

2 m 11. 3527.1

Université de Montréal

Validation de deux méthodes d'évaluation de la douleur aiguë  
en médecine d'urgence

Par Raoul Daoust

Facultés de médecine  
Programme de maîtrise en science biomédicales  
Option recherche clinique

Mémoire présenté à la Faculté des études supérieures en vue de l'obtention du  
grade de Maître ès science (M.Sc.) en sciences biomédicales  
(option recherche clinique)

Janvier 2007  
©, Raoul Daoust, 2007



W

4

058

2007

v. 103

**Direction des bibliothèques**

**AVIS**

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

**NOTICE**

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal

Faculté des études supérieures

Ce mémoire s'intitule :

Validation de deux méthodes d'évaluation de la douleur aiguë  
en médecine d'urgence

Présenté par :

Raoul Daoust

A été évalué par un jury composé des personnes suivantes :

Président-rapporteur  
Pierre Rainville

Directeur de recherche  
Gilles Lavigne

Codirecteur  
Pierre Beaulieu

Membre Jury  
François Madore

Mémoire accepté le :

## REMERCIEMENTS

Ce mémoire n'aurait pas vu le jour sans l'apport important de mon directeur et codirecteur de maîtrise, Gilles Lavigne et Pierre Beaulieu. Ils ont très bien su adapter leurs enseignements et leur approche à ma situation particulière; d'urgentologue, assistant chef de département, directeur de programme, père de famille et étudiant à la maîtrise.

Merci à Christiane Manzini pour sa précieuse collaboration à la conception de l'échelle visuelle analogique plastifiée ainsi qu'aux urgentologues, résidents, étudiants en médecine qui ont participé au recrutement de patients pour cette étude.

Merci au Fonds de recherche en santé du Québec du Dr Gilles Lavigne et au Fonds de recherche des urgentologues de l'urgence de l'hôpital Sacré-Cœur de Montréal.

Et surtout merci à ma famille qui a toléré que je passe plusieurs heures à l'ordinateur et non avec eux.

## TABLE DES MATIÈRES

<u>REMERCIEMENTS</u> .....	III
<u>TABLE DES MATIERES</u> .....	IV
<u>LISTE DES ABRÉVIATIONS</u> .....	VI
<u>LISTE DES FIGURES ET TABLEAUX</u> .....	VII
<u>SOMMAIRE</u> .....	IX
<b><u>1 INTRODUCTION</u></b> .....	<b>1</b>
<u>1.1 IMPORTANCE DE LA MESURE DE L'INTENSITÉ DE LA DOULEUR</u> .....	1
1.1.1 <i>Mesurer la douleur dans un département de médecine d'urgence</i>	1
1.1.2 <i>Impact de la mesure de l'intensité de la douleur</i> .....	3
<u>1.2 MÉTHODE DE MESURE DE L'INTENSITÉ DE LA DOULEUR</u> .....	4
1.2.1 <i>Types d'échelles</i> .....	5
1.2.1.1 L'échelle visuelle analogique.....	5
1.2.1.2 Les échelles numériques.....	7
1.2.1.3 L'échelle catégorielle .....	9
1.2.1.4 Échelle descriptive différentielle .....	11
1.2.1.5 Échelle comportementale (descriptive) .....	12
1.2.1.6 Échelle physiologique.....	12
<u>1.3 ASPECTS PARTICULIERS DE L'ÉVALUATION DE LA DOULEUR EN PÉDIATRIE</u> ..	13
<u>1.4 ASPECTS PARTICULIERS DE L'ÉVALUATION DE LA DOULEUR EN GÉRIATRIE</u> ..	14

<u>1.5 L'INTENSITÉ DE LA DOULEUR COMME OBJECTIF PRIMAIRE EN RECHERCHE CLINIQUE.....</u>	15
<u>1.6 DIFFÉRENCE ENTRE LES SEXES DANS L'ÉVALUATION DE LA DOULEUR .....</u>	16
<u>1.7 INFLUENCE DE LA MÉMOIRE DE LA DOULEUR SUR LES ÉVALUATIONS .....</u>	17
<u>1.8 RÉSUMÉ DES ÉTUDES QUI COMPARENT UNE ÉCHELLE VISUELLE ANALOGIQUE PAPIER À UNE ÉCHELLE NUMÉRIQUE, UNE ÉCHELLE CATÉGORIELLE OU UNE ÉCHELLE VISUELLE ANALOGIQUE PLASTIFIÉE.....</u>	18
<u>1.9 CARACTÉRISTIQUES IMPORTANTES DE L'ÉCHELLE VISUELLE ANALOGIQUE RETENUE POUR LA CRÉATION D'UNE ÉCHELLE VISUELLE ANALOGIQUE PLASTIFIÉE..</u>	24
<u>1.10 MÉTHODE DE VALIDATION DES ÉCHELLES DE MESURE .....</u>	26
1.10.1 <i>Coefficient de corrélation et régression linéaire.....</i>	26
1.10.2 <i>Comparaison des moyennes .....</i>	29
1.10.3 <i>Coefficient de corrélation intraclasse .....</i>	30
1.10.4 <i>Concordance selon la méthode de Bland et Altman.....</i>	31
<u>1.11 OBJECTIF ET HYPOTHÈSE .....</u>	36
<b><u>2 ARTICLE .....</u></b>	<b>38</b>
<b><u>3 DISCUSSION ET CONCLUSION .....</u></b>	<b>53</b>
<b><u>4 BIBLIOGRAPHIE.....</u></b>	<b>61</b>
<b><u>APPENDICE.....</u></b>	<b>73</b>

## LISTE DES ABRÉVIATIONS

EC	: échelle catégorielle à l'aide de visage
EN	: échelle numérique
ENV	: échelle numérique verbale
EVA	: échelle visuelle analogique
EVAp	: échelle visuelle analogique plastifiée
IASP	: association internationale pour l'étude de la douleur

## LISTE DES FIGURES ET TABLEAUX

<b>Figure 1</b> : Échelle visuelle analogique .....	7
<b>Figure 2</b> : Échelle descriptive différentielle.....	11
<b>Figure 3</b> : Échelle visuelle analogique plastifiée.....	25
<b>Figure 4</b> : Exemple de graphique avec la méthode Bland et Altman.....	32
<b>Tableau 1</b> : Études comparant l'échelle visuelle analogique (EVA) à une échelle numérique (EN) ou catégorielle (EC).....	20
<b>Tableau 2</b> : Études comparant l'échelle visuelle analogique (EVA) à une échelle visuelle analogique plastifiée (EVAp).....	23

## SOMMAIRE

**Mots clés** : mesure douleur, validation, échelle numérique, échelle visuelle analogique, corrélation.

La douleur est très fréquente en médecine d'urgence et elle a de nombreux impact; elle module la priorité de triage, la prescription d'analgésie, la présence de soulagement et la perception générale de la qualité des soins. La « Joint Commission on the Accreditation of Health Care Organization » considère que l'évaluation de l'intensité de la douleur devrait faire partie des signes vitaux de tous les patients qui se présentent au département de médecine d'urgence. En clinique l'échelle numérique verbale de 11 points (0 à 10) est la plus utilisée. Par ailleurs l'échelle visuelle analogique est considérée comme l'étalon de référence lors de recherches cliniques. Néanmoins, celle-ci est difficile à utiliser dans le contexte clinique nécessitant, papier, crayon et un appui. Pour pallier à ceci, l'utilisation de versions plastifiées et électroniques de l'échelle visuelle analogique a été suggérée. Malheureusement la méthode de validation la plus fréquente de ses méthodes est le coefficient de corrélation qui a de nombreuses limites comme par exemple le fait d'être sensible à l'écart des résultats chez la population étudiée. Le coefficient de corrélation intraclasse pallie à certaines des lacunes de la corrélation simple mais conserve plusieurs limites. Bland et Altman ont conçu une autre méthode afin de valider un nouvel outil de mesure

qui n'a pas les limites de la corrélation et dont l'utilisation est plus intuitive. Il s'agit de calculer le biais, qui est la moyenne des différences entre les résultats obtenues avec l'étalon de référence et ceux obtenus avec la nouvelle méthode. On utilise aussi les limites de concordance qui sont  $\pm 1.96$  fois l'écart-type du biais, ceci représentant l'intervalle contenant 95% des différences entre les deux méthodes de chaque côté du biais (qui est la moyenne de ses différences). Si ces limites sont plus petites que : la plus petite différence significative ou des limites de la répétabilité (limite de concordance lorsque l'on répète une évaluation avec l'étalon de référence) alors la nouvelle technique peut être utilisée à la place du standard de référence. Nous avons postulé qu'une échelle verbale numérique et une échelle analogique plastifiée seraient équivalentes à une échelle visuelle analogique sur papier. Nous avons donc évalué chez des patients souffrant de douleur aiguë en médecine d'urgence, selon la méthode d'Altman-Bland et la corrélation intraclasse, la concordance d'une échelle numérique verbale de 11 points (0 à 10) et d'une échelle visuelle analogique plastifiée en utilisant comme référence l'échelle visuelle analogique standard sur papier de 100 mm. Les patients de 18 ans et plus, qui se présentait au département de médecine d'urgence d'un hôpital universitaire, ayant une douleur de moins de 24 heures étaient admissibles. Ceux avec altération de l'état de conscience, incapable de comprendre ou d'utiliser les échelles ou ceux dont la qualité des soins pouvait être altérée par la participation à l'étude étaient exclus. Les résidents et enseignants ont recruté 1 176 patients entre août 2005 et janvier 2006, les résultats ont été transformés en une échelle de 0 à 100 unités aux fins

d'analyse. Nous avons démontré une excellente corrélation intraclasse et un faible biais pour l'échelle verbale numérique et l'échelle visuelle analogique plastifiée lorsque comparées à l'échelle visuelle analogique sur papier. Néanmoins, nous avons aussi noté un biais variable et des limites de concordances plus larges que la limite de répétabilité pour l'échelle verbale numérique, ce qui selon la méthode de Bland et Altman rend cette échelle non valide. Nos données suggèrent que l'échelle visuelle analogique plastifiée est valide pour évaluer la douleur aiguë en médecine d'urgence, et que l'échelle numérique verbale malgré une excellente corrélation intraclasse et un faible biais ne devrait pas être utilisée en recherche et avec prudence en clinique.

# **1 INTRODUCTION**

Selon l'Association internationale pour l'étude de la douleur (connu sous l'acronyme IASP pour « International Association for the Study of Pain » en langue anglaise), la douleur est « une expérience sensorielle et émotionnelle désagréable associée à un dommage actuel ou potentiel des tissus ou décrit comme tel ». La douleur est un motif de consultation fréquent à la salle d'urgence, et la mesure de son intensité à de nombreuses conséquences en clinique. Afin de préparer notre revue de la littérature, nous avons utilisé les moteurs de recherche OVID et Pubmed pour questionner les bases de données Medline et Embase en utilisant les mots « pain measure », « pain scale ». Nous avons limité notre revue des écrits aux langue anglaise et française. De plus, nous avons revu la bibliographie de ces articles.

## **1.1 Importance de la mesure de l'intensité de la douleur**

### **1.1.1 Mesurer la douleur dans un département de médecine d'urgence**

Au département de médecine d'urgence, la mesure de l'intensité de la douleur va modifier la priorité assignée, selon l'échelle de triage et gravité canadienne, par l'infirmière du triage et par le fait même influencer le délai

avant l'évaluation par un médecin (Murray, Bullard et al. 2004). De plus, l'infirmière qui n'utilise pas un outil standardisé sous-évalue systématiquement l'intensité de la douleur lorsque celle-ci est comparé à l'autoévaluation faite par le patient et cette différence est d'autant plus importante que la douleur est intense (Rundshagen, Schnabel et al. 1999). Ceci est aussi valable pour tous les intervenants auprès du patient; s'il n'y a pas une évaluation formelle de l'intensité de la douleur, ces derniers utiliseront le comportement du patient, ce qui n'est pas un reflet fiable de l'autoévaluation de l'intensité de la douleur (Whipple, Lewis et al. 1995) (Klopfenstein, Herrmann et al. 2000) (Coulthard, Pleuvry et al. 2000). Il semble que seul le patient connu porteur d'une maladie reconnue comme causant de la douleur intense aura une évaluation plus juste de son intensité même sans utilisation d'un outil (Chibnall, Tait et al. 1997). Même chez l'enfant l'évaluation de l'intensité de la douleur est sub-optimale en médecine d'urgence (Probst, Lyons et al. 2005) (Drendel, Brousseau et al. 2006). Finalement, la « Joint Commission on the Accreditation of Health Care Organization » américaine considère que l'évaluation de l'intensité de la douleur devrait faire partie des signes vitaux de tous les patients qui se présentent au département de médecine d'urgence (Knox 2005).

### 1.1.2 Impact de la mesure de l'intensité de la douleur

Le simple fait d'estimer l'intensité de la douleur en clinique est suivi par la prescription plus systématique et plus rapide d'un analgésique, un soulagement plus efficace et contribue à l'amélioration de la perception de la qualité des soins par le patient et le médecin (Thomas and Andruszkiewicz 2004) (Kovach, Weissman et al. 1999). Cette démarche permet de prédire la nécessité et la quantité de morphine à utiliser en période postopératoire. À titre d'exemple une échelle visuelle analogique (EVA) de 70/100 et plus, est associée à l'utilisation d'une dose de morphine supérieure à 0,15 mg/kg (Aubrun, Langeron et al. 2003). On sait aussi que le patient sera satisfait du soulagement en période postopératoire si on diminue l'intensité de douleur à moins de 40/100 sur l'EVA (Jensen, Martin et al. 2005). La situation est similaire en traumatologie, où l'utilisation d'analgésique augmente lorsque l'on mesure l'intensité de la douleur (Silka, Roth et al. 2004).

Globalement, il semble que si l'infirmière possède de faibles connaissances de la mesure de l'intensité de la douleur, la prise en charge de la douleur sera déficiente (Horbury, Henderson et al. 2005). Mais à la suite d'une formation et à une évaluation systématique de l'intensité de la douleur, il y a une diminution de 64 % de la fréquence de soulagement inadéquat (Salomaki, Hokajarvi et al. 2000).

## **1.2 Méthode de mesure de l'intensité de la douleur**

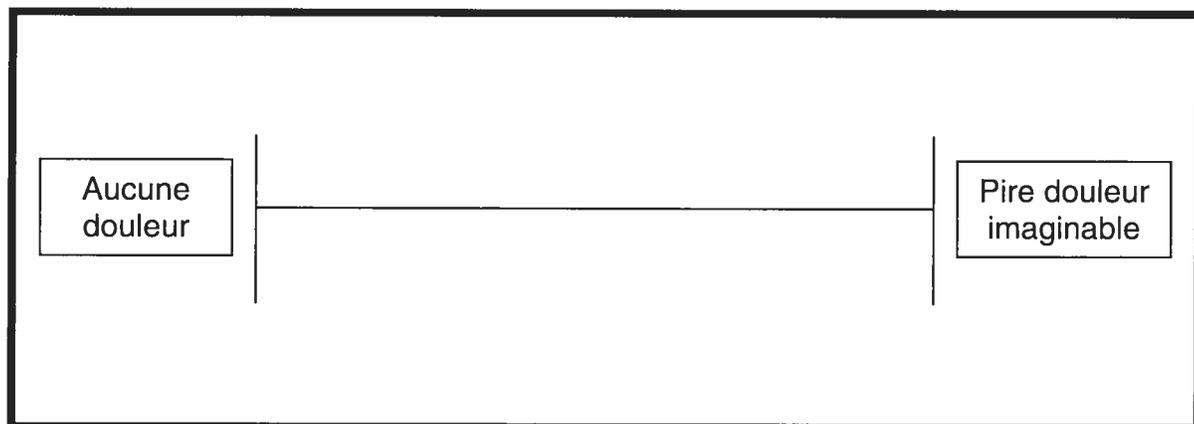
Il existe plusieurs méthodes de mesure de l'intensité de la douleur qui seront décrites dans cette section. Chacune a des avantages et des inconvénients qui doivent être pris en considération selon : le milieu clinique, l'âge, la condition du patient ou selon que l'on effectue une simple évaluation clinique ou une étude randomisée avec traitement témoins ou placebo (Beaulieu 2005) (Jensen and Karoly 2001). (Melzack and Katz 2006) (McGrath and Unruh 2006). Les échelles de type visuel analogique, numérique et catégoriel sont les plus utilisées. On retrouve également des échelles descriptives, différentielles et comportementales (observationnelles), cette dernière étant utilisée pour les patients qui ne peuvent communiquer ou comprendre les autres échelles. Afin d'évaluer l'impact d'un épisode douloureux, la mesure de l'intensité est possiblement l'équivalent d'échelle plus complexe et qui nécessite plus de temps à administrer, comme le McGill pain Questionnaire (Jenkinson, Carroll et al. 1995). Ces types d'échelles sont fréquemment comparés à l'aide d'outils statistiques que nous présenterons ultérieurement (section 1.8).

