Comparison of different scales measuring pain intensity in patients with Whiplash Associated Disorders

Sanneke Grootjans, fysiotherapie student 1557148

Hogeschool Utrecht, 2011-07-25

Supervisors: Hans Heneweer *and Steven Kamper**

* *Hogeschool Utrecht*

** *George Institute for Global Health, Sydney- EMGO institute, VU*

Abstract

Background: There are many ways clinicians measure pain intensity in patients with musculoskeletal conditions, such as a Whiplash Associated Disorder.

Objective: To assess if some ratings of pain are scored systematically higher than others and whether this pattern changes with increasing symptom duration (acute and chronic WAD)

Methods: This study involves secondary analysis of data collected as part of three large studies (n= 361 at baseline) conducted in Sydney and Brisbane, Australia. Patient reported pain questions and scales were extracted from the existing data set. In each cohort, at each follow - up time point, mean scores for all the pain measures were converted to a 0-100 point by simple multiplication and plotted along with their 95% confidence intervals. Mean scores were compared by an analysis of variance (ANOVA).

Results: Mean pain scores from the acute cohort (Study 1) showed a clear pattern with the highest scores coming from the SF-36 Bodily Pain question, second the pain intensity item of the Neck Disability Index (NDI), and third the VAS scores ($p< 0.05$). Mean pain scores from the chronic patients in Study 2 showed large differences between some measures. The highest ratings again came from the SF-36 and the lowest from the NDI. Scores from the Functional Rating Index (FRI) and Numeric Rating Scale (NRS) felt between these two measurements but did not appear to be different from one another ($p > 0.05$)

Mean pain scores from the chronic patients in Study 3 showed a similar pattern in that SF-36 and NDI item scores were again the highest and lowest respectively. There was however a smaller, but not significant difference between the mean scores from all the measures in this cohort (14 weeks follow – up, $p = 0.145$, 6 months follow – up, $p = 0.127$)

Conclusion: Pain ratings from different measurement instruments reveal different outcomes.

The passage of time (after injury) does not appear to significantly influence the way pain is rated, the pattern of scores from the different measures is fairly consistent, no matter how long patients have had their symptoms.

1. Introduction

1.1 Background

Whiplash Associated Disorder (WAD) is a common musculoskeletal condition which typically begins after a rear- end motor vehicle accident where acceleration-deceleration energy is transferred to the neck[1]. The most common symptom of WAD, is neck pain (90-100%), other symptoms include decreased neck mobility, headaches, pain in arm/shoulder, dizziness, and several social and psychological complaints[2]. The global incidence and prevalence rates of WAD widely vary between countries and settings, from 16 (New Zealand) to 70 (Canada) per 100.000 residents per year. These numbers are based on the insurance administration data[3]. In the Netherlands, an estimation of 94 to 188 per 100.000 is made based on the statistics of accidents that could be related to a whiplash trauma (accidents reported as a rear of end accident[4]. A distinction can be made between 'acute WAD' (usually described as < 30 days) and 'chronic WAD' (> 3 months)[5].

1.2 Measuring Pain Intensity

In general, there many ways researchers and clinicians measure pain intensity. Although the Visual Analog Scale and Numerical Rating Scale are used most commonly in clinical research[6-9], various likert scales are also used often as part of larger, multi-dimensional outcome measures. These pain rating scales are 'subjective outcomes' or 'patient reported outcomes', because they measure perception of pain as experienced by the patient and often form the primary research outcome in musculoskeletal conditions. Although they purport to measure the sensory aspect of pain, it is well accepted that various psychological and cognitive variables also influence pain intensity measures.

