Numeric Rating Scale (NRS)

Plusieurs auteurs, dérivée de la VAS

Instrument de mesure	Numeric Rating Scale
Abréviation	NRS
Auteur	Plusieurs auteurs
Thème	Management des symptômes de la douleur
Objectif	Evaluer la douleur chez le patient
Population	Non spécifiée
Relevé	Dispensateur de soins
Nombre d'items	1 item
Présence du patient requise	Oui
Localisation de l'instrument de mesure	http://www.ndhcri.org/

Objectif

Contrôle de la douleur chez le patient à l'aide d'un autorapport sur une échelle à un item.

Groupe cible

La Numeric Rating Scale (NRS) a déjà été validée au sein d'une population diversifiée. Des groupes de patients au sein desquels la validité et la fiabilité de l'échelle ont été étudiées sont entre autres des personnes âgées, des patients en service de chirurgie, des enfants, des patients atteints de problèmes musculo-squelettiques, des patients oncologiques et d'autres groupes souffrant de problèmes de douleur (Carey, Turpin, Smith, Whatley, & Haddox, 1997; Gagliese, Weizblit, Ellis, & Chan, 2005; Herr, Spratt, Mobily, & Richardson, 2004; Jensen, Engel, McKearnan, & Hoffman, 2003; Kaasalainen & Crook, 2003; Mawdsley, Moran, & Conniff, 2002; Paice & Cohen, 1997; Puntillo & Neighbor, 1997; Spadoni, Stratford, Solomon, & Wishart, 2004; Taylor & Herr, 2003; Ware, Epps, Herr, & Packard, 2006; Wong & Baker, 1988).

Par conséquent, l'utilisation de la NRS peut être généralisée pour une large population de patients souffrant de douleurs.

Description

La NRS est clairement dérivée de la VAS. Il s'agit également d'une échelle dont les extrêmes sont limités par les termes « Aucune douleur » et « Pire douleur possible ».

Contrairement à la VAS, la NRS est numérotée à intervalles égaux. La division de l'échelle va de 0 à 10. Cela donne 11 possibilités de choix pour le patient, 0 étant égal à « aucune douleur » et 10 à « pire douleur possible ». L'intention est que le patient indique sur la ligne l'intensité de la douleur qu'il ressent au moment du relevé.

Tout comme pour la VAS, il existe de nombreuses variantes de la NRS. Des différences se retrouvent notamment dans la terminologie que l'on emploie pour délimiter le point final (ex. douleur extrême, douleur la plus atroce, pire douleur possible, …), dans le nombre de points pour diviser l'échelle (ex. NRS 6 points, NRS 21 points), dans l'orientation de l'échelle (verticale – horizontale), dans l'intervalle temporel auquel se rapporte l'évaluation de la douleur (ex. « Evaluez votre douleur moyenne sur les 24 dernières heures »), …

Fiabilité

La fiabilité de la NRS a déjà été étudiée dans le cadre de plusieurs études. Plusieurs études font état d'une *intra-rater reliability* supérieure à r = 0.70 (Kaasalainen & Crook, 2003; Mawdsley et al., 2002; Ware, Epps, Herr, & Packard, 2006; Taylor & Herr, 2003). Dans l'étude de Spadoni et al. (2004), cette corrélation atteignait r = 0.63. La valeur peut en être remise en question, étant donné que le retest n'a eu lieu que 72h plus tard. Un changement dans le vécu de la douleur chez le patient peut peut-être être à l'origine de la corrélation inférieure.

Le *coefficient alpha de Cronbach* entre la NRS et d'autres échelles de mesure à un seul item a été vérifié dans l'étude de Herr et al. (2004) et dans l'étude de Carey et al. (1997). La corrélation atteignait respectivement α = 0.97 (VAS – Numeric Rating Scale – Verbal Descriptor Scale – Verbal Numeric Scale – Faces Pain Scale) et α = 0.88 (VAS – Faces Rating Scale – Numeric Rating Scale).

L'*interrater reliability* a généré une concordance de 100% entre les différents évaluateurs (Herr et al., 2004).

Validité

Tout comme pour la fiabilité, la validité de la NRS a déjà été amplement étudiée. Gagliese et al. (2005), Jensen et al. (2003), Kim & Buschmann (2006), Paice & Cohen (1997), Puntillo & Neighbor (1997), Spadoni et al. (2004), Taylor & Herr (2003), Ware et al. (2006), Wong & Baker (1988) ont étudié la *concurrent validity* de la NRS en corrélant l'instrument de mesure à d'autres échelles de douleur (entre autres la VAS, Verbal Rating Scale, le McGill Pain Questionnaire, la Faces Pain Scale, l'Iowa Pain Thermometer…). La moyenne des corrélations obtenues dans ces études était de r = 0.48 – 0.96. La plupart des corrélations obtenues sont toutefois majoritairement élevées. Des comparaisons entre des personnes avec un fonctionnement cognitif amoindri et des personnes cognitivement intactes démontrent que ces corrélations sont plus faibles pour le premier groupe cité (Gagliese et al., 2005; Kaasalainen & Crook, 2003; Taylor & Herr, 2003; Ware et al., 2006). Il se peut que la NRS convienne moins bien pour une utilisation comme instrument de mesure chez des patients avec un fonctionnement cognitif amoindri (ex. déments). Une alternative peut consister à utiliser la Faces Pain Scale ou la Verbal Rating Scale.

Pour contrôler la *construct validity* de la VAS, Gagliese et al. (2005), Jensen et al. (2003) et Herr et al. (2004) ont effectué une *analyse des principaux composants.* Les 3 études ont abstrait 1 facteur au départ des scores sur les différentes échelles de douleur à un item. La charge factorielle de la NRS était chaque fois supérieure à 0.95, excepté dans l'étude de Jensen et al. (2003) où la charge factorielle était de 0.80 pour la NRS.