## 1.2.1 Types d'échelles

### 1.2.1.1 L'échelle visuelle analogique

Il s'agit d'une ligne continue qui est habituellement d'une longueur de 100 mm (ou 10 cm), servant à estimer la douleur perçue par un sujet en un temps donné. Elle est délimitée à chaque extrémité par des mots qui permettent une graduation entre « aucune douleur » et la « pire douleur imaginable », le patient détermine l'intensité de la douleur à l'aide d'un trait coupant la ligne. On mesure alors la distance entre l'extrémité « pas de douleur » et ce trait, et ainsi estimer l'intensité de la douleur (aussi utilisé pour estimer les aspects déplaisants de la douleur) (voir Fig. 1). Cette échelle peut-être présentée de façon verticale ou horizontale et elle est devenue l'étalon de référence en recherche sur la douleur (Beaulieu 2005) car elle est plus sensible aux modifications d'intensité de la douleur (Paul-Dauphin, Guillemin et al. 1999). Un autre avantage de l'échelle visuelle analogique (EVA) est le fait qu'elle est linéaire (Price, McGrath et al. 1983); c'est-à-dire que les espaces sur l'échelle sont équivalents et qu'un changement de 1cm sur une échelle de 10 cm a la même signification, que ce changement soit de 2 à 3 cm ou de 8 à 9 cm, ceci est valable pour les douleurs légères à modérées (Myles, Troedel et al. 1999) et sévères (Myles and Urquhart 2005).

Bien qu'elle soit souvent l'outil pour mesurer la douleur en recherche clinique (Jensen and Karoly 2001) elle est peu utilisée en clinique et encore plus rarement en médecine d'urgence. En milieu clinique, on croit qu'elle est plus difficile à manipuler et que son taux de réponse est plus faible que l'échelle numérique (Briggs and Closs 1999). Cependant, certains auteurs rapportent une facilité d'utilisation identique (Williams, Davies et al. 2000) ou supérieure aux autres échelles (échelle numérique verbale, catégorielle), ceci même chez les patients ayant un trouble cognitif léger ou modéré (Pautex, Herrmann et al. 2005). Toutefois, son usage demande l'utilisation d'un crayon et d'un appui rigide pour la feuille de papier, ce qui a donné naissance à des versions plastifiées (Grossman, Sheidler et al. 1992) (Hagino, Thompson et al. 1996) (Choiniere and Amsel 1996) et électronique (Calam, Jimmieson et al. 2000) (Gaertner, Elsner et al. 2004) (Jamison, Gracely et al. 2002) (Stone, Shiffman et al. 2003). Certains auteurs ajoutent de la couleur à cette EVA plastifiée pour en faciliter l'interprétation, mais l'impact d'ajouter la couleur n'a jamais été étudié.

**Figure 1 :****Échelle visuelle analogique****1.2.1.2 Les échelles numériques**

L'échelle numérique (EN) est composée d'une série de chiffres écrits sur une feuille, ou que l'on présente habituellement verbalement, le « 0 » étant l'absence de douleur et le chiffre le plus élevé représente la pire douleur imaginable. Ensuite, on demande au patient de choisir le chiffre qui représente le mieux l'intensité de sa douleur. L'échelle numérique est habituellement de « 0 à 10 », mais on retrouve aussi le « 0 à 5 » et plus rarement les « 0 à 20 » ou « 0 à 100 ». Celle de 0 à 10 (11 points) est la plus utilisée en clinique, car elle semble plus facile à comprendre et manipuler (Briggs and Closs 1999). De plus, elle semble être la préférée des patients (Williams, Davies et al. 2000) (Choiniere and Amsel 1996). Soulignons qu'elle est facilement utilisable verbalement (ENV) et qu'elle est celle recommandée par un groupe d'experts de l'IASP qui a établi un consensus

sur les méthodes à utiliser lors de l'évaluation de l'efficacité des médicaments analgésiques; ce groupe est nommé IMMPACT (Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials) (Dworkin, Turk et al. 2005).

Ainsi, l'EN la plus étudiée semble être celle de 11 points : selon certains auteurs cette échelle est aussi valable que la EVA (Breivik, Bjornsson et al. 2000) et elle serait tout aussi valide au département de médecine d'urgence (Bijur, Latimer et al. 2003) (Fosnocht, Dahle et al. 2003). Néanmoins, les méthodes de validation de ces études ont des limites importantes que nous allons discuter plus loin. Et à la lumière de données d'autres études, elle semble avoir tendance à surévaluer l'intensité de la douleur, ceci surtout lorsqu'elle est faible ou légère (Carpenter and Brockopp 1995) (Choiniere and Amsel 1996). Par ailleurs, il a été proposé que chez des patients d'un âge moyen de 76 ans, incluant des patients avec déficit cognitif léger à modéré, l'échelle numérique de 21 points (0 à 20) horizontale sur papier serait la plus valide lorsque comparé à : une échelle verbale numérique de cinq points, une échelle catégorielle de sept points et une échelle numérique verticale de 21 points (Chibnall and Tait 2001). Toutefois, chez des patients souffrants de troubles cognitifs sévères une échelle numérique de 5 points serait plus valide qu'une de 10 points (Morrison, Ahronheim et al. 1998). Finalement, une autre étude a mis en évidence que les patients âgés préfèrent utiliser l'échelle catégorielle afin d'évaluer l'intensité de leur douleur (Jones, Fink et al. 2005).

### 1.2.1.3 L'échelle catégorielle

L'échelle catégorielle (EC) est composée d'une série de mots qui décrivent une intensité graduelle de la douleur habituellement en 4 points (mais également décrite en 15 points); elle est soit sur papier soit verbal. Le patient doit choisir entre les catégories descriptives offertes, par exemple : aucune douleur, douleur faible, douleur modérée, douleur forte. Chez l'enfant, les « pokers chip » constituent une autre version de ce type d'échelle qui utilise quatre jetons, chaque jeton correspondant à une partie de la « pire douleur imaginable », donc un jeton est une partie de la pire douleur imaginable, deux jetons sont deux parties, etc., jusqu'à un maximum de quatre jetons (Suraseranivongse, Montapaneewat et al. 2005). L'échelle à l'aide de pictogramme (photos ou dessins de visage) est un autre type d'échelle catégorielle. Il s'agit d'illustrations montrant l'expression du visage de patient avec un niveau de douleur faible à très intense. Ceci est surtout utilisé en pédiatrie (Bulloch and Tenenbein 2002) (Hunter, McDowell et al. 2000) et pour les patients ayant un handicap de lecture ou d'écriture (Skovlund and Flaten 1995). Malgré sa simplicité, les images doivent être expliquées ce qui limite son utilisation chez les très jeunes enfants et les patients avec troubles cognitifs sévères. C'est d'ailleurs dans cette situation que cette méthode semble avoir le plus faible taux de réponse (Pautex, Herrmann et al. 2005). Malheureusement, elle ne semble pas être linéaire, les changements

d'intensité entre les différents niveaux ne sont pas identiques ce qui peut limiter et rendre complexe l'interprétation et l'analyse des résultats (Hunter, McDowell et al. 2000).

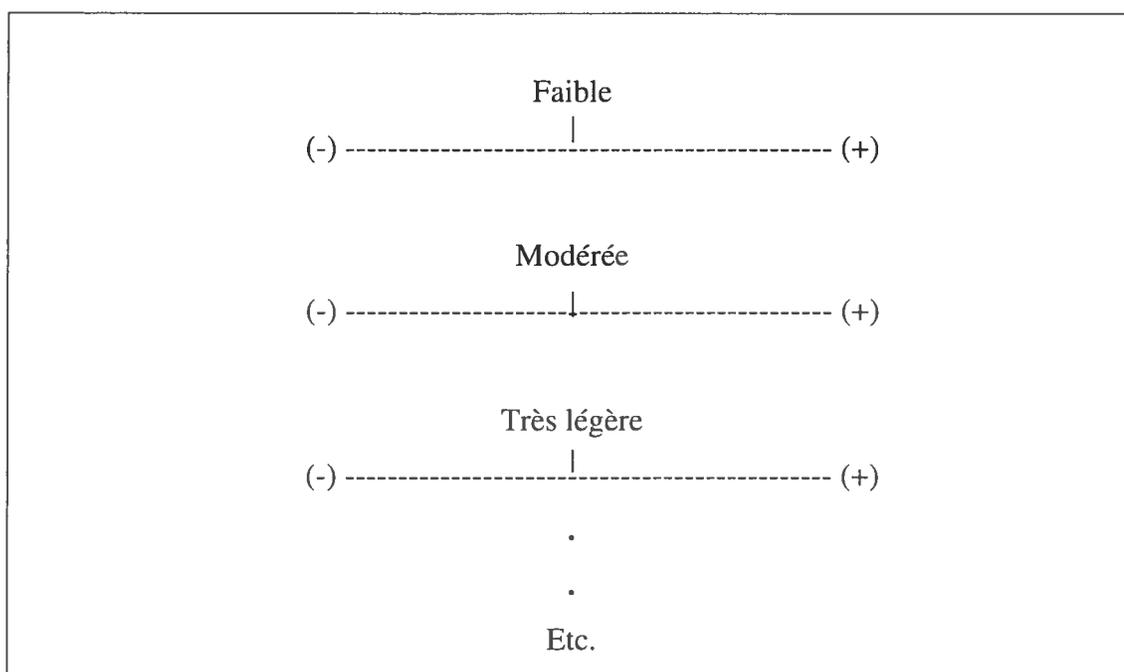
Pour certains auteurs il y a une bonne correspondance entre cette échelle et l'EVA (Collins, Moore et al. 1997) (Lines, Vandormael et al. 2001). Mais les patients doivent apprendre et comprendre les différentes catégories avant de les utiliser et il est possible que le nombre limité de choix d'intensité de douleur ne corresponde pas au niveau de douleur du patient (Joyce, Zutshi et al. 1975). De plus, il existe une large superposition entre les catégories lorsque cette méthode est comparée à une EVA (Averbuch and Katzper 2004) (Bolognese, Schnitzer et al. 2003) (Skovlund and Flaten 1995). De plus, les patients doivent apprendre et comprendre les différentes catégories avant de les utiliser. En général, les échelles catégorielles sont moins utilisées en recherche, car elles sont moins sensibles aux modifications d'intensité de la douleur (Breivik, Bjornsson et al. 2000). Elles sont surtout utiles chez les patients qui ont des troubles cognitifs, car elles s'avèreraient plus faciles à comprendre (Pautex, Herrmann et al. 2005) et c'est souvent celles que ces patients préfèrent (Jones, Fink et al. 2005).

#### 1.2.1.4 Échelle descriptive différentielle

Lors de l'usage de ce type d'échelle, le sujet reçoit une liste de mots décrivant différent niveau d'intensité de douleur (légère modérée, intense, très intense, etc.), pour chaque niveau d'intensité de la liste, le patient doit spécifier sur une échelle si sa douleur est moindre ou supérieure au niveau qu'il évalue (Fig. 2). Cette échelle permet de vérifier la constance des réponses, elle donne possiblement des résultats plus fiables, mais elle est complexe, difficile à comprendre pour les patients et il y a peu de recherche sur sa validité (Turk and Melzack 2001) (Gracely and Kwilosz 1988).

Figure 2 :

#### Échelle descriptive différentielle



### 1.2.1.5 Échelle comportementale (descriptive)

Utilisée surtout en pédiatrie et en gériatrie chez les patients incapables d'utiliser les autres échelles, cette méthode se base sur l'évaluation de l'expression faciale, l'agitation et la vocalisation. Par exemple, chez les patients avec un déficit intellectuel, jusqu'à 35 % d'entre eux sont incapables d'utiliser une EVA (Lachapelle, Hadjistavropoulos et al. 1999). Malheureusement, l'échelle comportementale est peu fiable; on rapporte une corrélation de seulement 0,23 à 0,40 avec l'EVA (Coulthard, Pleuvry et al. 2000) (Pautex, Herrmann et al. 2005) (Choiniere, Melzack et al. 1990). De plus, elle sous-estime la présence de la douleur; en effet, un signe détectable de souffrance par cette méthode étant recensé chez seulement 61 % des patients qui ont de la douleur (Coulthard, Pleuvry et al. 2000).

### 1.2.1.6 Échelle physiologique

Surtout utilisée en néonatalogie, on évalue l'intensité de la douleur en mesurant le pouls, le niveau d'oxygène transcutané, la réponse biologique au stress (ex. : mesure des catécholamines, glucagon, corticostéroïde, hyperglycémie et acidose lactique causées par la suppression de la sécrétion d'insuline) et la diaphorèse. Son utilisation est très limitée, et elle est difficile à valider vu l'absence de standard de référence. Il y a encore moins d'évidences pour l'utilisation du dosage des endorphines, le rythme respiratoire et la pression artérielle pour évaluer la présence et l'intensité de la douleur (McGrath and Unruh 2006). Enfin, l'imagerie semble pouvoir offrir

des avenues prometteuses pour corréler la perception de la douleur à diverses régions sensorielles ou émotives du cerveau, mais ceci dépasse le contexte théorique et pratique de ce mémoire.

### **1.3 Aspects particuliers de l'évaluation de la douleur en pédiatrie**

En pédiatrie, environ 50 % des enfants souffrants qui se présentent au département de médecine d'urgence vont recevoir de l'analgésie (McCarthy, Hewitt et al. 2000) et le soulagement de la douleur est jugé insuffisant chez 80 % de ceux-ci (Ferrell and Rhiner 1991). Cette sous-utilisation de l'analgésie serait plus importante que celle rapportée chez l'adulte (Walco, Cassidy et al. 1994). La situation est similaire dans la période postopératoire où l'utilisation de l'analgésie est insuffisante (Beyer and Bournaki 1989), ceci étant exacerbé par le fait qu'on surestime l'efficacité de cette analgésie (Romsing, Moller-Sonnergaard et al. 1996).