There is a considerable heterogeneity in the way pain intensity outcomes are collected and reported in clinical research[10]. At the same time there is very little research into how different pain measures can be compared. Sources of heterogeneity include the words used in the question, descriptors on the scale, the number of scale points and the time period over which patients are asked to recall their pain. This can create issues of interpretability and comparability for readers of primary studies and in particular for researchers conducting systematic reviews. Of particular relevance to meta-analyses is the question of whether it is appropriate for researchers to rescale any pain measures available in the primary studies to a common base e.g. 0-100 points for the purposes of pooling pain outcomes[11]. There is very little research assessing whether different pain intensity measures yield comparable ratings in the same patients, in particular to our knowledge this has never been done in patients with whiplash associated disorders (WAD).

1.3 Pain intensity and WAD

Although the psychosocial aspect of WAD has received more research attention in recent years[12-13], the physiotherapy guidelines focus on patient's pain rating outcomes, measured with pain rating tools such as

the VAS and the NRS. It could be important for physiotherapists to be aware of differences in a patient's pain rating depending in the way the question is asked, which scale is used and if there is a difference between acute and chronic patients. If there is a difference in the way pain is rated in the same subjects, it is important for physiotherapists to be consistent with the selection of a pain measurement tool for patients with WAD. This could be important on an individual level when assessing a patient's progress and also for comparing patients with each other regarding their treatments and outcome, especially across multiple – physiotherapist practices.

1.4 Definition of the problem

There are several questions relevant to considering whether ratings collected from different pain scale are comparable. These include: (1) whether some ratings of pain are scored systematically higher than others (2) whether this pattern change with increasing symptom duration

The outcome of this study could be useful for two groups; (1) researchers conducting systematic reviews and (2) physiotherapists treating and evaluating patients with WAD.

2. Methods

2.1 Participants

This study involves secondary analysis of data collected as part of three large studies conducted in Sydney and Brisbane, Australia (Table1). Study 1 was a longitudinal cohort study investigating the prognosis of acute WAD, Studies 2 and 3 were RCT's testing the effectiveness of exercise programs in people with chronic WAD. The following inclusion criteria were common to all studies; neck pain due to a car accident, age between 18 and 65 and fluency in written and spoken English. Participants were excluded if cervical scans showed fracture or dislocation on cervical scan or they had a diagnosis of serious spinal pathology or major psychiatric illness.

The principle point of difference between the cohorts was with regard to the duration of symptoms on entry to the study. Participants in Study 1 were enrolled within 1 month of their car accident and were recruited from hospital emergency rooms, via newspaper advertisements and through referral from physiotherapy practices. Participants whose symptoms had persisted for greater than 3 months and less than 12 months (Study 2) and greater than 3 months but less than 5 years (Study3) made up the chronic cohorts. Participants in Studies 2 and 3 were recruited via newspaper advertisement and from the records of the third party insurance administrator (Motor Accidents Authority)

Table 1 Demographics and Illness characteristics total N=361 at baseline

	Study 1 Acute	Study2 Chronic	Study 3 Chronic
Age; years (SD)	42.00 (13.36)	43.27 (14.68)	43.72 (12.92)
Gender% female	69.23	66.42	64.52
Duration Symptoms: Days (SD)	19 (9)	285 (117)	456 (688)
Neck Disability Index: % (SD)	36.4 (17.3)	38.0 (13.2)	36.2 (15.9)
Pain at baseline (SD) VAS / NRS	36.4 (22.9)	52.6 (20.0)	51.9 (20.4)
Follow-up points	Baseline (n=104) 3/12 (n=91) 6/12 (n=86) 12/12 (n=89)	Baseline (n= 134) 6/52 (n=132) 12/12 (n=125)	Baseline (n= 123) 14/52 (n=99) 6/12 (n=68) 12/12 (n=21)

2.2 Measures

Assessments were carried out at baseline and at either two or three follow-up points in each study (Table 1). Data were collected in an assessment booklet containing various questionnaires and scales including; socio–demographic variables (e.g. age, gender), pain severity, psychological measures, and (functional) disability questionnaires. The pain related questions were pooled from the various questionnaires and items (Table 2).