Des corrélations significatives ont aussi été rapportées entre des scores NRS et l'intensité de la douleur causée (Herr et al., 2004), des observations comportementales du patient ressentant la douleur (Kaasalainen & Crook, 2003) et des éléments en rapport avec la douleur comme la dépression et des limitations fonctionnelles dues à la douleur (Jensen et al., 2003).

Convivialité

La NRS n'a pas été testée sur le plan de sa convivialité, étant donné que le management de la douleur dans les hôpitaux que nous avons interrogés était assuré au moyen de la VAS. Nous pensons toutefois pouvoir dire qu'il existe un chevauchement important entre la convivialité de la NRS et de la VAS. Pour ces raisons, nous faisons référence à la rubrique « Convivialité » de la VAS. En tenant compte du fait que l'écart entre le point zéro et le marquage du patient ne doit pas être mesuré (ce contrairement à la VAS), nous pouvons partir du principe que moins d'erreurs seront commises par les dispensateurs de soins et que le relevé de la NRS se fera plus rapidement.

Remarques

Une comparaison entre les différents instruments de mesure démontre que la NRS est une échelle hautement fiable et valide au niveau de l'appréciation de la douleur. Bien que la VAS soit utilisée le plus fréquemment dans la pratique, la NRS mérite peut-être d'être préférée à la VAS. Plusieurs études rapportent en effet des difficultés éprouvées par des patients pour remplir la VAS (Gagliese et al., 2005; Herr et al., 2004; Paice & Cohen 1997; Pautex et al., 2006). Des comparaisons entre la VAS et la Numeric Rating Scale (NRS) font état de moins de problèmes avec l'utilisation de cette dernière échelle. De plus, la NRS bénéficie d'une fiabilité et d'une validité comparables. En outre, des patients eux-mêmes indiquent qu'ils préfèrent la NRS et d'autres échelles de douleur à la VAS (Carey et al., 1997; Dworkin et al., 2005; Gagliese et al., 2005; Herr et al., 2004; Paice & Cohen 1997). Plusieurs études de validation indiquent donc que la NRS est l'instrument le plus adéquat pour l'évaluation de la douleur (Downie et al., 1978; Dworkin et al., 2005; Jensen et al., 2003). Cela doit être quelque peu nuancé, puisque aucune échelle de douleur ne peut être utilisée pour toutes les formes de douleur ou pour toutes les catégories d'âge. En contraste avec la VAS, une étude approfondie de la sensitivité de la NRS fait également défaut.

Bien que la fiabilité et la validité de la NRS soient soutenues chez les enfants (Jensen & Karoly in Gagliese et al., 2005), il peut être indiqué d'opter pour la FPS pour le contrôle de la douleur chez ce groupe de patients. En particulier chez les jeunes enfants (< 12 ans), étant donné que la FPS est plus claire et plus

compréhensible pour eux. De plus, les propriétés psychométriques de la FPS ont déjà été amplement étudiées et cela confirme l'utilisation de la FPS chez les enfants (Bieri, Reeve, Champion, Addicoat, & Ziegler, 1990; Bosenberg, Thomas, Lopez, Kokinsky & Larsson, 2003; Hicks, von Baeyer, Spafford, van Korlaar & Goodenough, 2001; Paik & Ahn, 2002; Stinson, Kavanagh, Yamada, Gill & Stevens, 2006; Wong & Baker, 1988).

Références

Bieri, D., Reeve, R. A., Champion, G. D., Addicoat, L., & Ziegler, J. B. (1990). The Faces Pain Scale for the self-assessment of the severity of pain experienced by children: development, initial validation, and preliminary investigation for ratio scale properties. *Pain, 41,* 139-150.

Bosenberg, A., Thomas, J., Lopez, T., Kokinsky, E., & Larsson, L. E. (2003). Validation of a six-graded faces scale for evaluation of postoperative pain in children. *Paediatr Anaesth., 13,* 708-713.

Carey, S. J., Turpin, C., Smith, J., Whatley, J., & Haddox, D. (1997). Improving pain management in an acute care setting. The Crawford Long Hospital of Emory University experience. *Orthop.Nurs., 16,* 29-36.

Downie, W. W., Leatham, P. A., Rhind, V. M., Wright, V., Branco, J. A., & Anderson, J. A. (1978). Studies with pain rating scales. *Ann Rheum Dis, 37,* 378-381.

Dworkin, R. H., Turk, D. C., Farrar, J. T., Haythornthwaite, J. A., Jensen, M. P., Katz, N. P. et al. (2005). Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain, 113,* 9-19.

Gagliese, L., Weizblit, N., Ellis, W., & Chan, V. W. (2005). The measurement of postoperative pain: a comparison of intensity scales in younger and older surgical patients. *Pain, 117,* 412-420.

Herr, K. A., Spratt, K., Mobily, P. R., & Richardson, G. (2004). Pain intensity assessment in older adults: use of experimental pain to compare psychometric properties and usability of selected pain scales with younger adults. *Clin.J.Pain, 20,* 207-219.

Hicks, C. L., von Baeyer, C. L., Spafford, P. A., van, Korlaar, I, & Goodenough, B. (2001). The Faces Pain Scale-Revised: toward a common metric in pediatric pain measurement. *Pain, 93,* 173-183.

Jensen, M. P., Engel, J. M., McKearnan, K. A., & Hoffman, A. J. (2003). Validity of pain intensity assessment in persons with cerebral palsy: a comparison of six scales. *J Pain, 4,* 56-63.

Kim, E. J. & Buschmann, M. T. (2006). Reliability and validity of the Faces Pain Scale with older adults. *Int J Nurs Stud., 43,* 447-456.

Kaasalainen, S. & Crook, J. (2003). A comparison of pain-assessment tools for use with elderly long-term-care residents. *Can J Nurs Res, 35,* 58-71.

