 1 expert

Fiche 12 Classification des sujets -
Thermographie (D^r Yves Bergeron)

Fiche 12B Fiche d'informations -
Grille d'évaluation pour la thermographie
(D^r Yves Bergeron)

Fiche 13 Classification des sujets -
Thermographie (D^r Roger Vadeboncoeur)

Fiche 13B Fiche d'informations -
Grille d'évaluation pour la thermographie
(D^r Roger Vadeboncoeur)

Fiche 14 Classification des sujets -
B-200 (D^r Marcel Morand)

Fiche 15 Classification des sujets -
B-200 (D^r Jean Fleury)

Fiche 16 Classification des sujets -
Spinoscopie (D^r Serge Gracovetsky) Annulé

Fiche 17 Classification des sujets -
Spinoscopie (D^r Nicolas Newman) Annulé

Fiche 32* Classification des sujets -
Spinoscopie (D^r Patrick Loisel)

Fiche 33* Classification des sujets -
Spinoscopie (D^r Jean-Paul Ouellette)

2^e lecture : technologie + clinique, 1 expert

Fiche 18 Classification des sujets -
Thermographie + examen clinique (D^r Yves Bergeron)

Fiche 18B Fiche d'informations -
Grille d'évaluation pour la thermographie (D^r Yves Bergeron)

Fiche 19 Classification des sujets -
Thermographie + examen clinique (D^r Roger Vadeboncoeur)

- Fiche 19B Fiche d'informations -
Grille d'évaluation pour la thermographie (D^r Roger Vadeboncoeur)
- Fiche 20 Classification des sujets -
B-200 + examen clinique (D^r Marcel Morand)
- Fiche 21 Classification des sujets -
B-200 + examen clinique (D^r Jean Fleury)
- Fiche 22 Classification des sujets -
Spinoscopie + examen clinique (D^r Serge Gracovetsky) Annulé
- Fiche 23 Classification des sujets -
Spinoscopie + examen clinique (D^r Nicolas Newman) Annulé
- Fiche 34* Classification des sujets -
Spinoscopie (D^r Patrick Loisel)
- Fiche 35* Classification des sujets -
Spinoscopie (D^r Jean-Paul Ouellette)

3^e lecture : technologie + clinique, consensus 2 experts

- Fiche 24 Classification des sujets -
Thermographie (consensus)
(D^r Yves Bergeron et D^r Roger Vadeboncoeur)
- Fiche 24B Fiche d'informations -
Grille d'évaluation pour la thermographie (consensus)
- Fiche 25 Classification des sujets -
B-200 (consensus)
(D^r Marcel Morand et D^r Jean Fleury)
- Fiche 26 Classification des sujets -
Spinoscopie (consensus) Annulé
(D^r Serge Gracovetsky et D^r Nicolas Newman)
- Fiche 36* Classification des sujets -
Spinoscopie (consensus)
(D^r Patrick Loisel et D^r Jean-Paul Ouellette)

4^e lecture : technologie + clinique + 1 médecin expert

5^e lecture : technologie + clinique + consensus 2 médecins experts

- Fiche 27 Classification des sujets -
Formulaires des technologies + conclusions antérieures
(D^r Luc Fortin)

A126

Fiche 28 Classification des sujets -
Formulaires des technologies + conclusions antérieures
(D^r Guy Bouvier)

Fiche 30 Classification des sujets -
Formulaires des technologies + conclusions antérieures
(consensus)

Technologies : résultats du B-200 et de la thermographie seulement

* Fiches additionnelles pour les lectures de remplacements :
32, 33, 34, 35, 36

Note : Le canevas des instructions données aux sujets lors de l'entraînement
à la motivation fait suite aux fiches.