Table 2 Pooled Ratings of pain with different scales and questionnaires extracted per study

Questionnaire	Question	Scale	study1	study 2	study 3
Neck Disability Index (NDI)	What is your pain intensity right now?	6 point likert scale	Y	Y	Y
Visual Analog Scale (VAS)	What is your average pain intensity in the past 24 hours?	10 cm visual analog scale	Y	N	N
SF 36	How much bodily pain did you have during the past 4 weeks?	6 point likert scale	Y	Y	Y
Functional Rating Index (FRI)	What is your pain intensity right now?	5 point likert scale	N	Y	N
Numeric Rating Scale 24/24 (NRS)	What is your average pain intensity in the last 24 hours?	11point box scale	N	Y	Y
NRS 1/52	What is your average pain intensity in the Last week?	11 point box scale	N	N	Y

Whiplash Disability Questionnaire (WDQ)	How much pain do you have today?	11 point NRS scale	N	N	Y

2.2.1The Neck Disability Index (NDI) is a 10 item pain intensity and daily activity questionnaire regarding daily limitations after cervical spine injury[14] Item 1 was extracted for this study, asking the patient to rate 'pain intensity right now' on a six point likert scale, 'none (0) to severe (5)'.

2.2.2 ***The Visual Analogue Scale*** **(VAS)** scale is a 10 cm horizontal line, with extremes marked 'no pain' (left) and the 'worst pain imaginable' (right)[15]. Patients were asked to mark the spot on the line that best represents their pain intensity over the last 24 hours.

2.2.3 ***SF 36*** is a health-related quality of life questionnaire which captures 36 questions divided into eight domains [16-17]. For this study, question 7 was used, asking 'how much bodily pain did you have during the past four weeks', and is rated on a 6-point likert scale ranging from 'none (1) to severe (6)'.

2.2.4 ***The Functional Rating Index*** **(FRI)** is a self reported instrument measuring the degree of disability after spinal injury consisting out of 10 items. For this study the item 'pain intensity right now' was extracted on a 5 point likert scale, ranging from '0 = no pain' to '5 worst possible pain'[18].

2.2.5. ***The Numerical Rating Scale*** **(NRS)** scale has the same terminal anchors as a VAS scale but consists of numbers from 0-10[19]. Patients were asked to circle the number which best represents their pain over the last 24 hours.

2.2.6 ***The Whiplash Disabilty Questionnaire*** **(WDQ)** is a modified version of the NDI with 13 items designed to evaluate WAD[20]. For this study the item 'how much pain do you have today' was extracted. Pain is scored on an 11-point NRS scale from 0 = no pain to 10 = worst pain imaginable.

2.3 Data Analysis

In each cohort, at each follow up time point mean scores for all the pain measures were converted to a 0-100 point by simple multiplication and plotted along with their 95% confidence intervals. Data were analysed using the SPSS statistical program. To compare the multiple means with each other, an analysis of variance (ANOVA) was performed.

3. Results

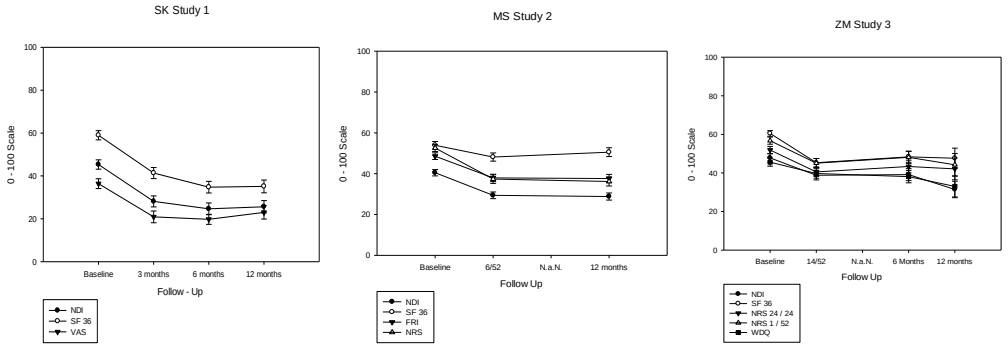
Mean pain scores (Table 1) from the acute cohort (Study 1) showed a clear pattern with the highest scores coming from the SF-36 Bodily Pain question, second NDI pain intensity item, and third theVAS scores. While mean pain levels felt over the study period, the pattern remained the same although differences

between the scores became smaller (Figure1a). Results from the ANOVA's indicated that there are significant differences between scores on the different measures at all time points (p = < 0.05)