Mawdsley, R. H., Moran, K. A., & Conniff, L. A. (2002). Reliability of two commonly used pain scales with elderly patients. *Journal of Geriatric Physical Therapy, 25,* 16-20.

Paice, J. A. & Cohen, F. L. (1997). Validity of a verbally administered numeric rating scale to measure cancer pain intensity. *Cancer Nurs, 20,* 88-93.

Paik, H. J. & Ahn, Y. M. (2002). Measurement of acute pain after eye surgery in children. *Korean J Ophthalmol., 16,* 103-109.

Pautex, S., Michon, A., Guedira, M., Emond, H., Le Lous, P., Samaras, D. et al. (2006). Pain in severe dementia: self-assessment or observational scales? *J Am Geriatr Soc, 54,* 1040-1045.

Puntillo, K. A. & Neighbor, M. L. (1997). Two methods of assessing pain intensity in English-speaking and Spanish-speaking emergency department patients. *J Emerg Nurs, 23,* 597-601.

Spadoni, G. F., Stratford, P. W., Solomon, P. E., & Wishart, L. R. (2004). The evaluation of change in pain intensity: a comparison of the P4 and single-item numeric pain rating scales. *J Orthop Sports Phys Ther, 34,* 187-193.

Stinson, J. N., Kavanagh, T., Yamada, J., Gill, N., & Stevens, B. (2006). Systematic review of the psychometric properties, interpretability and feasibility of self-report pain intensity measures for use in clinical trials in children and adolescents. *Pain, 125,* 143-157.

Taylor, L. J. & Herr, K. (2003). Pain intensity assessment: a comparison of selected pain intensity scales for use in cognitively intact and cognitively impaired African American older adults. *Pain Manag.Nurs, 4,* 87-95.

Ware, L. J., Epps, C. D., Herr, K., & Packard, A. (2006). Evaluation of the Revised Faces Pain Scale, Verbal Descriptor Scale, Numeric Rating Scale, and Iowa Pain Thermometer in older minority adults. *Pain Manag.Nurs, 7,* 117-125.

Wong, D. L. & Baker, C. M. (1988). Pain in children: comparison of assessment scales. *Pediatr Nurs, 14,* 9-17.

Localisation de l'instrument de mesure

http://www.ndhcri.org/hcqip/

NUMERIC RATING SCALE (NRS)

VERSCHEIDENE AUTEURS SINDS1974

Author (year)	Setting	Sample (n)	Design	Reliability	Validity
Mawdsley, R. H., Moran, K. A., & Conniff, L. A. (2002)	An outpatient setting.	Elderly with musculoskeletal injury or disorders with pain aged 61 to 84. (n = 32)	Repeated measures design: VAS and NRS measurements 15 minutes apart.	S	
Gagliese, L., Weizblit, N., Ellis, W., & Chan, V. W. (2005)	Toronto General Hospital, University Health Network and Mount Sinai Hospital, Toronto, Ontario.	Patients scheduled to receive patient-controlled analgesia following general surgery. (n = 504)	Comparative study: the VAS was compared with the Numeric Rating Scale (NRS), the Verbal Descriptor Scale (VDS) and the pain intensity measures of the McGill Pain Questionnaire (MPQ).		FV CrV CsV

Reliability: Stability (S), Internal consistency (IC), Equivalence (E)
Validity: Face validity (FV), Content validity (CtV), Criterion validity (CrV), Construct validity (CsV)
Sensitivity (Sen), Specificity (Sp), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Receiver Operating Curve (ROC), Likelihood Ratio (LR), Odds Ratio (OR)

Results reliability	Results validity	Commentary
(S) *Intra-rater reliability NRS:* r = 0.74 *Intra-rater reliability VAS:* r = 0.92		
	(FV) *Face validity;* Patients rated the VAS as the least accurate and least preferred instrument for future assessments.The VAS had the highest proportion of patients making errors (p ≤ 0.005). **(CrV)** *Concurrent validity:* - Correlation between VAS and VDS and NRS: 3. Patients aged > 60 years: r = 0.60-0.93 (p<0.001) 4. Patients aged ≤ 60 years: r = 0.72-0.91 (p<0.001) - Correlation between VAS and MPQ was lower. **(CsV)** *Principal components analysis:* A one-factor model was extracted from the different measurement scales and accounted for 82% of the total variance in the young group and 85% of the variance in the older group. Factor loading was 0.96.	

Reliability: Stability (S), Internal consistency (IC), Equivalence (E)
Validity: Face validity (FV), Content validity (CtV), Criterion validity (CrV), Construct validity (CsV)
Sensitivity (Sen), Specificity (Sp), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Receiver Operating Curve (ROC), Likelihood Ratio (LR), Odds Ratio (OR)

Author (year)	Setting	Sample (n)	Design	Reliability	Validity
Spadoni, G. F., Stratford, P. W., Solomon, P. E., & Wishart, L. R. (2004)	20 outpatient physical therapy clinics. The clinics were located in the Canadian Provinces of Alberta, Ontario, Quebec and the State of Georgia, USA.	220 patients with musculoskeletal problems receiving physical therapy. 213 patients (97%) were available for retest; and 183 patients (83%) provided data at follow-up. (n = 220)	Comparative study: To determine whether the P4 was more adept at assessing pain change than 2 versions of a single-item NRS. One version inquired about pain intensity over the past 24 hours and the second version asked about pain intensity over the past 2 days. NRSs and the P4 were administered on 3 occasions - initial visit, within 72 hours of baseline, and 12 days following baseline assessment.	S	CrV
Paice, J. A. & Cohen, F. L. (1997)	A large tertiary care hospital.	A convenience sample of 50 adult oncological patients with pain. (n = 50)	Comparative study: the NRS was compared with VAS and Simple Descriptor Scale (SDS).		CrV
Carey, S. J., Turpin, C., Smith, J., Whatley, J., & Haddox, D. (1997)	Inpatient units of the Crawford Long Hospital, Atlanta.	The admitted diagnosis for 39.5% of the sample was acute pain, 40.3 with chronic pain and 20.2 with no pain. (n = 267)	Comparative study: the VAS was compared with FRS and NRS.	IC	