En pédiatrie, l'infirmière utilise peu (~ 1/3 des cas) les échelles de mesure d'intensité de la douleur, malgré le fait que comme chez l'adulte, leur évaluation basée sur le comportement est peu fiable, lorsque comparé à l'autoévaluation à l'aide de EVA. Par exemple, on rapporte une corrélation de 0,579 entre l'évaluation de l'intensité de la douleur selon le comportement, et l'autoévaluation de la douleur à l'aide d'une échelle visuelle analogique plastifiée (EVAp) (Colwell, Clark et al. 1996) L'observation du comportement

est souvent la méthode privilégiée d'évaluation de l'intensité de la douleur, mais elle en sous-estime l'intensité surtout si cette dernière est intense (Beyer, McGrath et al. 1990). Par ailleurs, il est possible que les parents avec des outils d'observations systématisée puissent évaluer l'intensité de la douleur de leur enfant de façon adéquate (Finley, Chambers et al. 2003). Les échelles catégorielles à l'aide de pictogrammes d'expression de visage comportant différentes intensités de douleur sont également utilisées dans ce contexte (Bulloch and Tenenbein 2002). Finalement, la sous-évaluation de la douleur et le soulagement insuffisant en pédiatrie sont de plus en plus jugés non éthiques, d'où la nécessité d'une évaluation et d'un soulagement systématique de la douleur (Walco, Cassidy et al. 1994).

#### **1.4 Aspects particuliers de l'évaluation de la douleur en gériatrie**

Contrairement à la croyance générale, la prévalence de la douleur est fréquente dans les hôpitaux gériatriques, où l'on y recense jusqu'à 43 % de patients souffrant (Cornu, Chah-Wakilian et al. 1997). Un autre mythe est que la personne âgée ne peut utiliser une méthode d'autoévaluation de la douleur. Mais, même chez des patients âgés atteints de troubles cognitifs légers à modérés, on a démontré que 65 % de ces patients utilisent adéquatement une échelle numérique sur papier de 21 points. Celle-ci serait plus valide dans cette situation qu'une échelle numérique verbale de 5 points ou qu'une échelle catégorielle de visage illustrant différente intensité de

douleur (ECV) (Chibnall and Tait 2001). De façon paradoxale, et tel que décrit plus haut, il est rapporté que les patients âgés préfèrent les échelles verbales descriptives plutôt que les échelles numériques ou les ECV (Jones, Fink et al. 2005), mais ils auraient quand même plus de facilité à utiliser les outils non verbaux (Alexander, Plank et al. 2005).

### **1.5 L'intensité de la douleur comme objectif primaire en recherche clinique.**

La mesure de l'intensité de la douleur est essentielle aux études portant sur les analgésiques, car les échelles basées sur le soulagement ont peu de sensibilité ou spécificité pour en quantifier l'efficacité (Mamie, Morabia et al. 2000). D'ailleurs, certains patients rapportent un soulagement malgré une augmentation de l'intensité de la douleur (Feine, Lavigne et al. 1998). La mesure de l'intensité de la douleur serait aussi sensible pour détecter les impacts d'une intervention que des formes plus complexes d'évaluations comme le questionnaire de la douleur de McGill (Jenkinson, Carroll et al. 1995). D'autres auteurs croient que la mesure de l'intensité est insuffisante pour évaluer une intervention et que l'on doit aussi en mesurer l'impact fonctionnel, mais ceci est difficilement applicable à la douleur aiguë qui se présente à la salle d'urgence (Turk, Dworkin et al. 2003) (Chapman and Dunbar 1998).

D'ailleurs, les résultats d'une étude peuvent varier énormément selon la méthode d'évaluation de la douleur utilisée. En effet, on a comparé la mesure de l'intensité de la douleur, une échelle de soulagement, une échelle de satisfaction, et trois échelles de prise en charge de la douleur (mettant en relation la prise de médication et l'intensité de la douleur) et ils ont démontré que l'efficacité du même traitement peut varier de 16 % à 91 % selon la méthode de mesure retenue (de Wit, van Dam et al. 1999). Le choix de la méthode d'évaluation de l'efficacité d'une modalité de traitement est donc primordial et la mesure de l'intensité semble la plus applicable dans le contexte de la médecine d'urgence.

### **1.6 Différence entre les sexes dans l'évaluation de la douleur**

Les femmes rapporteraient une intensité plus grande de douleur lors d'études cliniques, en particulier pour les douleurs postopératoires (Morin, Lund et al. 2000) (Rosseland and Stubhaug 2004) et lors d'étude avec un stimulus de douleur contrôlé (Feine, Bushnell et al. 1991). Néanmoins, d'autres études surtout dans le domaine de la douleur chronique ne rapportent pas plus de douleur chez la femme (Berkley 1997) et la différence significative minimale ne serait pas influencée par le sexe (Kelly 1998) . De plus, l'impact émotionnel de cette douleur serait identique entre les deux sexes (Unruh, Ritchie et al. 1999). Le sexe du patient est donc un facteur confondant dont

on doit probablement tenir compte dans le cadre d'études expérimentales, mais son impact demeure controversé.

### **1.7 Influence de la mémoire de la douleur sur les évaluations**

Il est essentiel d'obtenir une évaluation de l'intensité durant l'épisode douloureux puisqu'il est probable que la mémoire de la douleur soit à la source de biais important. A notre connaissance, une seule étude rapporte une bonne corrélation entre l'évaluation actuelle de la douleur et son rappel (Babul, Darke et al. 1993). Par contre, on rapporte que lorsque l'évènement est chargé d'émotion l'intensité de la mémoire de la douleur serait augmentée (Everts, Karlson et al. 1999), et la situation serait similaire en pédiatrie (van den Brink, Bandell-Hoekstra et al. 2001). La mémoire de la douleur est également influencée par l'intensité initiale de la douleur, lorsqu'elle est faible, elle a tendance à augmenter alors que lorsqu'elle est forte son évaluation va diminuer avec le temps (Feine, Lavigne et al. 1998) (Smith, Gracely et al. 1998).

De plus, au moment du rappel, si la douleur est faible on sous-estime l'ancienne douleur et si au moment du rappel elle est intense c'est l'inverse qui semble se produire (Eich, Reeves et al. 1985) (Smith and Safer 1993). Soulignons que la mémoire de la douleur va aussi varier selon le site de la

douleur. En effet, le rappel d'une douleur au niveau de la colonne cervicale et lombaire serait plus juste que le rappel au niveau abdominal et de la tête (Jamison, Sbrocco et al. 1989). On note aussi une situation analogue chez l'enfant dont le rappel de l'intensité de la douleur est influencé par l'âge, la dépression et l'intensité de la douleur initiale (van den Brink, Bandell-Hoekstra et al. 2001). Il serait donc peu fiable de prendre une échelle de soulagement pour valider une échelle qui mesure l'intensité de la douleur.

### **1.8 Résumé des études qui comparent une échelle visuelle analogique sur papier (EVA) à une échelle numérique (EN), une échelle catégorielle (EC) ou une échelle visuelle analogique plastifiée (EVAp)**

Chez l'adulte en médecine d'urgence, c'est donc l'échelle numérique (EN) qui est la plus utilisée, mais à la lumière des avantages et des inconvénients déjà discutés précédemment, une EVA semble préférable et une EVAp serait probablement plus pratique. Le Tableau 1 résume les études qui comparent EVA à une EN ou une EC. Voir le Tableau 2 qui est un résumé des études qui comparent une EVA à une EVA plastifiée. La principale limite de toutes ces études est leur méthode de validation dont nous allons discuter ultérieurement.

Les EVAp semblent avoir une excellente corrélation avec l'EVA qui est en général reconnue comme l'étalon de référence. La corrélation entre les EN et l'EVA semble également excellente, mais on note la possibilité que l'EN surestime la douleur (comparé à l'EVA) lorsqu'elle est de faible intensité. A la lumière de ces informations, une EVAp ou une EN semblent des alternatives acceptables à la EVA, avec une possible surestimation de la faible intensité de douleur avec une EN. L'étude de Price ne se retrouve pas dans le tableau 2 car elle utilisait un stimulus expérimental standardisé pour évaluer le coefficient de corrélation entre l'EVAp, une EN et ce stimulus chez 23 patients (Price, Bush et al. 1994). Toutefois, aucune corrélation entre les différentes méthodes n'a été effectuée, à la lumière de l'excellente corrélation entre l'EVAp, EN et le stimulus douleur, il est probable que la corrélation soit excellente entre les différentes méthodes (Price, Bush et al. 1994). Dans la deuxième partie de l'étude de Price, dix patients ont noté leur intensité de douleur sur une EVA papier et une EVAp. On y présente un graphique qui tend à démontrer des résultats similaires, mais aucune analyse statistique n'a été effectuée.

**Tableau 1 :**  
Études comparant l'échelle visuelle analogique (EVA) à une échelle numérique (EN) ou catégorielle (EC)

Études EVA vs	Méthode de comparaison	Nombre de patients	Résultats	Commentaires
<b>EN (verbale) 5 points</b>	Corrélation ----- Moyenne EVA (10cm) : 4,28	50	0,78 ----- EN sur 5 : 2,9	EN en moyenne plus élevée que l'EVA lorsque l'unité du dénominateur est identique, mais il est difficile d'en tirer une conclusion utile.
Carpenter et al 1995	Médiane EVA (10cm) : 3,85		3	Dans un tableau de l'étude qui donne le détail des résultats; on note une tendance à la surestimation des résultats avec une EN lors de douleur de faible intensité (< 4/10).
<b>EN (verbale) 11 points et EC**</b>	Comparaison moyenne/catégories : EVA (10cm) <i>Faible</i> ..... 0,9 <i>Agaçante</i> ..... 2,6 <i>Irritante</i> ..... 3,6 <i>Cruelle</i> ..... 8,4 <i>Atroce</i> ..... 8,8	243**	EN sur 10 1,9 3,5 4,4 8,5 9,0	L'EN semble démontrer une tendance à surestimer l'intensité de la douleur lorsqu'elle est faible, agaçante ou irritante.  Il n'y a pas de présentation de la signification statistique pour chaque catégorie dans le tableau de l'article.
Choinière et al 1996				

\*\* Résultats tirés de l'étude déjà présentée dans le tableau 1

**Tableau 1 (suite 1) :**

Études comparant l'échelle visuelle analogique (EVA) à une échelle numérique (EN) ou une échelle catégorielle (EC)

Études EVA vs	Méthode de comparaison	Nombre De Patients	Résultats	Commentaires et limites
<b>EN (verbale) 11 points</b> DeLoach et al 1998	Corrélation à trois moments différents	60	0,94 0,91 0,95	Il s'agit d'une étude effectuée en période postopératoire.
<b>EN (verbale) 11 points</b> -----	Corrélation Cas de trauma Cas de non-trauma ----- Comparaison moyenne/catégories :	200 90	0,795 0,911 ----- Résultats Moyen sur EVA	Il s'agit d'une étude effectuée en France et les cas de trauma ont été recrutés une journée différente des cas de non-trauma. L'EVA utilisée ici est plastifiée et les patients de 15 ans et plus étaient éligibles.
Berthier et al 1998	<i>Cas de trauma</i> Légère..... Modéré..... Sévère ..... Intolérable.....  <i>Cas non-trauma</i> Légère..... Modéré..... Sévère ..... Intolérable.....		21,8 42,7 61,0 67,1  18,5 39,9 63,1 80,8	La différence entre la douleur « sévère » et « intolérable » n'était pas significative pour les cas de trauma.

**Tableau 1 (suite 2) :**

Études comparant l'échelle visuelle analogique (EVA) à une échelle numérique (EN) ou une échelle catégorielle (EC)

Études EVA vs	Méthode de comparaison	Nombre De Patients	Résultats	Commentaires et limites
<b>EN (papier) 11 points</b> Breivik et al 2000	Corrélation	28	0,94	L'EN a tendance à être plus élevé que l'EVA. De plus, il y a une grande variabilité de VAS pour chaque point de l'EN
<b>EN (papier) 11 points</b> Williams et al 2000	Descriptive	78		Les patients préfèrent l'EN
<b>EN (verbale) 11 points</b> Bijur et al 2003	Corrélation	108	0,94	Une étude faite au département médecine d'urgence
<b>EN (papier) 11 points</b> ----- <b>EC</b> Van Laerhoven et al 2004	Corrélation ----- Corrélation ----- Descriptive	120	0,80 ----- 0,76 -----	Une étude réalisée en pédiatrique  Les enfants préfèrent l'échelle catégorielle

**Tableau 2 :**  
Études comparant l'échelle visuelle analogique (EVA) à une échelle visuelle analogique plastifiée (EVAp)

Études EVA vs	Méthode de comparaison	Nombre de patients	Résultats	Commentaires et limites
<b>EVAp</b> Grossman et al 1992	Corrélation	71	0,99	Répétabilité (test-retest) $r= 0,97$ pour l'EVA
<b>EVAp</b> Hagino et al 1996	Corrélation intraclasse	36	0,86	Pas d'effet du site anatomique de la douleur
<b>EVAp</b>	Corrélation	65	0,85	
	----- Comparaison moyenne/catégories :	----- 243	----- EVAp (10 cm)	Pas de présentation de la signification statistique pour chaque catégorie dans le tableau de l'article.  Il y a eu de multiples tests statistiques, mais l'ajustement de Bonferroni qui aurait été fait n'est pas présenté.
Choinière et al 1996	<i>Faible</i> ..... <i>Agaçante</i> ..... <i>Irritante</i> ..... <i>Cruelle</i> ..... <i>Atroce</i> .....		1,2 2,7 3,6 8,4 8,9	

### **1.9 Caractéristiques importantes de l'échelle visuelle analogique retenue pour la création d'une échelle visuelle analogique plastifiée**

Scott et Huskisson (Scott and E.C. 1976) ont démontré que plusieurs aspects de l'EVA influencent sa validité. La simple insertion de mots sur l'EVA la transforme en échelle catégorielle, en créant des points d'ancrage. Il est probable que la couleur et la forme de l'échelle aient les mêmes impacts. De plus, pour être pratique à l'urgence cette échelle devrait être portable, ne pas nécessiter de crayons ou papier et être facile d'interprétation. Nous avons donc conçu une EVA plastifiée en respectant les principes suivants :

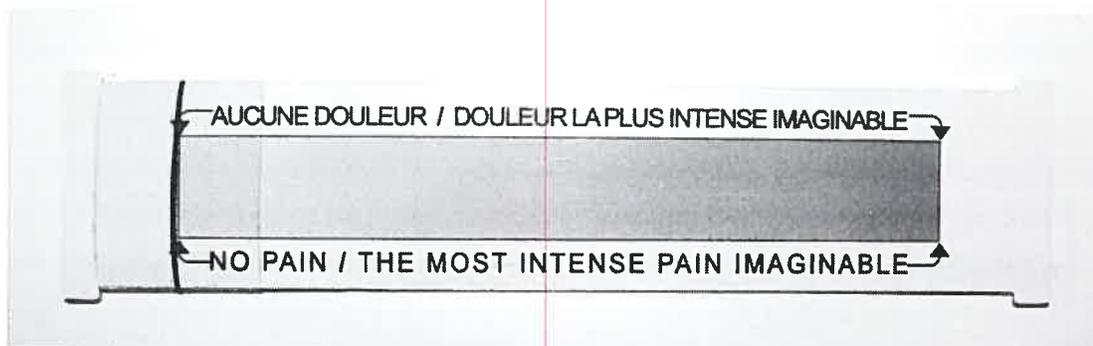
- Aucun point d'ancrage par le texte, la couleur, la forme, etc.
- Bilingue (français, anglais), ceci correspond au contexte social de Montréal.
- Manipulation facile par le patient.
- Lecture facile et sans ambiguïté du résultat.