Mean pain scores from the chronic patients in Study 2 showed large differences between some measures but not others. The highest ratings again come from the SF-36 and the lowest from the NDI, scores from the FRI item and NRS felt between these two but did not appear to be different from one another (Figure 1b).

Mean pain scores from the chronic patients in Study 3 showed a similar pattern in that SF-36 and NDI item scores were again the highest and lowest respectively. However, there was a smaller difference between the means from all the measures in this cohort (Figure 1c).

Fig. 1. Mean ratings (a) study 1 (NDI, SF36 and VAS). (b) Mean ratings study 2 (NDI, SF 36, FRI and NRS). (c) Mean ratings study 3 (NDI, SF 36, NRS 24/24 and 24/7, WDQ).



1.a 1.b 1. c

Table 3 Means and ANOVA

Study 1 Acute *** P value < 0.05 is significant**

			95% CI For mean			
Baseline	**Mean**	**SD**	**Lower bound**	**Upper bound**	**N (Val.)**	**P**
NDI	45.35	22.1	40.99	49.70	101	
SF 36	59.00	21.53	54.73	63.27	100	
VAS	36.40	22.94	31.85	40.95	100	
						0.000
3/12						
NDI	28.13	24.00	23.05	33.21	91	
SF 36	41.35	24.46	36.20	46.50	89	
VAS	20.90	24.24	15.43	26.36	78	
						0.000
6/12						
NDI	24.65	25.42	19.20	30.10	86	
SF 36	34.76	24.86	29.37	40.16	84	
VAS	19.76	21.14	15.11	24.40	82	
						0.000
12/12						
NDI	25.62	26.97	19.94	31.30	89	
SF 36	35.53	27.5	29.60	41.46	85	
VAS	23.00	26.01	16.80	29.20	70	
						0.008

Study 2 Chronic

			95% CI For mean			
Baseline	**Mean**	**SD**	**Lower bound**	**Upper bound**	**N(Val.)**	**P**
NDI	40.45	17.93	37.38	43.51	134	
SF 36	53.92	20.51	50.29	57.55	125	
NRS	52.61	20.00	49.19	56.03	134	
FRI	48.59	18.40	45.32	51.86	124	
						0.000
6/52						
NDI	29.39	18.73	26.17	32.62	132	
SF 36	48.17	21.92	44.20	52.13	120	
NRS	37.35	24.08	33.20	41.49	132	
FRI	37.92	19.98	34.31	41.53	120	
						0.000
12/12						
NDI	28.80	19.58	25.33	32.27	125	
SF 36	50.55	22.78	46.24	54.85	110	
NRS	36.16	24.62	31.80	40.52	125	
FRI	37.61	21.29	33.61	41.62	111	
						0.000

Study 3 Chronic

			95% CI For mean			
Baseline	**Mean**	**SD**	**Lower bound**	**Upper bound**	**N(Val.)**	**P**
NDI	47.80	24.11	43.50	52.11	124	
SF36	60.49	16.49	57.54	63.43	123	
NRS 24/24	51.88	20.42	48.14	55.62	117	
NRS 1/52	56.75	20.00	53.09	60.41	117	
WDQ	45.61	23.65	41.39	49.83	123	
						0.000
14/52						
NDI	38.79	24.04	33.99	43.58	99	
SF 36	45.31	21.79	40.94	49.67	98	
NRS 24/24	40.52	23.00	35.86	45.18	96	
NRS 1/52	45.10	23.62	40.32	49.89	96	
WDQ	39.69	24.09	34.84	44.55	97	
						0.145
6/12						
NDI	39.12	25.73	32.89	45.34	68	
SF 36	48.36	23.13	42.72	54.00	67	
NRS 24/24	43.33	26.09	36.76	49.90	63	
NRS 1/52	48.10	25.71	41.62	54.57	63	
WDQ	38.09	26.56	31.66	44.52	68	
						0.127
12/12						
NDI	31.43	19.57	22.52	40.34	21	
SF 36	47.62	24.06	36.67	58.57	21	
NRS 24/24	42.11	24.63	30.24	53.97	19	
NRS 1/52	44.21	25.67	31.84	56.58	19	
WDQ	33.00	24.52	21.53	44.47	20	
						0.000