Reliability: Stability (S), Internal consistency (IC), Equivalence (E)
Validity: Face validity (FV), Content validity (CtV), Criterion validity (CrV), Construct validity (CsV)
Sensitivity (Sen), Specificity (Sp), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Receiver Operating Curve (ROC), Likelihood Ratio (LR), Odds Ratio (OR)

Results reliability	Results validity	Commentary
(S) *Test-retest reliability:* The second assessment was obtained within 72 hours of the baseline visit. - P4: r = 0.78 (95% CI [0.72, 0.83]) - NRS 24-hour: r = 0.63 (95% CI [0.54, 0.70]) - NRS 2 day : r = 0.70 (95% CI [0.62, 0.77])	**(CrV)** *Concurrent validity:* The criterion validity analysis produced a correlation coefficient of 0.57 between the P4 and 24-hour NRS. This difference is statistically significant (z = 1.73, p = .043). The longitudinal validity analysis yielded correlation coefficients of 0.61 and 0.56 between the retrospective rating of change and the P4 and 2-day NRS, respectively. This difference is statistically significant (z = 2.53, p = .006).	Ratings are retrospective. This might bias the validity of the NRS-scores. However, the aim of the NRS is to rate the present pain of the patient and not the pain of the past few days. We cast doubt on such longitudinal measurements; an adequate pain policy requires pain assessment on a regularly base and not every two days.
	(CrV) *Concurrent validity:* -The strong positive correlation between the NRS and VAS was statistically significant: r = 0.85, p < 0.001. -NRS was also correlated to SDS: r = 0.83, p < 0.001	A majority of subjects (50%) preferred the use of the NRS when comparing the three scales used to measure pain intensity. Fewer patients preferred the SDS (38%), with the VAS chosen least often (12%). Eleven subjects (20%) in this study were unable to complete the VAS or did so with great difficulty. All subjects were able to complete the NRS and SDS without apparent difficulty.
(IC) *Cronbach's alpha:* Intercorrelations between 3 scales was α = 0.88.		Patients selected the FPS as 'easiest to use' (48.6%), followed by the NRS (35.3%) and the VAS (16.1%).

Reliability: Stability (S), Internal consistency (IC), Equivalence (E)
Validity: Face validity (FV), Content validity (CtV), Criterion validity (CrV), Construct validity (CsV)
Sensitivity (Sen), Specificity (Sp), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Receiver Operating Curve (ROC), Likelihood Ratio (LR), Odds Ratio (OR)

Author (year)	Setting	Sample (n)	Design	Reliability	Validity
Ware, L. J., Epps, C. D., Herr, K., & Packard, A. (2006)	Not specified.	68 subjects aged 60 years and older whith cognitive impairments. The mean score for the Mini Mental State Examination (MMSE) was 23 (standard deviation = 5.4) with a range from 10 to 30. Fifty-nine percent (n = 40) of the sample scored 24 or greater indicating no cognitive impairment (CI). Forty-one percent (n = 28) scored less than 24 indicating some degree of CI. (n = 68)	Comparative study: Subjects were instructed to recall a vividly remembered pain and rate this remembered pain using the Iowa Pain Thermometer (IPT), the Verbal Descriptor Scale (VDS), a 0 to 10 Numeric Rating Scale (NRS), and the Faces Pain Scale Revised (FPS-R).	S	CrV

Reliability: Stability (S), Internal consistency (IC), Equivalence (E)
Validity: Face validity (FV), Content validity (CtV), Criterion validity (CrV), Construct validity (CsV)
Sensitivity (Sen), Specificity (Sp), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Receiver Operating Curve (ROC), Likelihood Ratio (LR), Odds Ratio (OR)

Results reliability	Results validity	Commentary
(S) *Test-retest reliability:* The researcher returned 2 weeks later and reminded subjects of the "vividly remembered pain" identified at the first assessment and asked them to rate that pain again using all four scales. In the cognitively intact group, Spearman rank correlation coefficients between the two-week vividly remembered pain ratings were 0.87 (NRS), followed by the VDS (0.86), IPT (0.81), and FPS-R (0.76). Correlations ranged from 0.77 to 0.89 in the CI group. The FPS-R had the strongest test–retest reliability coefficient (0.89) in this group followed by the IPT (0.82), VDS (0.79), and NRS (0.77).	**(CrV)** *Concurrent validity:* Spearman correlations between pain ratings on the selected scales ranged from 0.64 to 0.90 in the cognitively intact group and from 0.56 to 0.83 in the CI group. The lowest correlations in the CI group and intact group (0.64–0.84) were found between the FPS-R and the other scales (0.56–0.66).	In terms of the concurrent validity, moderate to high inter-tool correlations for the CI and cognitively intact groups were found with the exception of low correlations associated with the FPS-R, suggesting that the FPS-R may measure overall affect as opposed to pain. Four participants with moderate CI were unable to follow directions and complete the VDS and IPT. The NRS had the highest failure rate with six participants with moderate CI and one mildly impaired participant unable to use the scale. No failures occurred when using the FPS-R to evaluate pain. The NRS (n = 12/36, 33%) was the preferred scale in the cognitively intact group, and the FPS-R (n = 13/24, 54%) was the preferred scale in the CI group.