Les caractéristiques que nous avons retenues sont :

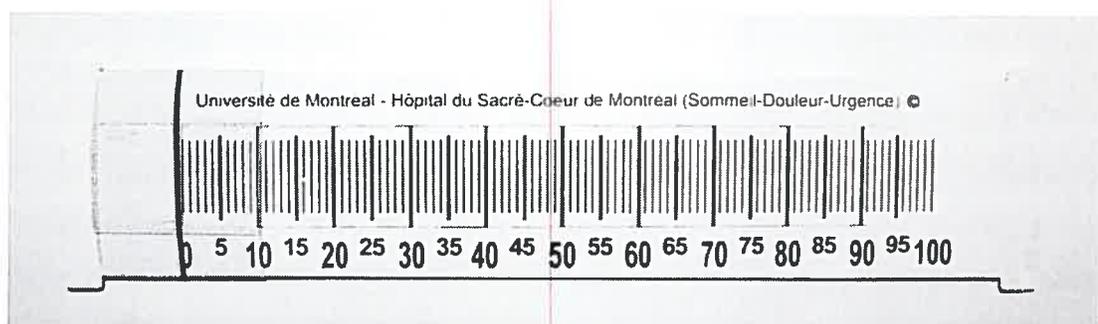
- Explication de l'échelle au dessus en français et en dessous en anglais, manipulation des espaces et polices des caractères pour éviter les points d'ancrages.
- Utilisation d'intensités de rouge dans l'échelle, ceci sans autre couleur afin d'éviter les points d'ancrages et un volume uniforme de l'échelle.
- Chiffres au dos afin de faciliter la lecture du résultat de l'échelle avec un espace entre les lignes ne laissant aucun doute sur son interprétation.

**Figure 3 :**

EVA aspect qui est présenté au patient :



EVA aspect pour la lecture du résultat :



## **1.10 Méthode de validation des échelles de mesure**

La méthode de validation d'une échelle pour mesurer la douleur est une décision importante et qui doit être basée sur les avantages et les inconvénients de chacune des méthodes. L'étalon de référence reconnu est l'EVA sur papier. Voici un résumé des différentes techniques utilisées pour valider des échelles d'intensité de la douleur avec leurs principales caractéristiques.

### **1.10.1 Coefficient de corrélation et régression linéaire**

Il est important de faire la distinction entre la corrélation et la concordance. La corrélation est l'importance de l'association entre un résultat « A » avec nouvelle méthode de mesure et le résultat « B » de la méthode de référence, mais ne dit rien sur la proximité des résultats « A » et « B », qui est en fait la caractéristique utile en clinique. Par exemple, une nouvelle échelle qui donne systématiquement des résultats équivalant à deux fois la valeur de l'échelle de référence aura un excellent coefficient de corrélation, mais l'utilisation de la nouvelle technique sera cliniquement inacceptable (Porter 1999) (Bland and Altman 1986). La concordance est une évaluation de la

capacité de la nouvelle méthode de donner le même résultat que la méthode de référence, qui est en fait ce que l'on recherche en clinique.

Le coefficient de corrélation et la régression linéaire simple sont une mesure de la corrélation (association linéaire) entre deux variables continues. Le coefficient de corrélation et la régression sont descriptifs, mais la régression permet aussi de concevoir un modèle prédictif (Porter 1999). La régression linéaire est très rarement utilisée (Bijur, Latimer et al. 2003), on y considère que les deux méthodes sont équivalentes si la pente est de 1 et que l'intercepte de l'axe y est 0, il y aurait donc à ce moment corrélation et concordance. Cette situation idéale est très rare il est donc habituellement difficile de faire une extrapolation en clinique de la formule de régression. Celle-ci a également les mêmes limites que le coefficient de corrélation (discutée dans cette section), qui est la méthode de validation habituellement utilisée pour les échelles de douleur (Porter 1999).

Un coefficient de corrélation supérieur à 0 implique que lorsqu'une variable augmente l'autre aura aussi tendance à augmenter. Par ailleurs, un coefficient de corrélation négatif implique que si une variable augmente l'autre aura tendance à diminuer. Si le coefficient est nul, alors il n'y a pas de corrélation et s'il est de 1 alors on peut prédire exactement la valeur de la deuxième variable, mais ce ne sera pas nécessairement la même valeur.

Il est aussi démontré qu'une population ayant des résultats hétérogènes (ex. : une population avec des niveaux de douleur qui couvrent toute la dispersion de l'échelle de mesure) présentera une meilleure corrélation qu'une population homogène (ex. : une population avec un niveau de douleur près de 50 mm sur 100 mm) (Gill, Beevers et al. 1985). D'autre part, plus les résultats de la population étudiée sont différents ou étendus (p. ex : un groupe qui présente une étendue de l'intensité de la douleur de 0 à 100 mm) plus le coefficient de corrélation aura tendance à être grand. Plus les résultats sont agglutinés ou similaires (p. ex : étendue de l'intensité de la douleur entre 40 et 60 mm) plus le coefficient de corrélation aura tendance à être petit (Bland and Altman 1996). De plus, l'inversion de l'ordre de présentation des valeurs va fréquemment modifier le coefficient de corrélation. Autre lacune, la corrélation considère que l'étalon de référence est nécessairement un reflet fiable de la valeur réelle du résultat alors que la moyenne des deux résultats a plus de chance d'être près de la réalité (Bland and Altman 1996). Pour pallier ces lacunes, Bland et Altman ont introduit la notion de biais et ses limites de concordance tel que discuté ci-dessous.

### 1.10.2 Comparaison des moyennes

Certaines études ont tenté de valider des échelles en comparant des moyennes selon certaines catégories d'intensité de douleur (Choiniere and Amsel 1996) (Berthier, Potel et al. 1998). Cette méthode a également des limites; par exemple, plusieurs patients avec une douleur à 50 mm sur une échelle de 100 mm mesurés avec l'outil A pourraient donner des résultats très variables avec l'outil B. Mais si ces résultats varient de chaque côté du 50 mm ( ex ; de 20 à 80 mm) il est possible que la moyenne des résultats avec la méthode B soit également près de 50, mais la méthode B est clairement non valide. Les données de l'article de Bland et Altman nous donnent un exemple de ceci; la moyenne des mesures avec le débit de pointe standard est de 450 litres/minute et la moyenne avec la nouvelle méthode est de 452 litres/minute. Ces moyennes nous donnent l'impression que les deux méthodes nous donnent des résultats similaires, mais on démontre dans l'article des limites de concordance large inacceptable en clinique (Bland and Altman 1986). Witter note des problèmes similaires avec l'utilisation de moyenne comme objectif primaire en recherche. A titre d'exemple, dans une étude sur l'effet d'une médication sur la douleur, il est possible que la douleur diminue pour certains patients et augmente pour d'autres et que l'on obtienne des moyennes similaires. Dans cette étude, si on considère la variation individuelle de la douleur, il est possible que la médication ait eu un effet chez une proportion significative de patients (Witter, Simon et al. 2003).

### 1.10.3 Coefficient de corrélation intraclasse

Dans plusieurs études présentées au **point 1.8**, des paires de résultats pour une même intensité de douleur ont été utilisées, suivi de calculs de leur coefficient de corrélation simple. Les limites du coefficient de corrélation ont déjà été discutées : il est malheureusement souvent influencé par l'étendue des résultats chez les sujets, l'ordre de présentation des données et il mesure seulement la force de l'association entre les deux méthodes. Par exemple : un groupe qui présente une étendue de l'intensité de la douleur de 0 à 100 mm aura tendance à avoir un coefficient plus grand et si l'étendue de l'intensité de la douleur est de 40 et 60 mm, le coefficient aura alors tendance à être plus petit. Le coefficient de corrélation intraclasse, qui est aussi une mesure de l'association linéaire, corrige plusieurs de ces problèmes car il mesure la reproductibilité chez un même sujet et il fait la moyenne des coefficients de corrélations obtenus lors de l'inversion des paires. Il compare globalement la variabilité des mesures chez un sujet, à la variabilité entre les sujets. Ainsi, il est donc supérieur au coefficient de corrélation simple pour l'évaluation de différentes méthodes de mesure de l'intensité de la douleur (Bland and Altman 1996) (Rosner 2000).

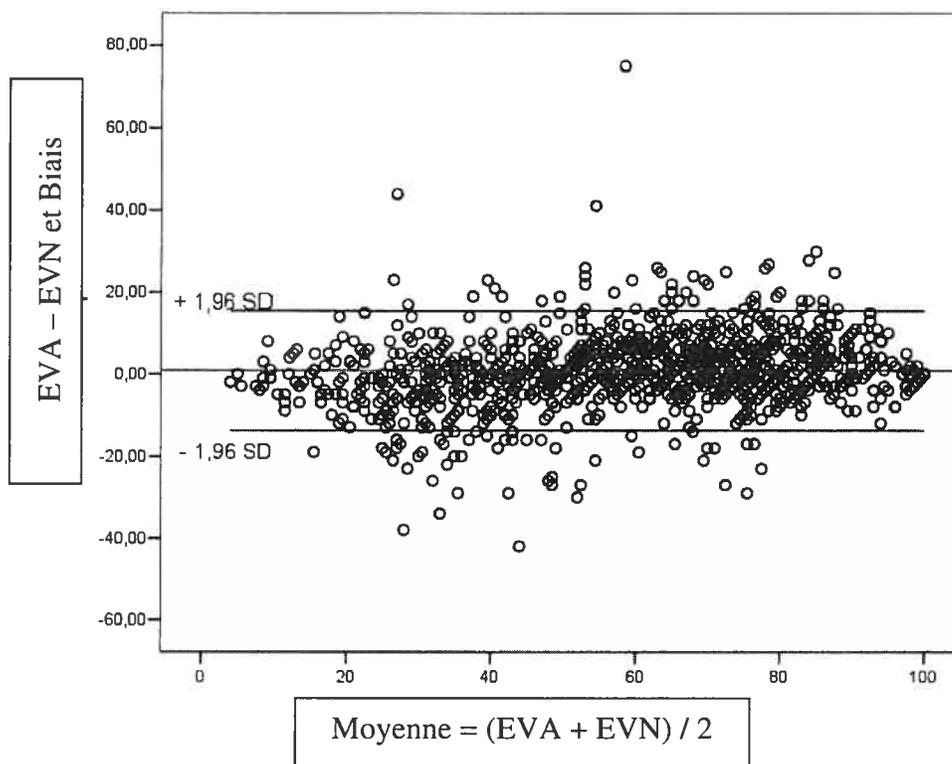
#### 1.10.4 Concordance selon la méthode de Bland et Altman

Tel que mentionné précédemment, la concordance entre les résultats de la méthode que nous tentons de valider et l'étalon de référence est ce qui nous intéresse en clinique. Bland et Altman ont démontré à l'aide d'un exemple portant sur les débits expiratoires maximaux, qu'un nouvel outil peut posséder un excellent coefficient de corrélation avec l'outil de référence, mais qu'il peut souffrir d'un biais ou des limites de concordance qui le rende non valide en clinique (Bland and Altman 1986). Le *biais*, est une moyenne des différences entre les résultats de la nouvelle méthode et ceux de l'étalon de référence. Le résultat de la nouvelle technique additionné au biais est ce qu'on devrait obtenir avec l'outil de référence. Les différences peuvent varier de façon plus ou moins grande autour de la moyenne de ces différences (biais). Pour quantifier cette variation, on utilise les *limites de concordances* qui représentent l'intervalle qui contient 95 % des différences. En pratique ceci représente  $\pm 1.96$  fois l'écart-type du biais. Ceci permet également d'évaluer la constance du biais sur l'étendue de l'intensité de la douleur, car il est possible que la différence entre les deux techniques varie différemment si l'intensité de la douleur est faible versus intense. Les résultats sont souvent présentés sous une forme graphique avec les différences entre les deux méthodes et leur moyenne (le biais) sur l'axe des « x », et la moyenne des deux techniques sur l'axe des « y ». On y ajoute deux lignes qui représentent les limites de concordances (Fig. 4).

## Figure 4

### Exemple de graphique avec la méthode Bland et Altman

Chaque point représente la différence entre le résultat avec l'EVA et le résultat avec l'ENV sur une échelle de 0 à 100 unités. La ligne rouge représente la moyenne des différences (le biais), les deux lignes noires sont les limites de concordances.



On peut aussi calculer le biais et les limites de concordances en répétant la même technique. On obtient alors un biais qui devrait être de zéro s'il n'y a pas un apprentissage entre les répétitions et des limites de concordances qui deviennent les limites de la répétabilité intra-test, aussi nommé limite de reproductibilité intra-test. Ces limites de répétabilité intra-test sont utilisées pour valider un nouvel outil, si les limites de concordance ce celui-ci sont identiques ou plus petites que les limites de répétabilité intra-test, alors l'outil est valide.

Cette technique est basée sur certains principes :

- 1 *Le biais* : représente la moyenne des différences entre le résultat de la nouvelle méthode et la méthode de référence. Il doit être constant, donc ne pas varier selon le niveau d'intensité de la douleur, pour être utilisable.
- 2 La moyenne entre le résultat obtenu par la nouvelle méthode et l'ancienne est plus représentative de la réalité et doit être utilisée sur l'axe des « x » pour l'interprétation du biais.
- 3 Si les *limites de concordances*, établies par deux écarts-types de chaque côté de ce biais, ne dépassent pas la plus petite différence clinique significative ou les limites de la répétabilité intra-test, la nouvelle méthode peut remplacer l'ancienne.
- 4 Il est possible d'établir un intervalle de confiance pour le biais selon la formule ( $t$  = valeur selon la table de distribution de  $t$ ,  $n$  = nombre de patients,  $s$  = écart-type) :

$$\text{IC 95\%} = \text{biais} \pm (t_{(n-1, 95\%)} \times \text{erreur-type})$$

$$\text{Erreur type} = \sqrt{s^2/n}$$

- 5 Il est aussi possible d'établir un intervalle de confiance pour les limites de concordance selon la formule :

$$\text{IC 95\%} = \text{limite de concordance} \pm (t_{(n-1, 95\%)} \times \text{erreur-type})$$

$$\text{Erreur type} = \sqrt{3s^2/n}$$

On a établi des limites de répétabilité intra-test lors d'une étude sur la douleur en période postopératoire; les auteurs ont mesuré l'intensité de la douleur à l'aide d'une EVA de 100 mm à deux reprises, à trois minutes d'intervalle, et ont démontré des limites de répétabilité pour l'EVA de  $\pm 20$  mm (Deloach, Higgins et al. 1998). Bijur a aussi étudié la répétabilité intra-test de l'EVA papier lors de douleurs aiguës à l'urgence (Bijur, Silver et al. 2001). Ils ont mesuré l'intensité de la douleur aiguë à une minute d'intervalle et obtenu des limites de répétabilité de  $\pm 16$  mm sur une échelle de 100 mm. Le choix des limites de répétabilité intra-test pour la validation d'une échelle n'est donc pas évident et ce sujet est traité dans une autre section (voir **Discussion**).

L'autre élément important pour établir la validité d'une méthode de mesure est la plus petite différence significative. Il existe une controverse la concernant dans les études sur la douleur. On retrouve selon les études, sur une échelle de 100 mm : 9, 10, 12, 13, 15, 18 et 30 mm (Kelly 1998; Kelly 2001) (Powell, Kelly et al. 2001) (Gallagher, Liebman et al. 2001) (Bulloch and Tenenbein 2002) (Hagg, Fritzell et al. 2003) (Forouzanfar, Weber et al. 2003) (Todd and Funk 1996) (Lee, Hobden et al. 2003). Dans la littérature on retrouve également des auteurs qui considèrent qu'une différence absolue n'est pas un reflet de la réalité et qu'un pourcentage de la douleur initiale serait plus valide pour détecter un effet significatif d'une intervention thérapeutique (Bird and Dickson 2001). De plus, ce pourcentage fait lui-même l'objet de controverse; on retrouve 30%,

33%, 35% et 50 %, et il semble exister un certain consensus pour la valeur de 33% (Farrar, Portenoy et al. 2000) (Farrar, Young et al. 2001) (Farrar, Berlin et al. 2003) (Cepeda, Africano et al. 2003) (Forouzanfar, Weber et al. 2003) (Jensen, Chen et al. 2003), mais un pourcentage n'est pas applicable à la méthode de Bland et Altman.