4. Discussion

4.1 Patterns

Certain patterns are noticeable in both acute and chronic cohorts. The SF 36 scored the highest in each study compared to the other pain rating scores. In study 1, the VAS scored the lowest on each follow up moment. In study 2 the SF 36 had the highest pain rates, closely followed by the NRS at baseline. At the two follow – up moments in study 2, the NRS, FRI and the NDI pain ratings dropped, the SF 36 consistently stayed higher than the others at the same level. At study 3, the pain rating scores showed a more synchronous pattern at the 14 week and 6 month follow - up points, $p > 0.05$, as only two points of the study. The higher scoring pain items were the SF36 and the NRS (both pain intensity '24 hours' and 'last' week), the lower scoring items are the WDQ and the NDI.

4.2 A closer look

The SF 36 scored the highest at both the acute as the chronic study, the NDI scored relatively low compared to the SF 36. Both of the items are 6 – point likert scales and are extracted from a multi-item questionnaire. The way the question is asked to rate their pain do differ from each other; 'what is your pain intensity right now?' (NDI) and 'how much bodily pain did you have during the past 4 weeks?' (SF 36) Both the SF 36 bodily pain subscale[21] as the NDI are valid and reliable assessment tools[14] . This could suggest the way a patient is asked to rate their pain does makes a difference. Recall ratings of pain could cause bias, although there are studies advocating, recall may be valid for use in clinical research and physiotherapy practices[22.10].

The higher NRS pain rating scores from study 3 could support the theory of higher pain recall ratings. The question 'what is your average pain intensity in the last week?' scores higher at all follow – up points than 'what is your average pain intensity in the last 24 hours?' A contra against this theory is the fact that the FRI with the question 'what is your pain intensity right now?' in study 1 scored higher than the NRS 'what is your average pain intensity in the last 24 hours?'

The VAS and the NRS both belong to the three most used pain rating measurement tools (third is the Verbal Rating Scale)[19]. In this study the VAS in study 1 scored lower than the NRS in study 2 and 3. This could be the difference between the acute and chronic stage, yet the difference of the mean scores between the two scales is that great that further investigation in these scales in patients with WAD is desirable.

4.3. The research process and recommendations

This study was aimed upon the analysis of possible patterns or differences between pain scores. Pain ratings from different measurement instruments revealed different outcomes.

However, more research is required for the in between pain rating correlations and values in patients with WAD. In this study the three studies did not use the same pain measurement tools consistently, with exception of the NDI and the SF 36, which were used in all three studies. It would be preferable to conduct a second study, within the same subjects with different scales and questions.

5. Conclusion

The sigma Plots and ANOVA results show that there are differences between pain ratings at each time point, in each study, so pain ratings from different measurement instruments are not the same.
The passage of time (after injury) does not appear to significantly influence the way pain is rated, the pattern of scores from the different measures is fairly consistent, no matter how long patients have had their symptoms. Researchers and clinicians should treat the data of pain rating measurement tools with extreme cautiousness, both at the level of groups as at the individual level.

6. Acknowledgements

I would like to thank the George Institute for Global Health Sydney for giving me the opportunity to learn as much about research as possible in three months. A very special thanks to Steve Kamper, in my opinion the best supervisor a student could have. A very special thanks to Hans Heneweer, for introducing me to 'the George' and being a wonderful and great supervisor as well.