Reliability: Stability (S), Internal consistency (IC), Equivalence (E)
Validity: Face validity (FV), Content validity (CtV), Criterion validity (CrV), Construct validity (CsV)
Sensitivity (Sen), Specificity (Sp), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Receiver Operating Curve (ROC), Likelihood Ratio (LR), Odds Ratio (OR)

Author (year)	Setting	Sample (n)	Design	Reliability	Validity
Herr, K. A., Spratt, K., Mobily, P. R., & Richardson, G. (2004)	Subjects were recruited through college bulletin board displays, community faith centers, senior-citizen centers, senior-housing, and long-term facilities.	86 younger adults (age 25-55). 89 older adults (age 65-94). (n = 175)	Repeated measures design: Noxious heat stimili were delivered to the ventral forearm by an electronically controlled contact thermode. The heat stimili were programmed to last 5 seconds and to present randomly 43°C, 45°C, 46°C, 47°C, 48°C, 49°C or 51°C. A 2 minute trial interval followed each stimulus presentation during which the subject rated the stimulus on 5 scales.	IC E	CsV Sen

Reliability: Stability (S), Internal consistency (IC), Equivalence (E)
Validity: Face validity (FV), Content validity (CtV), Criterion validity (CrV), Construct validity (CsV)
Sensitivity (Sen), Specificity (Sp), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Receiver Operating Curve (ROC), Likelihood Ratio (LR), Odds Ratio (OR)

Results reliability	Results validity	Commentary
(IC) *Internal consistency:* Intercorrelations between the scales were all statistically significant at every temperature ($p < 0.001$). *Cronbach's alpha within each scale across the 7 temperature values:* VAS: 0.88 NRS: 0.88 VDS: 0.86 VNS: 0.88 FPS: 0.88 *Cronbach's alpha within temperature across the 5 scales:* 43°C: 0.96 45°C: 0.96 46°C 0.96 47°C: 0.97 48°C: 0.97 49°C: 0.97 51°C: 0.97 **(E)** *Interrater reliability:* VAS: 93.5% agreement NRS: 100% agreement VDS: 100% agreement VNS: 100% agreement FPS: 100% agreement	**(CsV)** *Principal components analysis:* A one-factor model was extracted from the different measurement scales. The correlation of each scale to the isolated factor was as follows: VAS: 0.94 NRS: 0.96 VDS: 0.95 VNS: 0.95 FPS: 0.86 **(Sen)** Each tool demonstrated significant increases in score associated with increase in temperature ($p < 0.001$).	The VAS had a significant higher failure rate (6.7%) in comparison with the NRS, VDS, VNS, and FPS. The VNS demonstrated significantly higher levels of pain report than the other 4 scales. The psychometric scale evaluation was conducted using an experimental pain stimulus instead of using clinical pain stimili. The scales most preferred in order by the total sample: NRS (35.3%); VDS (25.3%); VNS (15.9%); FPS (12.9%); VAS (10.6%).

Reliability: Stability (S), Internal consistency (IC), Equivalence (E)
Validity: Face validity (FV), Content validity (CtV), Criterion validity (CrV), Construct validity (CsV)
Sensitivity (Sen), Specificity (Sp), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Receiver Operating Curve (ROC), Likelihood Ratio (LR), Odds Ratio (OR)

Author (year)	Setting	Sample (n)	Design	Reliability	Validity
Kaasalainen, S. & Crook, J. (2003)	A 240-bed long-term-care facility in urban southwestern Ontario, Canada.	4 groups of 130 elderly long-term-care residents: (1) cognitively intact, (2) mildly cognitively impaired, (3) moderately cognitively impaired, and (4) extremely cognitively impaired. (n = 130)	Repeated measures design: FPS, NRS, Present Pain Intensity Scale (PPI) were conducted twice 48 hours apart.	S	CsV

Reliability: Stability (S), Internal consistency (IC), Equivalence (E)
Validity: Face validity (FV), Content validity (CtV), Criterion validity (CrV), Construct validity (CsV)
Sensitivity (Sen), Specificity (Sp), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Receiver Operating Curve (ROC), Likelihood Ratio (LR), Odds Ratio (OR)

Results reliability	Results validity	Commentary
(S) *Test-retest reliability:* Test-retest reliability for the three verbal-report scales was moderate to strong for the cognitively intact group (FPS: ICC = 0.84; PPI: ICC = 0.55; NRS: ICC = 0.87) but decreased for the other groups. In addition, the error variances were low for the cognitively intact group (FPS: s2 error = 0.53; PPI: s2 error = 0.71; NRS: s2 error = 1.45) but increased with increasing cognitive impairment.	**(CsV)** *Convergent validity:* The Pearson r correlations of the Pain Assessment in the communicatively Impaired (PACI) tool, a behavioural-observation measure, with the three verbal-report scales (FPS, PPI, NRS) were low to moderate. For the cognitively intact group, all of these correlations were moderate and significant (FPS: $r = 0.66$, $p < 0.001$; PPI: $r = 0.62$, $p < 0.01$; NRS: $r = 0.65$, $p < 0.01$). For the mildly impaired group, none were significant at the $p < 0.05$ level. For the moderately impaired group, the PACI correlated moderately and significantly with the FPS ($r = 0.63$, $p < 0.001$) and PPI ($r = 0.64$, $p < 0.001$). However, the correlation between the PACI and NRS for those with moderate impairment was low and non significant ($r = 0.30$, $p < 0.12$).	Test-retest for the 3 scales was moderate to strong for elderly persons with no cognitive impairment but decreased for the other groups. Similarly, error variances were low for those with no cognitive impairment but increased with increasing cognitive impairment. These findings indicate that the level of cognitive impairment decreases the reliability of these tools. Test-retest reliability for both the NRS and the FPS was strong for residents without cognitive impairment but declined considerably for those with mild and moderate impairment, suggesting that these tools may not be good choices for use with these two groups. For those with moderate cognitive impairment, the PPI seems to be a more appropriate and reliable tool than the FPS or the NRS.