Puisque que toutes ces différences sont plus petites que les limites de répétabilité intra-test, il est logique d'utiliser cette dernière comme limite de concordance acceptable lors d'étude clinique. Appliquer des limites plus sévères que celles obtenues lorsque l'on répète l'étalon de référence ne serait pas adéquat.

### 1.11 Objectif et hypothèse

La douleur est cette sensation intense et désagréable associée à une blessure ou à une pathologie qui, dans certains cas, peut menacer l'intégrité physique et émotionnelle du patient (Merskey 1994). En médecine d'urgence, les cliniciens évaluent habituellement l'intensité de la douleur à l'aide d'une échelle numérique verbale (EVN) de 11 points (0 à 10), et rarement avec une échelle visuelle analogique plastifiée à glissière (EVAp) qui devrait être un outil plus facile à utiliser et valide. Par ailleurs, la littérature sur la recherche clinique reconnaît que l'échelle visuelle analogique standard de 10 cm sur papier (EVA) est l'étalon de référence.

L'objectif de notre étude est de valider deux méthodes d'évaluation de l'intensité de la douleur aiguë en médecine d'urgence :

- Une nouvelle échelle visuelle analogique plastifiée à glissière
- Une échelle numérique verbale à 11 niveaux (0-10)

L'étalon de référence pour cette validation est l'échelle visuelle analogique de 100 mm sur papier.

Ce qui justifie cette étude est le fait qu'à ce jour, l'usage des EVN, des échelles catégorielles, de l'EVA pour évaluer l'intensité de la douleur aiguë à l'urgence et en recherche n'a pas été validée de façon adéquate. Pour ce faire; nous allons comparer les estimations de la douleur avec une EVN (populaire en médecine d'urgence) et une EVA plastifiée chez les patients souffrant de douleur aiguë en médecine d'urgence, en utilisant la méthode de Bland et Altman (Bland and Altman 1986). Ces derniers recommandent d'estimer le biais et les limites de la concordance. Et puisque certains auteurs considèrent le coefficient de corrélation intraclass acceptable comme méthode de validation, malgré ces lacunes, nous allons également la calculer afin de permettre de comparer ces différentes méthodes.

**Hypothèse :**

Lorsque comparée avec l'EVA papier comme étalon de référence :

- Le biais de l'EVA et de l'EVN est constant
- Les limites de concordances sont moindres que les limites de la répétabilité pour l'EVAp et EVN

## 2 ARTICLE

**Title:** Estimation of pain intensity in emergency medicine: a validation study

**Authors:** Raoul Daoust MD  
Associate professor  
Faculty of Medicine  
Université de Montréal  
Department of emergency medicine  
Research center  
Sacré-Coeur Hospital of Montreal

Pierre Beaulieu MD PhD  
Associate professor  
Faculty of Medicine  
Department of pharmacology  
Université de Montréal  
Research center  
CHUM, Hotel Dieu Hospital of Montreal

Christiane Manzini, research assistant  
Research center – Department of surgery  
Sacré-Coeur Hospital of Montreal

Gilles Lavigne DMD PhD  
Faculties of Dental Medicine & Medicine  
Université de Montréal  
Research center – Department of surgery  
Sacré-Coeur Hospital of Montreal

**Correspondence:** Raoul Daoust MD  
Sacré-Coeur Hospital of Montreal  
Emergency Department  
5400 Gouin Ouest  
Montreal  
Quebec, Canada  
H4J 1B3  
Telephone: 514-338-2222 #3234  
Facsimile number: 514-338-3513  
E-Mail : [REDACTED]

**Summary:** (keys words: pain measure, validation, numerical scale, visual analog scale, correlation)

**Objectives:** To estimate and to compare the validity of a verbal numerical scale (VNS) of 11 points (0 – 10), a plastified visual analogue scale (VASp) of 100units and a paper visual analogue scale (VAS) of 100mm as a gold standard, and to recommend the best method of reporting the intensity of acute pain in an emergency department (ED).

**Methods:** A convenience sample of 1,176 patients with acute pain (< 24 h) was recruited in the ED of a teaching hospital. Patients > 18 years and able to use the different scales were eligible for inclusion. Scales were presented randomly. Results were converted to a scale of 0 –100units and validity was quantified using intraclass correlation (ICC) and the Bland-Altman method. The limits of acceptability were previously set for the limits of agreement at  $\pm 20$ units, with a constant bias.

**Results:** The mean intensity of pain was 62units (3-100). The ICC was excellent between the VNS&VAS (0.88) and the VASp&VAS (0.92). The Bland-Altman method revealed a small bias of  $-4$  for the VNS and  $+1$  for VASp. However, the bias of the VNS varied with the intensity of pain from  $-10$  to  $+1$ units. The limits of agreement between the VNS&VAS and the VASp&VAS were  $-25;+17$ units and  $-17;+18$ units respectively.

**Conclusion:** In the ED, VASp has an excellent intraclass correlation, small bias and acceptable limits of agreement. However the VNS is less valid in that context because of its variable bias (mainly in lower scores) and wide limits of agreement.

## **Introduction**

Pain intensity assessment plays a key role in the emergency department (ED). It modulates the triage priority (Murray, Bullard et al. 2004), the prescription of analgesics, the incidence of pain relief and patient's perception of general quality of care (Kovach, Weissman et al. 1999) (Thomas and Andruszkiewicz 2004). The use of valid measures of pain intensity is also important in analgesic efficacy studies since intensity is described as a better primary outcome than pain relief because of its poor sensibility and specificity (Mamie, Morabia et al. 2000). Indeed, certain patients even express relief in the presence of more intense pain (Feine, Lavigne et al. 1998).

Emergency physicians usually evaluate the intensity of pain using a verbal numerical scale (VNS) like the 0 to 10 scale. This may be explained by the facility and rapidity of pain assessments although few studies support the VNS validity (Choiniere and Amsel 1996) (Carpenter and Brockopp 1995) (Deloach, Higgins et al. 1998) (Berthier, Potel et al. 1998) (Bijur, Latimer et al. 2003). The visual analogue scale (VAS), considered the gold standard in pain research, is rarely used in the ED (Turk and Melzack 2001). Other scales such as facial expression, behavioural, verbal descriptors, etc. are also seldom used in the ED, with the exception of facial expressions in a paediatric setting (Bulloch and Tenenbein 2002). To our knowledge there is a paucity of information regarding the validity of these tools in an ED.

To address these problems, some authors have made a case for the use of plastified mechanical visual analogue scale (VASp) and have mainly tried to validate them using a correlation coefficient. But this coefficient has limitations, for example, it is influenced by the range of results in the population studied, e.g. : if it is narrow it will have a tendency to be smaller and if it is wide it will be larger (Bland and Altman 1986) (Porter 1999). Bland-Altman proposed a method that measures the bias inherent to the

scaling issues. They defined it as the average of the differences estimated by the subtraction of the values obtained with the technique under evaluation from the ones obtained from the gold standard. Two standard deviations on each side of this bias are defined as the limits of agreement. If these limits are lower than the smallest clinically significant difference or than the limits of repeatability (limits of agreement when we retest with the gold standard) then the technique under evaluation is acceptable. The intra-class coefficient of correlation (ICC) is another method described as more valid than a simple coefficient of correlation (Bland and Altman 1996). However, the Bland-Altman method seems more intuitive and informative than the ICC.

In the present study, we design a patient friendly user plastified mechanical visual analogue scale (VASp; Fig1) with good graphic representation as described by Scott and Huskisson (Scott and E.C. 1976). We hypothesized that the VASp and the VNS are equally valid (null hypothesis) for measuring the intensity of acute pain in the ED.

## **Method**

### **Study design**

This prospective validation study was designed to compare the validity of the VASp and the 11 points VNS (usually employed in our ED) with the Bland-Altman method (Bland and Altman 1986) and the ICC, using a paper visual analogue scale (VAS) as a gold standard. The study received approval from our institutional review ethic board and informed consent was obtained from each participant.

## **Study setting and population**

This study was conducted in an adult tertiary care academic hospital with an affiliated emergency medicine residency program with an annual census of approximately 65,000 patients. We planned a convenience sample of 1,000 adults patients ( $\geq 18$  years) with acute pain (less than 24 h) presenting in our ED from August 2005 to January 2006. We also planned to monitor, during the study period, the characteristics of all the eligible patients (approximately 10,000 patients). Exclusion criteria were: alteration of the level of consciousness, inability to understand or use the different scales and the possibility that participation in the study could delay urgent treatment.

## **Study protocol**

Each attending physician was supplied with a copy of our VASp before the beginning of the study. Following a short training session from one of the authors on the study's protocol and the techniques used to evaluate the intensity of pain, medical students or attending physicians recruited participants in the ED over a 24 hours/day and seven days a week schedule. After obtaining informed consent, patient demographic information and pain location were collected (Table 1). Then, the 11 points VNS (0-10), the VASp and the 100 mm VAS were presented and explained to the patient in a predetermined random order. The words to describe the extremes of the VNS and VAS scales were: "no pain" for the lower end and "worst pain imaginable" for the higher end. In French the words used were "aucune douleur" and "pire douleur imaginable". The words written on the VASp were slightly different (but words used for the verbal presentation of the VASp were the same as above) so it would occupy all the space on the VASp to limit anchorage point as describe by Scott and Huskisson (Scott and E.C. 1976). On the higher end of the VASp, as seen on Fig 1, it is written: "The most

intense pain imaginable” in English and “Douleur la plus intense imaginable” in French. The patient completed the 100mm VAS directly on the questionnaire without having access to the written results from the other scales. If the patient had multiple sites of pain, he was asked to rate pain globally. If the patient could not decide between two numbers on the VNS, a half point was added to the lower number.

After data collection, researchers reviewed the questionnaires for completeness. Missing demographic and pain location information was obtained from the charts. Forty two questionnaires missing the VAS data, which was our gold standard, were excluded. The VAS was measured by one of the researchers without knowledge of the result from the other scales.

### **Data analysis**

Results were transferred to a Microsoft Access 2003 database, 15% of randomly selected files were reviewed for accuracy, and statistical analyses were performed using SPSS 13.0. The VNS was multiplied by 10 to convert results to a numerical scale of 0 to 100 units. The VAS and VASp were also converted to a 0 to 100 unit scale. Bias and limit of agreement were calculated using the Bland-Altman method. The bias represents the result with the gold standard (VAS) minus the result from the new technique (VNS or VASp).

The limits of acceptability selected for the limits of agreement were  $\pm 20$  units (limits of repeatability) on a 100 units scale derived from a study made with a VAS in post operative pain (Deloach, Higgins et al. 1998). Using the smallest clinical significant difference of intensity of pain as our limit of agreement would have been too severe because the reported differences are smaller than the limits of repeatability (Kelly 1998; Kelly 2001) (Powell, Kelly et al.

2001) (Gallagher, Liebman et al. 2001) (Bulloch and Tenenbein 2002) (Hagg, Fritzell et al. 2003) (Forouzanfar, Weber et al. 2003) (Todd and Funk 1996) (Lee, Hobden et al. 2003). We also determined that a variable bias was a reason to reject the scale (Bland and Altman 1986). To quantify the variability of the bias, the range of intensity was transformed in five equal categories and the bias calculated for each of categories, the statistical level of significance was determined by an ANOVA (limit of significance  $P < 0.05$ ). The ICC and its confidence interval was also calculated : we considered an ICC of less than 0.4 as a poor reproducibility, from 0.4 to 0.75 as good and 0.75 or more as a excellent reproducibility (Rosner 2000).

## **Results**

During the study period 10,231 patients were evaluated for pain intensity by the triage nurse (every patient is evaluated except if he needs resuscitation) and we enrolled 1,176 patients. The sample's characteristics (e.g.: age, language, gender and intensity of pain) were similar to eligible patients (Table 1). The enrollment was mostly limited by the flow of patients and sometimes by the incapacity of patients to use the different scales or the possibility that treatment might be delayed. The average pain in this sample was 62 units over a range of 3 to 100 units. The abdomen was the most frequent site of pain (31%), followed by extremities (22.4%), thorax (15.3%), spine (13.6%) and head-neck (13.6%). The ICC, bias and the limits of agreement are presented in Table 2. For both scales, the ICC was in the excellent range ( $> 0.75$  as described above) and their bias was small, but the limits of agreement were outside the limits of acceptability (limits of test retest) for the VNS. The bias in each of five categories along the range of intensity of pain is presented in Table 3 for the VASp and Table 4 for the VNS. The VNS bias is significantly variable along the range of the intensity of pain. These results were not affected by the patient's language, sex, age, site of pain, order of presentation or who explained the scales (data not shown).

## Discussion

In this study, VASp and the VNS both demonstrated excellent ICC and a small level of bias. Although the limits of agreement were within the repeatability limits for the VASp they were slightly outside those limits for the VNS. If we add the 95% confidence interval, the limits of agreement for the VNS would be -26 and +18. Such findings can be significant in the clinical and research setting, involving triage or decision of prescribing a pain medication for a patient in the ED. As an example, a result of 40 units on the VNS scale could represent a VAS pain intensity ranging from 14 to 58 units. We also demonstrated with the VNS a clinically and statistically significant variation in the bias along the range of intensity of pain. The data of another study also supports our findings of a variable bias; they compared different scales using average pain in five categories of pain intensity, and found that average pain was overestimated by the VNS (compared to a VAS) for light pain and that it was very similar for intense pain (Choiniere and Amsel 1996). Although VNS is the scale recommended by the IMMPACT group (Dworkin, Turk et al. 2005), studies designed to assess the analgesic effect of an intervention using a VNS scale may be subject to such a methodological bias. For example, if we use 12 units as a significant difference between treatment (Kelly 2001), a change in the average patient pain intensity from 40 to 30 units on a VNS would not have been significant. However, if we add the variable bias to such values, data will then be 35 and 20 units and the same study would have been considered as positive. Other authors suggest that a percentage of the initial pain be used as the significant difference (Farrar, Berlin et al. 2003), a variable bias should have a similar impact in this situation.

To our knowledge there is no other published study comparing the validity of a VASp or VNS for acute pain using the Bland-Altman method and a VAS on paper as a gold standard. Other studies on the subject use correlation, linear regression or compare average pain (Choiniere and Amsel 1996) (Carpenter and Brockopp 1995) (Breivik, Bjornsson et al. 2000) (Williams, Davies et al. 2000) (Bijur, Latimer et al. 2003) (van Laerhoven, van der Zaag-Loonen et al. 2004). Correlation has many limits, as described by Bland and Altman, the most important one being the impact of the range of pain intensity in the studied population on the coefficient of correlation (Bland and Altman 1986). Bijur and colleagues have suggested that a VNS can be substituted to a VAS for measuring acute pain in the ED (Bijur, Latimer et al. 2003). However, simple linear regression and correlation coefficient were used, both of which are very similar and have the same limitations (Porter 1999) (Bland and Altman 1986). Comparing average pain intensity to validate a new method also has limits. For example, if the estimate of average pain with the gold standard is 50mm, results ranging from 48 to 52mm (on a 100mm scale), the new method can have very different results (ex: 20 to 80mm), but if these results vary equally on each side of 50mm the average pain might be also 50mm. Data from the Bland And Altman study support this; the average peak flow from the two methods studied were very similar (450 and 452 L/min) but they demonstrated that the two techniques were not equivalent (Bland and Altman 1986). Most studies had other limitations: they were done during week days, excluding the types of patients and pain that usually present themselves during the evenings and nights, (ex: myocardial infarction); the number of patients also seem small (usually less than 100 patients), when the large variability of repeated measures is taken into account (Deloach, Higgins et al. 1998); the scales were often administered by one or two researchers, experts in their fields; the result might be different in daily use by an emergency physician.