References

1. Spitzer WO, Skovron ML, Salmi LR, et al. Scientific monograph of the Quebec Task Force on Whiplash-Associated Disorders: redefining "whiplash" and its management. Spine. 1995;20:1S-73S.

2. Stovner LJ. The Status of the whiplash syndrome: a critical review based on a methodological approach. Spine 1996; 21:2735-46.

3. Skovron ML.Current concepts in prevention, diagnosis, and treatment of the cervical whiplash syndrome. Lippincott-Raven Publishers, 1998: 61-8.

4. Huijskens CG. Incidentie en preventive van het 'whiplash'-trauma. TNO Road- Vehicles Research Institute. 1994.

5. Gross AR, Hoving JL, Haines TA, et al. Manipulation and mobilization for mechanical neck disorders. Cochrane database Syst. Rev.2004; issue 1.

6. Ohnhaus EE, Adler R. Methodological problems in the measurement of pain: a comparison between the verbal rating scale and the visual analogue scale. Pain.1975.

7. Jensen MP, Karoly P, Braver S: The measurement of clinical pain intensity: a comparison of six methods. Pain.1986.27(1):117-126.

8 Williamson A ,Hoggart B. Pain: a review of three commonly used pain rating scales. Journal of Clinical Nursing. 2005. 14(7), 798-804.

9. Ogon M, Krismer M, Söllner W, et al. Chronic low back pain measurement with visual analogue scales in different settings. Pain. 1996. 64(3), 425-428.

10. Broderick J, Schwartz J,Schneider S, et al.Can End-of-Day Reports Replace Momentary Assessment of Pain and Fatigue? The Journal of Pain. 2009. 10(3), 274-281.

11. Kamper SJ, Rebbeck J, Maher C, et al.Course and prognostic factors of whiplash: A systematic review and meta-analysis. Pain. 2008.Volume 138, Issue 3 , Pages 617-629.

12. Ferarri R, Russell AS. Epidemiology of whiplash: an international dilemma. Ann Rheum Diss.1999; 58: 1-5

13 Ratliff AHC. Whiplash injuries. Journal of Bone Joint Surgery. 1997.;79-B:517-9.

14. Vernon H. The Neck Disability Index: state-of-the-art. Journal of Manipulative Physiotherapy. 2008.;31:491-502.

15 Peters ML, Patijn JP,Lamé I,Pain Assessment in Younger and Older Pain Patients: Psychometric Properties and Patient Preference of Five Commonly Used Measures of Pain Intensity. Pain Medicine.2007.

16. Litcher-Kelly L, Martino SA, Broderick JE, et al. A Systematic Review of Measures Used to Assess Chronic Musculoskeletal Pain in Clinical and Randomized Controlled Clinical Trials. The Journal of Pain,2007 8(12), 906-913.

17 Ware J, Dewey KM. How to score Version 2 of the Sf-36 Health Survey. Quality Metric Incorporated. 2000.

18 Feise R, Menke M. Functional Rating Index: A New Valid and Reliable Instrument to Measure the Magnitude of clinical Change in Spinal Conditions Spine. 2001, Volume 26 - Issue 1 - pp 78-87

19 Dijkers M. Comparing Quantification of Pain Severity by Verbal Rating and Numeric Rating Scales. Mount Sinai School of Medicine, New York, New York. 2010.

20. Willis C, Niere KR, Hoving JL, et al. Reproducibility and responsiveness of the Whiplash Disability Questionnaire. School of Physiotherapy, La Trobe University, Melbourne, Vic. 3086, Australia .2004.

21. Von Korff M, Jensen MP, Karoly P, Spine. 2000.Volume 25, number 24.

22 Bolton JE, Humphreys BK, van Hedel HJ. Validity of weekly recall ratings of average pain intensity in neck pain patients. Journal of Manipulative Physiotherapy. 2010. Oct;33(8):612-7.