Reliability: Stability (S), Internal consistency (IC), Equivalence (E)
Validity: Face validity (FV), Content validity (CtV), Criterion validity (CrV), Construct validity (CsV)
Sensitivity (Sen), Specificity (Sp), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Receiver Operating Curve (ROC), Likelihood Ratio (LR), Odds Ratio (OR)

Author (year)	Setting	Sample (n)	Design	Reliability	Validity
Taylor, L. J. & Herr, K. (2003)	Recruitment took place through Elder Care Services, a private non-profit organization.	A convenience sample of 57 volunteers age 58 and older. Seventy-seven percent (n = 44) of the sample scored 24 or less on the mental status exam, indicating some degree of cognitive impairment. The remaining 23% (n = 13) were cognitively intact. (n = 57)	Comparative study: Subjects were instructed to recall a vividly remembered pain and rate this remembered pain using the FPS, the VDS, the NRS and the Iowa Pain Thermometer (IPT).	S	CrV

Reliability: Stability (S), Internal consistency (IC), Equivalence (E)
Validity: Face validity (FV), Content validity (CtV), Criterion validity (CrV), Construct validity (CsV)
Sensitivity (Sen), Specificity (Sp), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Receiver Operating Curve (ROC), Likelihood Ratio (LR), Odds Ratio (OR)

Results reliability	Results validity	Commentary
(S) *Test-retest reliability:* The researcher returned 2 weeks later and reminded subjects of the "vividly remembered pain" identified at the first assessment and asked them to rate that pain again using all four scales. Spearman rank correlation coefficients between the 2-week vividly remembered pain ratings ranged from 0.52 to 0.83 in both groups. In the cognitively impaired group, the FPS had the strongest reliability coefficient (0.79), followed by the VDS (0.63), NRS (0.57) and IPT (0.52). In the intact group, the strongest correlation was noted with the IPT (0.83), followed by the FPS (0.81), NRS (0.74), and VDS (0.73).	**(CrV)** *Concurrent validity:* Spearman correlations between present pain ratings on the selected scales were statistically significant ($p = 0.01$) and ranged from 0.81 to 0.96 in the intact group and from 0.74 to 0.83 in the impaired group. The lowest correlation was found between the FPS and VDS ($r = 0.74$).	All of the participants were able to use each of the pain intensity scales to rate their present pain in a manner that allowed interpretation of a single pain score (e.g., not selecting more than one response, selection of options outside the scale range). No failures were noted. The FPS showed strong test-retest stability in the cognitively impaired minority older adults. Of the 35 older adults who identified a scale preference, the FPS was the preferred scale in both the intact group (n = 5, 62.5%) and the impaired group (n = 15, 56%). However, it should be noted that almost 40% of both groups had no preference for a specific pain intensity scale.

Reliability: Stability (S), Internal consistency (IC), Equivalence (E)
Validity: Face validity (FV), Content validity (CtV), Criterion validity (CrV), Construct validity (CsV)
Sensitivity (Sen), Specificity (Sp), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Receiver Operating Curve (ROC), Likelihood Ratio (LR), Odds Ratio (OR)

Author (year)	Setting	Sample (n)	Design	Reliability	Validity
Puntillo, K. A. & Neighbor, M. L. (1997)	A level 1 trauma center emergency department in San Francisco.	Ninety-five English-speaking patients (67 male and 28 female) and 21 Spanish-speaking patients (16 male and 5 female) participated in the study. Study patients were treated in the emergency department for sprains or strains (35%), fractures (19%), contusions (14%), cellulitis or abscesses (12%), or other miscellaneous conditions such as headaches or abdominal/flank pain (20%). (n = 59)	Repeated measures design: Patients were asked to use the NRS and VRS (English or Spanish version according the language of the patient) seven times over a 2-hour period, once immediately before and six times after receiving an analgesic.		CrV
Wong, D. L. & Baker, C. M. (1988)	Pediatric units of two general hospitals in the South-Central United States.	Hospitalized children in three age groups: 3 to 7 (n = 52), 8 to 12 (n = 52), and 13 to 18 years (n = 46). (n = 150)	Comparative study. The following six scales were compared on reliability and validity: Simple Descriptive Scale (SDS), NRS, FPS, the Glasses Scale, the Chips Scale, and the Color Scale.	S	CrV

Reliability: Stability (S), Internal consistency (IC), Equivalence (E)
Validity: Face validity (FV), Content validity (CtV), Criterion validity (CrV), Construct validity (CsV)
Sensitivity (Sen), Specificity (Sp), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Receiver Operating Curve (ROC), Likelihood Ratio (LR), Odds Ratio (OR)

Results reliability	Results validity	Commentary
	(CrV) *Concurrent validity:* Spearman correlations between scores on the two scales were computed for each of the seven data collection times. Correlations between the two scales were moderate to very high ($r = 0.48$ to 0.96) and statistically significant ($p < 0.05$ to $p < 0.001$) at each of the seven time periods. In fact, there was an increase in the relationship between the two pain scores (NRS and VRS) from the beginning to the end of the study.	Slightly more English speaking patients preferred the NRS (50) over the VRS (35), and slightly more Spanish-speaking patients preferred the VRS (11) over the NRS (9). However, these findings were not statistically significant.
(S) *Test-retest reliability:* Retest occurred the day after the initial test except in a few instances when it was done 2 to 4 days later. Percentage agreement for the 6 scales was: - SDS: 72.73% - NRS: 75.44% - FPS: 74.24% - Glasses: 75.38% - Chips: 77.27% - Colors: 68.18%	**(CrV)** *Concurrent validity:* Children were asked to list painful events that they had experienced since being hospitalized and rank these from most to least painful. Each scale was also used to rate these painful events. A percentage consistency between the ranking of the events and the responses for each pain scale was calculated. Percentage agreement for the 6 scales was: - SDS: 62.81% - NRS: 60.00% - FPS: 60.43% - Glasses: 63.70% - Chips: 69.06% - Colors: 58.39%	The FPS was the most preferred scale by all age groups. The finding of an increase in validity and reliability with age is consistent with children' s advancing cognitive ability. However, reliability increased only from the 3 to 7 year age group. Reliability decreased in the 13 to 18 year age group for all the scales except for the color scale and the SDS. No one scale demonstrates superiority in validity or reliability. No significant differences exist among the scales for any age group.