One of our study's main limitations is the use of a convenience sample. We may have had a selection bias: selecting more cooperative patients, not reflecting the full range of pain, some patients expressing very intense or very light pain might not have been approached for the study and we do not know the proportion of patients that could not use the VASp. The fact that we recruited the full range of intensity of pain, around the clock and seven days a week, and that the scale was administered by a wide range of emergency physicians and medical students probably mitigates these limitations and adds to the external validity of the study. Another limitation is the fact that words on the VASp were slightly different from the words used to describe it. However this should make the limits of agreement larger and the ICC smaller, so globally add to the significance of our results. We also do not know the proportion of patients that were not be able to use the VASp, but a failure rate of 15% has been reported elsewhere (Berthier, Potel et al. 1998).

We did not use the limits of repeatability from the Bijur study; they repeated the measure of intensity of pain one minute apart, on 96 patients with acute pain in the ED (Bijur, Silver et al. 2001). The limits of agreement they obtained were  $\pm 16$ mm on a 100mm scale, but these results were achieved repeating, at 30 minutes interval, the paired VAS (done one minute apart) five times. Repeating paired VAS on the same patients seems to reduce the limits of agreement as shown in the DeLoach study. They repeated the paired pain assessments (done three minutes apart) up to three times and the limits of agreement were 17.6 and 23 for the two first set and only 13.5mm for the third (DeLoach, Higgins et al. 1998). This might be explained by a learning effect or that the patients with more difficulty using the VAS may do the first evaluation and not the subsequent, producing smaller limits of agreement. In the Bijur study, repeating the paired measures five times might exaggerate this effect, only 72 of the 96 patients did all the five paired VAS and another possible confounding factor is the rapid variation in acute pain during the study (two hours).

The findings of the present study also suggest that tools used in pain assessment for geriatric and paediatric population needs to be revisited with the Bland and Altman method. In fact, a common myth is that pain is less frequent in the geriatric population, but studies have demonstrated a prevalence of up to 43% of pain in a geriatric ward (Cornu, Chah-Wakilian et al. 1997). Most health providers think that patients with cognitive limitations cannot use a tool to measure the intensity of pain. However it was demonstrated that up to 65% of patients with light to moderate cognitive limitations can use a tool to measure the intensity of pain (Chibnall and Tait 2001). In our study we recruited patients aged up to 93 years old and we had no age effect (data not shown). Interestingly, it is also reported that older patients usually prefer verbal descriptive scales (Jones, Fink et al. 2005), although they have more ease using a non-verbal tool (Alexander, Plank et al. 2005) like the proposed VASp. The situation is similar in paediatric patients. Approximately one third of patients have the intensity of their pain evaluated with a scale (Colwell, Clark et al. 1996), only 50% of suffering children receive analgesics (McCarthy, Hewitt et al. 2000) and 80% of children receive inadequate analgesia (Ferrell and Rhiner 1991). As in the case of adults, pain documentation is associated with better use of analgesia (Drendel, Brousseau et al. 2006) Paediatric health care providers often use VNS, observational rating, or pictograms rating scale (facial expression) and think that a VASp is not appropriate or feasible. A VASp similar to ours has been validated in children 5-6 years and older (Hain 1997) (O'Rourke 2004). A color visual analog scale has also been studied and validated in children 4-5 years and older (Hain 1997). Although the validation techniques were not optimal, they were used successfully in children.

In conclusion, although the VNS is practical for measuring the intensity of acute pain in the ED, it's has limitations. It's variable bias and wide limits of agreement question it use for assessing pain in the clinical and research setting. However, the bias is less variable and the limits of agreement are only slightly outside the acceptable limits in the 40 to 100 units range, so its use for the follow up of clinical pain in this range could be acceptable. The VASp seem to be a practical and more valid alternative to VNS to quantify pain intensity in clinical and research settings, although it warrants further studies in geriatric and paediatric populations.

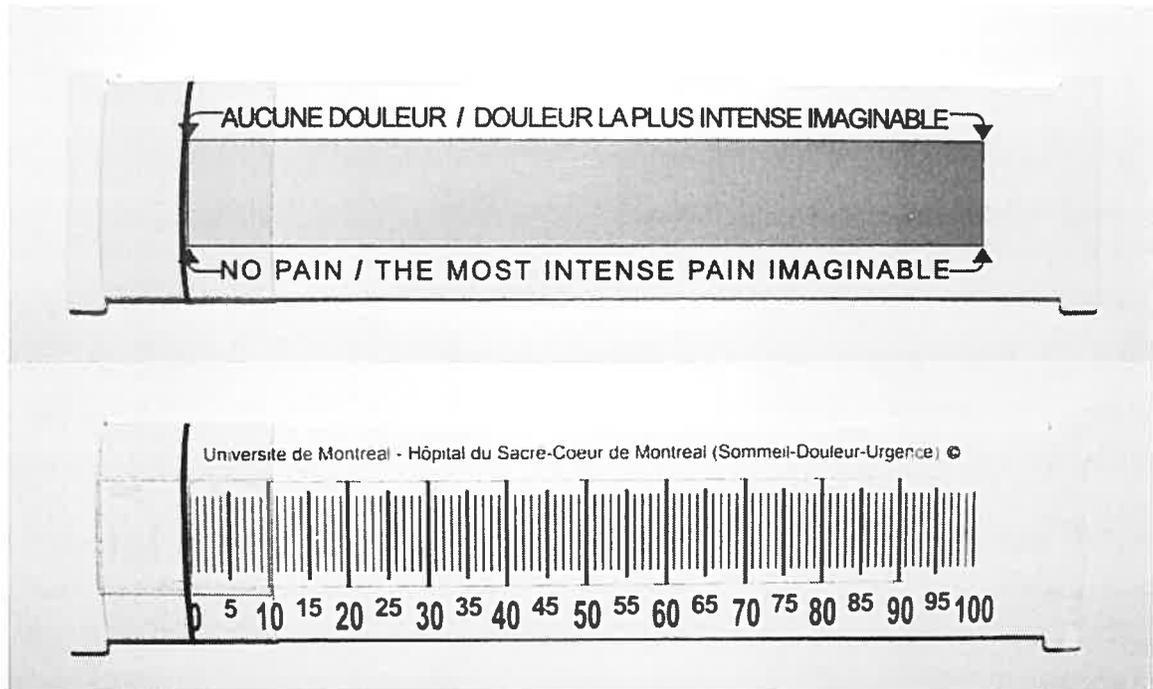


Figure 1

**Horizontal plastified mechanical visual analogue scale**

The coloured side, without the numbers, was presented to the patients and were instructed to slide the mechanical marker to the level of the perceived pain intensity. Then the attending physician or the medical student reads the number on the reverse. Note that to assure reliability of results the marker's line is just slightly smaller than the space between two numbers.

Table 1

**Characteristics of enrolled and of eligible patients during the study period.**

<b>Variable</b>	<b>Results 1,176 enrolled patients</b>	<b>Results 10,231 eligible patients</b>
<b>Gender</b>		
Female	54.3%	54.2%
Male	45.7%	45.8%
<b>Language</b>		
French	83.6%	
English	12.8%	
Other	2.3%	
<b>Age</b>		
Range	18-93 yr	18-98 yr
Average	48 yr	49 yr
Median	47 yr	48 yr
<b>Average pain intensity</b>		
	62 units	65 units
<b>Percentage of patients in each category of pain intensity (VNS converted to a 100 units scale)</b>		
0-19.9	4.3%	4.1%
20-39.9	17.3%	13.2%
40-59.9	23.8%	26.9%
60-79.9	34.4%	39.5%
80- 100	20.2%	16.2%

Table 2

**Global bias, limits of agreement and intra-class coefficient of correlation.**

	<b>Bias (0 – 100 units)</b>	<b>Limits of agreement (0 – 100 units)</b>	<b>ICC</b>
<b>VASp</b>	+ 1	-16.5 ; +18.3	0.92
<b>CI 95%</b>	± 0.5	± 0.9	0.92-0.93
<b>VNS</b>	- 4	- 24.7 ; +16.9	0.88
<b>CI 95%</b>	± 0.6	± 1.1	0.84-0.91

Table 3

**Variation of the bias along the range of intensity of pain for the VASp**

<b>Category of intensity of pain (scale on a 100 units)</b>	<b>Number of patients</b>	<b>Bias VAS – VASp (scale on a 100 units)</b>
0-19.9	51 (4.3%)	- 1.8
20-39.9	203 (17.3%)	- 3.4
40-59.9	280 (23.8%)	+1
60-79.9	404 (34.4%)	+ 2.6
80- 100	238 (20.2%)	+ 2.4

Table 4

**Variation of the bias along the range of intensity of pain for the VNS**

<b>Category of intensity of pain (scale on a 100 units)</b>	<b>Number of patients</b>	<b>Bias* VAS – VNS X 10** (scale on a 100 units)</b>
0-19.9	51 (4.3%)	- 10.2
20-39.9	203 (17.3%)	- 8.1
40-59.9	280 (23.8%)	- 5.0
60-79.9	404 (34.4%)	- 3.0
80- 100	238 (20.2%)	+ 1.0

\* ANOVA  $p < 0.0001$

\*\*Conversion to a numeral scale 0 to 100.

### 3 DISCUSSION ET CONCLUSION

L'importance de la mesure de l'intensité de la douleur aiguë à l'urgence est indéniable. Il est donc essentiel d'utiliser un outil de mesure fiable et validé. Il est reconnu que la méthode de Bland et Altman est la plus fiable pour valider un outil de mesure de l'intensité de la douleur (Bland and Altman 1986) et que l'EVA sur papier est l'étalon de référence idéal (Scott and E.C. 1976). A notre connaissance, notre étude est la seule qui utilise à la fois la méthode de Bland et Altman pour valider un outil de mesure de l'intensité de la douleur et l'EVA sur papier comme étalon de référence.

De nombreuses études ont utilisé le coefficient de corrélation simple avec les limites que nous avons déjà discutées (Bijur, Latimer et al. 2003) (Breivik, Bjornsson et al. 2000) (Carpenter and Brockopp 1995) (Choiniere and Amsel 1996) (Deloach, Higgins et al. 1998) (Grossman, Sheidler et al. 1992) (van Laerhoven, van der Zaag-Loonen et al. 2004). Le coefficient de corrélation intraclasse, qui compense pour certaines les limites du coefficient de corrélation simple, a été utilisé par Hagino avec un nombre restreint de patients (36) afin de comparer une EVA plastifié à une EVA papier et il a obtenu un excellent coefficient de 0,86 (Hagino, Thompson et al. 1996). Mais le coefficient de corrélation intraclasse n'a pas tous les avantages de la méthode d'Altman et Bland, et le résultat demeure difficile à interpréter en clinique. Choinière et Carpenter ont comparé des moyennes, avec les limites

déjà décrites (Carpenter and Brockopp 1995) (Choiniere and Amsel 1996). Étant donné la grande variabilité des réponses lors d'études de répétabilité d'une même échelle (Deloach, Higgins et al. 1998) (Berthier, Potel et al. 1998), toutes ces études semblent avoir utilisé un nombre insuffisant de patients (habituellement moins de 100), si l'on désire des intervalles de confiances restreints.

Notre étude a démontré un excellent coefficient de corrélation intraclasse et un faible biais pour l'EVAp et l'ENV. Les limites de concordances de l'EVAp sont larges, mais moindres que celle de la répétabilité de l'EVA papier (notre étalon de référence) que nous avons établi à priori comme nos limites d'acceptabilités (Deloach, Higgins et al. 1998). Pour l'ENV si on ajoute l'intervalle de confiance de 95 % à ces limites on obtient des limites de concordances de -26 et +18. L'impact de ces larges limites de concordances est important : par exemple, un résultat de 40 unités sur une ENV pourrait représenter de 14 à 58 unités sur une EVA papier. L'étendue possible de ces résultats sur une EVA est certainement cliniquement significatif, et il aurait aussi un impact en recherche puisque plusieurs retiennent de 9 à 30 unités comme la différence cliniquement significative (Powell, Kelly et al. 2001) (Lee, Hobden et al. 2003).

Nous avons également démontré la présence d'un biais variable lors de l'utilisation de l'ENV qui est cliniquement et statistiquement significatif. Ce biais variable pourrait avoir aussi des impacts importants. Théoriquement, ceci est suffisant pour rejeter une nouvelle technique, et par exemple : une étude portant sur un nouvel analgésique qui utiliserait une limite de 12 unités comme différence minimale significative (Kelly 2001) qui aurait une moyenne de douleur qui diminuerait de 40 à 30 unités serait négative. Mais en ajoutant le biais variable à cette étude, on obtient alors 35 unités et 21.9 unités, ce qui pourrait ainsi rendre l'étude positive. L'étude de Choinière et collaborateurs supporte aussi la présence d'un biais variable; ils ont comparé les moyennes d'intensité de douleur de l'EVA selon différentes catégories d'intensité (faible, agaçante, irritante, cruelle et atroce). Les données de cette étude suggèrent que l'ENV, lorsque comparée à une EVA, surestime l'intensité de la douleur lorsqu'elle est classée « faible » ou « agaçante » et donne des résultats similaires lorsque classée « atroce » (Choiniere and Amsel 1996).

La limite principale de notre étude est l'utilisation d'un échantillon de convenance; ceci nous amène possiblement à un biais de sélection. Nous avons peut-être instinctivement évité les patients qui semblaient très ou peu souffrants, et nous avons possiblement recruté certains patients plus coopératifs. Également, nous ne possédons pas de données sur la proportion de patients se présentant à l'urgence qui sont incapables d'utiliser

une EVAp. Une autre limite de notre étude est que les mots inscrits sur l'EVAp étaient légèrement différents de ceux utilisés pour décrire verbalement les différents types d'échelles. Cette limite devrait augmenter les limites de concordances et diminuer le coefficient de corrélation intraclasse; donc avec l'utilisation des mêmes mots les résultats de notre étude seraient possiblement encore plus significatifs. Toutefois, lors de l'utilisation sans information verbale de l'EVA, il est possible que les résultats soient légèrement différents, mais en clinique et en recherche cette situation est rare.

Nous avons utilisé les limites d'acceptabilité de l'étude de DeLoach, mais Bijur a aussi étudié la répétabilité de l'EVA papier lors de douleurs aiguës à l'urgence (DeLoach, Higgins et al. 1998) (Bijur, Silver et al. 2001). Ils ont mesuré l'intensité de la douleur aiguë à une minute d'intervalle chez 96 patients, et obtenu des limites de répétabilité de  $\pm 16$  mm sur une échelle de 100 mm. Mais, pour générer ces résultats, ils ont refait les paires de mesures à cinq reprises toutes les 30 minutes. Dans l'étude de DeLoach, ils ont mesuré la douleur aiguë en période postopératoire à trois minutes d'intervalle, mais à trois reprises toutes les 20 minutes. Ils nous présentent les résultats pour chaque séance de répétitions (informations non disponibles pour l'étude de Bijur) et lors de la troisième séance les limites de répétabilité sont de 13,5 mm sur une échelle de 100 mm (17.6 pour la

première et 23 pour la deuxième). Il semble donc que le fait de répéter la mesure diminue les limites de répétabilité et puisque l'étude de Bijur l'a fait cinq fois ceci pourrait expliquer pourquoi leurs résultats sont moindres que ceux de Deloach. De plus, seulement 72 patients ont complété les cinq répétitions, il est possible que ceux qui ont rencontré le plus de difficulté avec l'utilisation de l'EVA papier se soient désistés et qu'ils aient ainsi contribué à la diminution de leur limite de répétabilité. Pour ces raisons, nous avons conservé les limites de répétabilité de Deloach comme limite d'acceptabilité.