Reliability: Stability (S), Internal consistency (IC), Equivalence (E)
Validity: Face validity (FV), Content validity (CtV), Criterion validity (CrV), Construct validity (CsV)
Sensitivity (Sen), Specificity (Sp), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Receiver Operating Curve (ROC), Likelihood Ratio (LR), Odds Ratio (OR)

Author (year)	Setting	Sample (n)	Design	Reliability	Validity
Jensen, M. P., Engel, J. M., McKearnan, K. A., & Hoffman, A. J. (2003)	Not specified.	Persons with cerebral palsy (CP) who have reported the presence of a chronic pain problem. Pain intensity assessment data were available for 24 participants from the survey study and 45 participants from the longitudinal study (n = 69)	Comparative study . Study participants were recruited from 2 other ongoing studies (a single-assessment survey and a 2-year longitudinal study). A subgroup of those who reported ongoing problems with pain were recruited to participate in a longitudinal study and completed measures of pain and pain impact at 5 time points (11- and 21 point NRS, 5- and 16 point VRS, 6- and 7 point FPS). All of the current study participants came from the same population.		CrV CsV

Reliability: Stability (S), Internal consistency (IC), Equivalence (E)
Validity: Face validity (FV), Content validity (CtV), Criterion validity (CrV), Construct validity (CsV)
Sensitivity (Sen), Specificity (Sp), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Receiver Operating Curve (ROC), Likelihood Ratio (LR), Odds Ratio (OR)

Results reliability	Results validity	Commentary
	(CrV) *Concurrent validity:* There was a strong association among all measures, with the 21-point NRS showing the most consistently strong (all r 's > 0.80) association with the other measures. However, even the weakest association (r = 0.59), found between the NRS-11 and the FPS-7, indicated a great deal of overlap between these 2 measures. **(CsV)** *Factorananalyse:* A single factor emerged. The loadings, all 0.90 or greater (except NRS-11 0.80), support the validity of each of the scales as measures of pain intensity. *Convergent validity:* Correlation coefficients between each of the 6 measures and measures of pain interference (a modified version of the Pain Interference Scale of the Brief Pain Inventory) and depression (Center for Epidemiological Studies–Depression Scale, CES-D), were all in the expected direction, although some variability in the coefficients can be seen. The 7-point Faces scale appeared to be most strongly associated with these 2 measures, the NRS-11 and VRS-5 showed the weakest associations with pain interference, and the VRS-5 showed the weakest association with depression. The difference between the association between the NRS-11 and pain interference (r = 0.25) and the 7-point Faces scale (r = 0.50) was statistically significant (t (42) = 2.46, p < 0.05, for the difference between coefficients).	It is possible that the FPS-7 reflects other dimensions such as affect because of the strong correlation with depression.

Reliability: Stability (S), Internal consistency (IC), Equivalence (E)
Validity: Face validity (FV), Content validity (CtV), Criterion validity (CrV), Construct validity (CsV)
Sensitivity (Sen), Specificity (Sp), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Receiver Operating Curve (ROC), Likelihood Ratio (LR), Odds Ratio (OR)

Author (year)	Setting	Sample (n)	Design	Reliability	Validity
Kim, E. J. & Buschmann, M. T. (2006)	A general hospital and an oriental medical hospital in Korea.	85 older adults with chronic pain (i.e., a state of pain ≥ 6 months duration and for which the cause of the pain could not be removed) were recruited. (n = 85)	Comparative study.		CrV

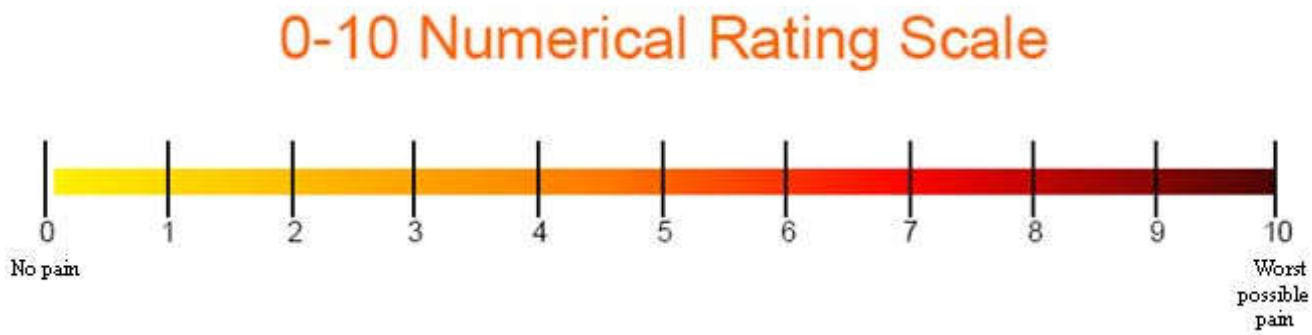
Reliability: Stability (S), Internal consistency (IC), Equivalence (E)
Validity: Face validity (FV), Content validity (CtV), Criterion validity (CrV), Construct validity (CsV)
Sensitivity (Sen), Specificity (Sp), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Receiver Operating Curve (ROC), Likelihood Ratio (LR), Odds Ratio (OR)

Results reliability	Results validity	Commentary
	(CrV) *Concurrent validity:* Results indicate that pain intensities by the FPS, the NRS and the VAS were not significantly different ($F(1.00, 68.00) = 2.93$, $p = 0.09$); likewise the tools were not different in reporting the intensity of patient's pain. The NRS had moderately strong correlation with the FPS ($r = 0.73$, $p < 0.001$) and the VAS ($r = 0.91$, $p < 0.001$). Correlation between the VAS and the NRS of the cognitively impaired group ($r = 0.88$, $p < 0.001$) was weaker than the correlation of the cognitively intact group ($r = 0..92$, $p < 0.001$), correlation between the FPS and the NRS of the cognitively impaired group ($r = 0.75$, $p< 0.001$) was slightly stronger than correlation of the cognitively intact group ($r = 0.70$, $p< 0.001$).	

Reliability: Stability (S), Internal consistency (IC), Equivalence (E)
Validity: Face validity (FV), Content validity (CtV), Criterion validity (CrV), Construct validity (CsV)
Sensitivity (Sen), Specificity (Sp), Positive Predictive Value (PPV), Negative Predictive Value (NPV), Receiver Operating Curve (ROC), Likelihood Ratio (LR), Odds Ratio (OR)

Numerical Rating Scale (NRS)

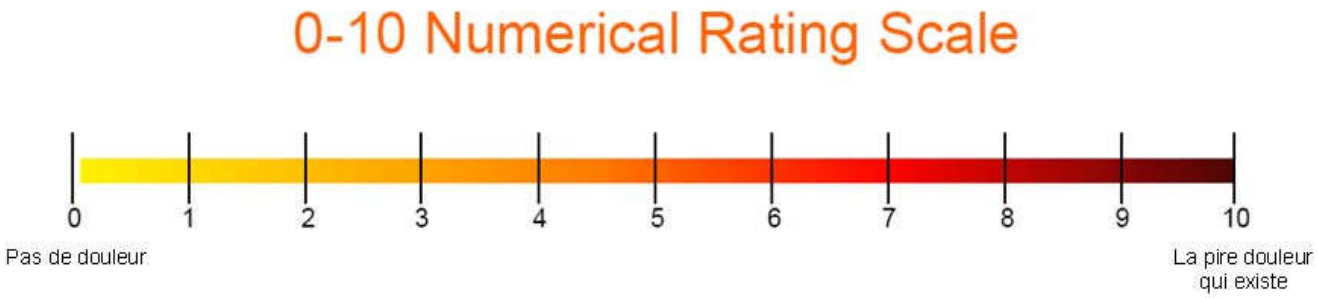
Instructions: Show the pain scale to the resident. Verbally read the scale to the resident and wait for a reply. On the 0-10 pain rating scale, 0 means no pain and 10 means the worst pain possible. The middle of the scale around 5 is moderate pain. A 2 or 3 would be mild pain, but 7 and higher is severe pain. Repeat the directions if the resident is having difficulty; use words other than "pain": aching, cramping, sore, uncomfortable, stiff, dull, pressure, burning, shooting. If the resident does not like it or understand it, switch to another scale. Always use the same scale for each follow-up assessment. Document the scale used as the Numerical Rating Scale (NRS).



Reprinted with permission from the North Dakota Health Care Review, Inc. (NDHCRI)

Numerical Rating Scale (NRS)

Instructions : Montrez l'échelle de douleur au patient. Expliquez l'échelle au patient et attendez sa réponse : "Sur cette échelle de la douleur de 0 à 10, 0 signifie 'pas de douleur' et 10 signifie 'la pire douleur qui existe'. Le milieu de l'échelle est à 5 et signifie 'une douleur moyenne'. Deux ou 3 devrait être une douleur légère, mais 7 et plus signifie une douleur sévère". Répétez les instructions si le patient a des difficultés, utilisez d'autres mots que "douleur" : mal, crampe, irritation, inconfort, courbature, engourdissement, pression, brûlure, lancement. Si le patient ne comprend toujours pas ou ne répond pas, utilisez une autre échelle. Utilisez toujours la même échelle si vous voulez suivre l'évolution du patient. Indiquez que vous avez utilisé l'échelle Numerical Rating Scale (NRS).



Qu'est-ce que BEST ?

BEST pour Belgian Screening Tools est le nom d'une étude réalisée par l'Université de Gand, service des Sciences Infirmières, à la demande du Service Public Fédéral de la Santé Publique, Sécurité Alimentaire et Environnement.

Objectif de BEST ?

Le but de ce projet est de construire une base de données contenant des instruments de mesures validés scientifiquement. Dans le but d'objectiver les diagnostics et résultats des interventions infirmières, des instruments de mesures fiables et valides doivent être disponibles pour démontrer l'efficience des soins infirmiers.

Notre attention se porte sur les instruments de mesure utilisables pour scorer les interventions infirmières du nouveau Résumé Infirmier Minimum ou DI-RHM.

Que pouvez-vous trouver dans ce rapport ?

Le rapport décrit les différents instruments de mesure. En plus, si nous en avons reçu l'autorisation des auteurs, l'instrument est mis à votre disposition. Les instruments de mesure présentant une fiabilité et une validité élevées ont également fait l'objet d'une traduction vers le néerlandais et le français.

Les chefs de projet UGent
Prof. dr. T. Defloor
Prof. dr. M. Grypdonck

Les collaborateurs du projet UGent
M. Daem
Dr. K. Vanderwee

Le chef de projet UCL
Dr. M. Gobert

Le collaborateur du projet UCL
C. Piron

Le chef de projet FOD
B. Folens

Le collaborateur du projet FOD
M. Lardennois

Daem, M., Piron, C., Lardennois, M., Gobert, M., Folens, B., Spittaels, H., Vanderwee, K., Grypdonck, M., & Defloor T. (2007). Mettre à disposition une base de données d'instruments de mesure validés: le projet BEST. Bruxelles: Service Publique Fédéral Santé Publique, Sécurité de la Chaîne alimentaire et Environnement.