Malgré la croyance générale alléguant que les patients âgés sont incapables d'utiliser un outil de mesure pour la douleur, Chibnall a démontré que 65 % des patients porteurs d'un trouble cognitif léger à modéré utilisent de façon adéquate différents outils de mesure (Chibnall and Tait 2001) et notre étude à incorporer des patients jusqu'à 93 ans. Malgré la préférence des patients âgés pour les échelles descriptives verbales, ceux-ci ont néanmoins démontré plus de facilité à utiliser un outil non verbal comme une VASp (Alexander, Plank et al. 2005). La fréquence d'événement douloureux est aussi élevée en gériatrie (Cornu, Chah-Wakilian et al. 1997). La situation est similaire en pédiatrie : seulement le tiers des enfants ont une évaluation de l'intensité de leur douleur à l'aide d'une échelle (Colwell, Clark et al. 1996), et environ 50 % des enfants souffrants vont recevoir un analgésique (McCarthy, Hewitt et al. 2000) et cette analgésie sera inadéquate chez 80 % (Ferrell and Rhiner 1991).

Comme chez adulte, le simple fait de mesurer l'intensité de la douleur est associé à une meilleure utilisation de l'analgésie, il est donc important de mesurer l'intensité de la douleur en pédiatrie (Drendel, Brousseau et al. 2006). Les professionnels de la santé utilisent souvent une ENV, une échelle observationnelle ou de pictogramme (visage avec expression de douleur). Malgré le fait que l'on croit fréquemment que l'utilisation d'une EVA est impossible chez l'enfant, des EVA plastifiées ont été étudiées chez des enfants de 4 à 6 ans (O'Rourke 2004) (Hain 1997). Bien que la technique de validation n'était pas idéale, cette étude a mis en évidence qu'une EVA est utilisable chez l'enfant et qu'une étude en pédiatrie est concevable.

Selon Lacroix et Gauvin, afin de déterminer la qualité d'un test on doit vérifier : *la justification scientifique, la validité de la méthode, la facilité d'utilisation, que le test est opérationnel, qu'il est conforme, l'exhaustivité, la reproductibilité, l'exactitude, que l'interprétation est facile et la validité externe* (Lacroix and Gauvin 2005). Notre revue de la littérature a démontré des bases scientifiques solides pour justifier l'utilisation de notre EVA. Notre EVAp utilise une méthode valide qui respecte les caractéristiques techniques d'une bonne EVA selon Scott et Huskisson. Nous n'avons pas évalué de façon formelle la facilité d'utilisation de notre EVAp, mais lors d'un sondage informel auprès des patients et enseignants qui ont utilisé cette échelle les commentaires suivants furent émis; elle est simple à utiliser et facile à comprendre pour les patients, la lecture des résultats par les médecins est

claire et sans ambiguïtés. Étant donné le grand nombre de patients, additionnée au fait que le recrutement était effectué par une grande variété de médecins et étudiants en médecine sur 24 heures et cela 7 jours par semaine, il est probable que l'EVAp donnerait les mêmes résultats dans différents contextes de médecine d'urgence. L'EVAp est donc opérationnelle et elle présente probablement une excellente validité externe. L'EVA est également conforme, car elle présente un excellent coefficient de corrélation intraclasse et de bonnes limites de concordance. L'exhaustivité est moins pertinente dans un contexte de médecine d'urgence et nous n'avons pas évalué la reproductibilité de notre EVA, mais il est probable qu'elle soit similaire aux autres EVA recensés dans la littérature (Berthier, Potel et al. 1998). Notre EVAp démontre un faible biais de +1, elle est donc juste et l'interprétation des résultats est simple. Notre EVAp semble avoir la majorité des caractéristiques d'un test de qualité.

L'ENV est facile d'utilisation pour mesurer la douleur aiguë à l'urgence. Son biais étant peu variable lorsque la douleur est intense (plus de 40 unités) et puisque ses limites de concordance sont seulement légèrement supérieures aux limites de répétabilité, il serait donc acceptable d'utiliser une ENV lorsque la douleur est entre 40 et 100 unités pour le suivi clinique. Mais, l'ENV ne devrait pas être utilisée en recherche; pour valider l'efficacité d'un analgésique ou d'une autre méthode de contrôle de la douleur et les résultats

d'études qui ont utilisé une ENV devraient être interprétés à la lumière ces limites.

L'EVA plastifiée nous semble une alternative pratique et valide qui devrait être utilisée en recherche et en clinique pour évaluer la douleur aiguë en médecine d'urgence. L'EVA plastifiée devrait faire l'objet d'études supplémentaire en pédiatrie et en gériatrie pour en vérifier le taux de complétion et sa validité ces contextes.

#### 4 BIBLIOGRAPHIE

- Alexander, B. J., P. Plank, et al. (2005). "Methods of pain assessment in residents of long-term care facilities: a pilot study." Journal of the American Medical Directors Association **6**(2): 137-43.
- Aubrun, F., O. Langeron, et al. (2003). "Relationships between measurement of pain using visual analog score and morphine requirements during postoperative intravenous morphine titration.[see comment]." Anesthesiology **98**(6): 1415-21.
- Averbuch, M. and M. Katzper (2004). "Assessment of visual analog versus categorical scale for measurement of osteoarthritis pain." Journal of Clinical Pharmacology **44**(4): 368-72.
- Babul, N., A. C. Darke, et al. (1993). "Using memory for pain in analgesic research." Annals of Pharmacotherapy **27**(1): 9-12.
- Beaulieu, P. (2005). "Pharmacologie de la douleur." Les presses de l'université de Montréal: 329.
- Berkley, K. J. (1997). "Sex differences in pain." Behaviour and Brain Sciences **20**(3).
- Berthier, F., G. Potel, et al. (1998). "Comparative study of methods of measuring acute pain intensity in an ED." The American Journal of Emergency Medicine **16**(2): 132-136.
- Beyer, J. E. and M. C. Bournaki (1989). "Assessment and management of postoperative pain in children." Pediatrician **16**(1-2): 30-8.
- Beyer, J. E., P. J. McGrath, et al. (1990). "Discordance between self-report and behavioral pain measures in children aged 3-7 years after surgery." Journal of Pain & Symptom Management **5**(6): 350-6.

- Bijur, P. E., C. T. Latimer, et al. (2003). "Validation of a Verbally administered Numerical Rating Scale of Acute Pain for Use in The Emergency Department." Academic Emergency Medicine **10**: 390-392.
- Bijur, P. E., W. Silver, et al. (2001). "Reliability of the Visual Analog Scale for Measurement of Acute Pain." Acad Emerg Med **8**(12): 1153-1157.
- Bird, S. B. and E. W. Dickson (2001). "Clinically significant changes in pain along the visual analog scale." Annals of Emergency Medicine **38**(6): 639-643.
- Bland, J. M. and D. G. Altman (1986). "Statistical methods for assessing agreement between two methods of clinical measurement." Lancet **i**: 307-310.
- Bland, J. M. and D. G. Altman (1996). "Statistic Notes: Measurement error and correlation coefficients." British Medical Journal **313**: 41-42.
- Bolognese, J. A., T. J. Schnitzer, et al. (2003). "Response relationship of VAS and Likert scales in osteoarthritis efficacy measurement." Osteoarthritis & Cartilage **11**(7): 499-507.
- Breivik, E. K., G. A. Bjornsson, et al. (2000). "A comparison of pain rating scales by sampling from clinical trial data.[see comment]." Clinical Journal of Pain **16**(1): 22-8.
- Briggs, M. and J. S. Closs (1999). "A descriptive study of the use of visual analogue scales and verbal rating scales for the assessment of postoperative pain in orthopedic patients." Journal of Pain & Symptom Management **18**(6): 438-46.
- Bulloch, B. and M. Tenenbein (2002). "Assessment of clinically significant changes in acute pain in children." Academic Emergency Medicine **9**(3): 199-202.
- Bulloch, B. and M. Tenenbein (2002). "Validation of 2 Pain Scales for Use in the Pediatric Emergency Department." Pediatrics **110**(3): e33-.

- Calam, R. M., P. Jimmieson, et al. (2000). "Can computer-based assessment help us understand children's pain?" European Journal of Anaesthesiology **17**(5): 284-8.
- Carpenter, J. S. and D. Brockopp (1995). "Comparison of patients' ratings and examination of nurses' responses to pain intensity rating scales." Cancer Nursing **18**(4): 292-298.
- Cepeda, M. S., J. M. Africano, et al. (2003). "What decline in pain intensity is meaningful to patients with acute pain?" Pain **105**(1-2): 151-157.
- Chapman, C. R. and P. J. Dunbar (1998). "Measurement in pain therapy: Is pain relief really the endpoint?" Current Opinion in Anaesthesiology **11**(5): 533-537.
- Chibnall, J. T. and R. C. Tait (2001). "Pain assessment in cognitively impaired and unimpaired older adults: a comparison of four scales." Pain **92**(1-2): 173-86.
- Chibnall, J. T., R. C. Tait, et al. (1997). "The effects of medical evidence and pain intensity on medical student judgments of chronic pain patients." Journal of Behavioral Medicine **20**(3): 257-71.
- Choiniere, M. and R. Amsel (1996). "A visual analogue thermometer for measuring pain intensity." Journal of Pain & Symptom Management **11**(5): 299-311.
- Choiniere, M., R. Melzack, et al. (1990). "Comparisons between patients' and nurses' assessment of pain and medication efficacy in severe burn injuries." Pain **40**(2): 143-152.
- Collins, S. L., R. A. Moore, et al. (1997). "The visual analogue pain intensity scale: what is moderate pain in millimetres?" Pain **72**(1-2): 95-7.
- Colwell, C., L. Clark, et al. (1996). "Postoperative use of pediatric pain scales: children's self-report versus nurse assessment of pain intensity and affect." Journal of Pediatric Nursing **11**(6): 375-82.

- Cornu, H. P., A. Chah-Wakilian, et al. (1997). "Emile Roux: One day pain came. Prevalence study in a gerontological hospital. [French]." Revue de Geriatrie **22**(3): 155-162.
- Coulthard, P., B. J. Pleuvry, et al. (2000). "Behavioural measurement of postoperative pain after oral surgery." British Journal of Oral & Maxillofacial Surgery **38**(2): 127-31.
- De Wit, R., F. van Dam, et al. (1999). "Empirical comparison of commonly used measures to evaluate pain treatment in cancer patients with chronic pain." Journal of Clinical Oncology **17**(4): 1280.
- Deloach, L. J., M. S. Higgins, et al. (1998). "The visual Analog Scale in the Immediate Postoperative Period : Intrasubject Variability and Correlation with a Numeric Scale." Anesthesia & Analgesia **86**: 102-6.
- Drendel, A. L., D. C. Brousseau, et al. (2006). "Pain assessment for pediatric patients in the emergency department." Pediatrics **117**: 1511-1518.
- Dworkin, R. H., D. C. Turk, et al. (2005). "Core outcome measures for chronic pain clinical trials: IMMPACT recommendations." Pain **113**(1-2): 9-19.
- Eich, E., J. L. Reeves, et al. (1985). "Memory for pain: relation between past and present pain intensity." Pain **23**(4): 375-80.
- Everts, B., B. Karlson, et al. (1999). "Pain recollection after chest pain of cardiac origin." Cardiology **92**(2): 115-20.
- Farrar, J. T., J. A. Berlin, et al. (2003). "Clinically Important Changes in Acute Pain Outcome Measures: A Validation Study." Journal of Pain and Symptom Management **25**(5): 406-411.
- Farrar, J. T., R. K. Portenoy, et al. (2000). "Defining the clinically important difference in pain outcome measures.[see comment]." Pain **88**(3): 287-94.

- Farrar, J. T., J. P. Young, Jr., et al. (2001). "Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale.[see comment]." Pain **94**(2): 149-58.
- Feine, J. S., M. C. Bushnell, et al. (1991). "Sex differences in the perception of noxious heat stimuli." Pain **44**(3): 255-262.
- Feine, J. S., G. J. Lavigne, et al. (1998). "Memories of chronic pain and perceptions of relief." Pain **77**(2): 137-41.
- Ferrell, B. R. and M. Rhiner (1991). "High-tech comfort : ethical issues in cancer pain management for the 1990s." The Journal of clinical ethics **2**: 108-115.
- Finley, G. A., C. T. Chambers, et al. (2003). "Construct validity of the parents' postoperative pain measure." Clinical Journal of Pain **19**(5): 329-34.
- Forouzanfar, T., W. E. J. Weber, et al. (2003). "What is a meaningful pain reduction in patients with complex regional pain syndrome type 1?" Clinical Journal of Pain **19**(5): 281-5.
- Fosnocht, D. E., J. M. Dahle, et al. (2003). "Measuring Pain Intensity in the Emergency Department: Correlation of a Numeric Rating Scale with a Visual Analog Scale." Acad Emerg Med **10**(5): 484-.
- Gaertner, J., F. Elsner, et al. (2004). "Electronic pain diary: a randomized crossover study.[erratum appears in J Pain Symptom Manage. 2004 Dec;28(6):626]." Journal of Pain & Symptom Management **28**(3): 259-67.
- Gallagher, E. J., M. Liebman, et al. (2001). "Prospective validation of clinically important changes in pain severity measured on a visual analog scale." Annals of Emergency Medicine **38**(6): 633-638.
- Gill, J. S., D. G. Beevers, et al. (1985). "Relation between initial blood pressure and its fall with treatment." The Lancet **325**(8428): 567-569.

- Gracely, R. H. and D. M. Kwilosz (1988). "The Descriptor Differential Scale: applying psychophysical principles to clinical pain assessment." Pain **35**(3): 279-288.
- Grossman, S. A., V. R. Sheidler, et al. (1992). "A comparison of the Hopkins Pain Rating Instrument with standard visual analogue and verbal descriptor scales in patients with cancer pain." Journal of Pain & Symptom Management **7**(4): 196-203.
- Hagg, O., P. Fritzell, et al. (2003). "The clinical importance of changes in outcome scores after treatment for chronic low back pain.[see comment]." European Spine Journal **12**(1): 12-20.
- Hagino, C., M. Thompson, et al. (1996). "Agreement between 2 pain visual analogue scales, by age and area fo complaint in neck and low back pain subjects : the standard pen and paper VAS versus plastic mechanical sliderule VAS." Journal of the Canadian Chiropractic Association **40**(4): 220.
- Hain, R. (1997). "Pain scales in children : A review." Palliative Medicine **11**: 341-350.
- Horbury, C., A. Henderson, et al. (2005). "Influences of patient behavior on clinical nurses' pain assessment: implications for continuing education." Journal of Continuing Education in Nursing **36**(1): 18-24.
- Hunter, M., L. McDowell, et al. (2000). "An evaluation of the Faces Pain Scale with young children." Journal of Pain & Symptom Management **20**(2): 122-9.
- Jamison, R. N., R. H. Gracely, et al. (2002). "Comparative study of electronic vs. paper VAS ratings: a randomized, crossover trial using healthy volunteers." Pain **99**(1-2): 341-7.
- Jamison, R. N., T. Sbrocco, et al. (1989). "The influence of physical and psychosocial factors on accuracy of memory for pain in chronic pain patients." Pain **37**(3): 289-94.

- Jenkinson, C., D. Carroll, et al. (1995). "Comparison of the sensitivity to change of long and short form pain measures." Quality of Life Research **4**(4): 353-7.
- Jensen, M. P., C. Chen, et al. (2003). "Interpretation of visual analog scale ratings and change scores: a reanalysis of two clinical trials of postoperative pain." Journal of Pain **4**(7): 407-14.
- Jensen, M. P. and P. Karoly (2001). "Handbook of pain assessment." Guilford Press: 15-34.
- Jensen, M. P., S. A. Martin, et al. (2005). "The meaning of pain relief in a clinical trial." Journal of Pain **6**(6): 400-6.
- Jones, K. R., R. Fink, et al. (2005). "Measuring pain intensity in nursing home residents." Journal of Pain & Symptom Management **30**(6): 519-27.
- Joyce, C. R. B., D. W. Zutshi, et al. (1975). "Comparison of fixed interval and visual analogue scales for rating chronic pain." European Journal of Clinical Pharmacology **8**: 415-420.
- Kelly, A. M. (1998). "Does the clinically significant difference in visual analog scale pain scores vary with gender, age, or cause of pain?" Acad Emerg Med **5**(11): 1086-1090.
- Kelly, A. M. (2001). "The minimum clinically significant difference in visual analogue scale pain score does not differ with severity of pain." Emerg Med J **18**(3): 205-207.
- Klopfenstein, C. E., F. R. Herrmann, et al. (2000). "Pain intensity and pain relief after surgery: A comparison between patients' reported assessments and nurses' and physicians' observations." Acta Anaesthesiologica Scandinavica **44**(1): 58-62.
- Knox, T. H. (2005). "Pain Assessment Instruments for Use in the Emergency Department." Emergency Medicine clinics of North America **23**: 285-295.

- Kovach, C. R., D. E. Weissman, et al. (1999). "Assessment and treatment of discomfort for people with late-stage dementia." Journal of Pain & Symptom Management **18**(6): 412-9.
- Lachapelle, D. L., T. Hadjistavropoulos, et al. (1999). "Pain measurement in persons with intellectual disabilities." Clinical Journal of Pain **15**(1): 13-23.
- Lacroix, J. and F. Gauvin (2005). "Validation et interprétation d'un test." Cours MSo 6027: introduction à l'épidémiologie clinique: 19-27.
- Lee, J. S., E. Hobden, et al. (2003). "Clinically important change in the visual analog scale after adequate pain control." Academic Emergency Medicine **10**(10): 1128-30.
- Lines, C. R., K. Vandormael, et al. (2001). "A comparison of visual analog scale and categorical ratings of headache pain in a randomized controlled clinical trial with migraine patients." Pain **93**(2): 185-90.
- Mamie, C., A. Morabia, et al. (2000). "Treatment efficacy is not an index of pain intensity.[see comment]." Canadian Journal of Anaesthesia **47**(12): 1166-70.
- McCarthy, C., S. Hewitt, et al. (2000). "Pain in young children attending an accident and emergency department." Journal of Accident & Emergency Medicine **17**(4): 265-267.
- McGrath, P. J. and A. M. Unruh (2006). "Measurement and assessment for paediatric pain." Textbook of pain: 305-15.
- Melzack, R. and J. Katz (2006). "Pain assessment adult patients." textbook of pain: 291-304.
- Merskey, H. (1994). "Logic, truth and language in concepts of pain." Quality of Life Research **3 Suppl 1**: S69-76.

- Morin, C., J. P. Lund, et al. (2000). "Differences between the sexes in post-surgical pain." Pain **85**(1-2): 79-85.
- Morrison, R. S., J. C. Ahronheim, et al. (1998). "Pain and discomfort associated with common hospital procedures and experiences." Journal of Pain & Symptom Management **15**(2): 91-101.
- Murray, M., M. Bullard, et al. (2004). "Revisions to the canadian emergency department triage and acuity scale implementation guidelines." Canadian journal of emergency medicine **6**(6): 421.
- Myles, P. S., S. Troedel, et al. (1999). "The pain visual analog scale: is it linear or nonlinear?[see comment]." Anesthesia & Analgesia **89**(6): 1517-20.
- Myles, P. S. and N. Urquhart (2005). "The linearity of the visual analogue scale in patients with severe acute pain.[see comment]." Anaesthesia & Intensive Care **33**(1): 54-8.
- O'Rourke, D. (2004). "The Measurement of Pain in Infants, Children, and Adolescents: From Policy to Practice." PHYS THER **84**(6): 560-570.
- Paul-Dauphin, A., F. Guillemin, et al. (1999). "Bias and precision in visual analogue scales: a randomized controlled trial." American Journal of Epidemiology **150**(10): 1117-27.
- Pautex, S., F. Herrmann, et al. (2005). "Feasibility and reliability of four pain self-assessment scales and correlation with an observational rating scale in hospitalized elderly demented patients." Journals of Gerontology Series A-Biological Sciences & Medical Sciences **60**(4): 524-9.
- Porter, A. M. (1999). "Misuse of correlation and regression in three medical journals." J R Soc Med **92**(3): 123-128.

- Powell, C. V., A.-M. Kelly, et al. (2001). "Determining the minimum clinically significant difference in visual analog pain score for children." Annals of Emergency Medicine **37**(1): 28-31.
- Price, D. D., F. M. Bush, et al. (1994). "A comparison of pain measurement characteristics of mechanical visual analogue and simple numerical rating scales." Pain **56**(2): 217-226.
- Price, D. D., P. A. McGrath, et al. (1983). "The validation of visual analogue scales as ratio scale measures for chronic and experimental pain." Pain **17**(1): 45-56.
- Probst, B., E. Lyons, et al. (2005). "Factors affecting emergency department assessment and management of pain in children." Pediatric Emergency Care **21**(5): 298-305.
- Romsing, J., J. Moller-Sonnergaard, et al. (1996). "Postoperative pain in children: comparison between ratings of children and nurses." Journal of Pain & Symptom Management **11**(1): 42-6.
- Rosner, B. (2000). "Fundamentals of biostatistic." (5e edition): 562-566.
- Rosseland, L. A. and A. Stubhaug (2004). "Gender is a confounding factor in pain trials: women report more pain than men after arthroscopic surgery." Pain **112**(3): 248-53.
- Rundshagen, I., K. Schnabel, et al. (1999). "Patients' vs nurses' assessments of postoperative pain and anxiety during patient- or nurse-controlled analgesia." British Journal of Anaesthesia **82**(3): 374-8.
- Salomaki, T. E., T. M. Hokajarvi, et al. (2000). "Improving the quality of postoperative pain relief." European Journal of Pain: Ejp **4**(4): 367-72.
- Scott, J. and H. E.C. (1976). "Graphic representation of pain." Pain **2**: 175-184.

- Silka, P. A., M. M. Roth, et al. (2004). "Pain scores improve analgesic administration patterns for trauma patients in the emergency department." Academic Emergency Medicine **11**(3): 264-70.
- Skovlund, E. and O. Flaten (1995). "Response measures in the acute treatment of migraine." Cephalalgia **15**(6): 519-22, 450-1.
- Smith, W. B., R. H. Gracely, et al. (1998). "The meaning of pain: Cancer patients' rating and recall of pain intensity and affect." Pain **78**(2): 123-129.
- Smith, W. B. and M. A. Safer (1993). "Effects of present pain level on recall of chronic pain and medication use." Pain **55**(3): 355-61.
- Stone, A. A., S. Shiffman, et al. (2003). "Patient compliance with paper and electronic diaries." Controlled Clinical Trials **24**(2): 182-99.
- Suraseranivongse, S., T. Montapaneewat, et al. (2005). "Cross-validation of a self-report scale for postoperative pain in school-aged children." Journal of the Medical Association of Thailand **88**(3): 412-8.
- Thomas, S. H. and L. M. Andruszkiewicz (2004). "Ongoing visual analog score display improves Emergency Department pain care." Journal of Emergency Medicine **26**(4): 389-94.
- Todd, K. H. and J. P. Funk (1996). "The minimum clinically important difference in physician-assigned visual analog pain scores." Acad Emerg Med **3**(2): 142-146.
- Turk, D. C., R. H. Dworkin, et al. (2003). "IMMPACT recommendations." Pain Management Nursing **106**: 337-45.
- Turk, D. C. and R. Melzack (2001). "Assessing pain intensity." Handbook of Pain Assessment: 23-24.

- Unruh, A. M., J. Ritchie, et al. (1999). "Does gender affect appraisal of pain and pain coping strategies?" Clinical Journal of Pain **15**(1): 31-40.
- van den Brink, M., E. N. Bandell-Hoekstra, et al. (2001). "The occurrence of recall bias in pediatric headache: a comparison of questionnaire and diary data." Headache **41**(1): 11-20.
- van Laerhoven, H., H. J. van der Zaag-Loonen, et al. (2004). "A comparison of Likert scale and visual analogue scales as response options in children's questionnaires." Acta Paediatrica **93**(6): 830-5.
- Walco, G. A., R. C. Cassidy, et al. (1994). "Pain, hurt, and harm - The ethics of pain control in infants and children." New England Journal of Medicine **331**(8): 541-544.
- Whipple, J. K., K. S. Lewis, et al. (1995). "Analysis of pain management in critically ill patients." Pharmacotherapy **15**(5 1): 592-599.
- Williams, A. C., H. T. Davies, et al. (2000). "Simple pain rating scales hide complex idiosyncratic meanings." Pain **85**(3): 457-63.
- Witter, J., L. S. Simon, et al. (2003). "Are Means Meaningless? The Application of Individual Responder Analysis to Analgesic Drug Development." American Pain Society Bulletin **13**(6): 1-7.

**APPENDICE : QUESTIONNAIRES ET FORMULAIRES UTILISÉS**

**Pour les archives :**  
NE PAS METTRE AU DOSSIER, SVP faire parvenir au Dr Raoul Daoust case # 502

**Carte du patient**

Date : \_\_\_\_\_

**Si carte non disponible :**

Prénom, Nom : \_\_\_\_\_

Numéro dossier : \_\_\_\_\_

**Cocher  la réponse**

Patient accepte \_\_\_\_\_

Patient refuse \_\_\_\_\_

Patient incapable de participer :

Incapacité physique \_\_\_\_\_

Ne comprend pas les consignes \_\_\_\_\_

Autres \_\_\_\_\_

**Langue parlée :** (Cocher  la réponse)

Français \_\_\_\_\_

Anglais \_\_\_\_\_

Autres avec traducteur \_\_\_\_\_

**Site de la douleur :** (Cocher  la réponse)

Tête et cou (sauf rachis) \_\_\_\_\_

Rachis \_\_\_\_\_

Thorax \_\_\_\_\_

Abdomen \_\_\_\_\_

Membres \_\_\_\_\_

Autres (spécifier) : \_\_\_\_\_

**VAS**

(Le patient doit mettre une ligne au niveau de l'intensité de la douleur, lui expliquer que la limite de gauche = pas de douleur et la limite de droite = la pire douleur imaginable)  
(En *anglais* utiliser les mots : la gauche = no pain, la droite = worst pain imaginable)

Pas de  
douleur

Pire douleur  
imaginable

Résultat selon la **VAS plastifiée** : \_\_\_\_\_

Résultat selon **l'échelle de 0 à 10** : \_\_\_\_\_

(Échelle verbale numérique)

(Utiliser les mêmes mots que pour la VAS)

Nom et signature du médecin examinateur : \_\_\_\_\_

**Pour les archives :**

NE PAS METTRE AU DOSSIER, SVP faire parvenir au Dr Raoul Daoust case # 502

**Carte du patient**

Date : \_\_\_\_\_

**Si carte non disponible :**

Prénom, Nom : \_\_\_\_\_

Numéro dossier : \_\_\_\_\_

**Cocher  la réponse**Patient accepte \_\_\_\_\_ Patient refuse \_\_\_\_\_ 

Patient incapable de participer :

Incapacité physique \_\_\_\_\_ Ne comprend pas les consignes \_\_\_\_\_ Autres \_\_\_\_\_ 
**Langue parlée :** (Cocher  la réponse)

Français \_\_\_\_\_

Anglais \_\_\_\_\_

Autres avec traducteur \_\_\_\_\_

**Site de la douleur :** (Cocher  la réponse)

Tête et cou (sauf rachis) \_\_\_\_\_

Rachis \_\_\_\_\_

Thorax \_\_\_\_\_

Abdomen \_\_\_\_\_

Membres \_\_\_\_\_

Autres (spécifier): \_\_\_\_\_

Résultat selon la **VAS plastifiée** : \_\_\_\_\_**VAS**

(Le patient doit mettre une ligne au niveau de l'intensité de la douleur, lui expliquer que la limite de gauche = pas de douleur et la limite de droite = la pire douleur imaginable)  
 (En *anglais* utiliser les mots : la gauche = no pain, la droite = worst pain imaginable)

Pas de  
douleur

Pire douleur  
imaginableRésultat selon **l'échelle de 0 à 10** : \_\_\_\_\_

(Échelle verbale numérique)

(Utiliser les mêmes mots que pour la VAS)

Nom et signature du médecin examinateur : \_\_\_\_\_

**Pour les archives :**

NE PAS METTRE AU DOSSIER, SVP faire parvenir au Dr Raoul Daoust case # 502

**Carte du patient**

Date : \_\_\_\_\_

**Si carte non disponible :**

Prénom, Nom : \_\_\_\_\_

Numéro dossier : \_\_\_\_\_

**Cocher  la réponse**Patient accepte \_\_\_\_\_ Patient refuse \_\_\_\_\_ 

Patient incapable de participer :

Incapacité physique \_\_\_\_\_ Ne comprend pas les consignes \_\_\_\_\_ Autres \_\_\_\_\_ **Langue parlée :** (Cocher  la réponse)

Français \_\_\_\_\_

Anglais \_\_\_\_\_

Autres avec traducteur \_\_\_\_\_

**Site de la douleur :** (Cocher  la réponse)

Tête et cou (sauf rachis) \_\_\_\_\_

Rachis \_\_\_\_\_

Thorax \_\_\_\_\_

Abdomen \_\_\_\_\_

Membres \_\_\_\_\_

Autres (spécifier) : \_\_\_\_\_

Résultat selon **l'échelle de 0 à 10** : \_\_\_\_\_

(Échelle verbale numérique)

(Utiliser les mêmes mots que pour la VAS)

Résultat selon la **VAS plastifiée** : \_\_\_\_\_**VAS**

(Le patient doit mettre une ligne au niveau de l'intensité de la douleur, lui expliquer que la limite de gauche = pas de douleur et la limite de droite = la pire douleur imaginable)

(En *anglais* utiliser les mots : la gauche = no pain, la droite = worst pain imaginable)Pas de  
douleurPire douleur  
imaginable

Nom et signature du médecin examinateur : \_\_\_\_\_

Formulaire Access pour le traitement des données.

Microsoft Access - [Échelle Douleur]

Fichier Edition Affichage Insertion Format Enregistrements Outils Fenêtre ?

No 14

Date 2005-08-19

Date de naissance 1968-08-25

Sexe F

Patient Refuse

Patient Accepte

Incapable physique

Incompréhension

Autre

Langue Anglais

Ordre des questions 1

Examineur Externe

Commentaire

Tête cou

Rachis

Thorax

Abdomen

Membres

Autres (spécifier)

Échelle Verbale 5

VAS plastifiée 37

VAS 33

