Inconsistency and ceiling effect in repeated measures of labor pain using VAS

ศศิธร พุมดวง[1]

Abstract:

Inconsistency and ceiling effect in repeated measures of labor pain using VAS
Phumdoung S.
Obstetric Gynecological Nursing and Midwifery Department,
Faculty of Nursing, Prince of Songkla University, Hat Yai, Songkhla, 90112, Thailand
Songkla Med J 2004;22(3):155-161

To examine the inconsistency and ceiling effect when using the visual analogue scale (VAS) in measuring labor pain, data from four repeated measures of the sensation of labor pain in the early active phase from a clinical trial were re-analyzed. That study examined the effect of music on the labor pain experienced by 55 primiparas, while 55 control cases underwent labor under normal conditions. Both groups were balanced for demographic and obstetric data and other pain confounding factors. The re-analysis shows a 5 to 13% inconsistency in both groups in reported pain between the perception of whether pain was changing or not compared to their previous rating on VAS. The control group had more cases of ceiling effect in their reports of labor pain than the music group but analysis with and without the 13 cases of ceiling effect showed similar results that the music group experienced less pain than the control group. In the analysis without ceiling effect cases, the homogeneity of variance improved. This study supports the premise that a strong design and large sample size helps to make the statistics robust to the violation assumption of homogeneity of variance.

Key words: *pain, labor pain, ceiling effect, inconsistency, repeated measures, VAS*

บทคัดย่อ:

การศึกษาครั้งนี้ศึกษาเกี่ยวกับความไม่สอดคล้อง (inconsistency) ของการรายงานการเปลี่ยนแปลงความปวดที่ติดต่อกัน 4 ครั้ง กับการรายงานความปวดโดยการใช้ Visual Analogue Scale (VAS) และการเกิดปัญหาเพดาน (ceiling effect) ของการรายงาน ความปวดของมารดาในระยะรอคลอด   โดยวิเคราะห์ข้อมูลจากการศึกษาเชิงทดลองที่ศึกษาผลของดนตรีต่อความปวดในมารดา ครรภ์แรกในระยะที่ 1 ของการคลอด ซึ่งมารดากลุ่มที่ได้รับดนตรีและกลุ่มควบคุมมีจำนวนกลุ่มละ 55 คน ทั้ง 2 กลุ่ม มีความเหมือนกัน ในปัจจัยด้านประชากรและด้านสูติกรรม รวมถึงปัจจัยภายนอกที่จะมีผลต่อความปวด   จากการวิเคราะห์ข้อมูลพบว่า ร้อยละ 5-13 ของมารดาทั้ง 2 กลุ่ม   รายงานความปวดโดย VAS ไม่ไปในทิศทางเดียวกัน (inconsistency) กับการรับรู้การเปลี่ยนแปลงของความปวด จริงเมื่อเปรียบเทียบกับระดับความปวดชั่วโมงที่ผ่านมา ส่วนการเกิดปัญหาเพดานของการรายงานความปวด พบในมารดากลุ่มควบคุม มากกว่ามารดากลุ่มที่ได้รับดนตรี   อย่างไรก็ตามจากการวิเคราะห์เปรียบเทียบความปวดของทั้ง 2 กลุ่ม ซึ่งวิเคราะห์ข้อมูลโดยรวม และไม่รวมมารดาจำนวน 13 ราย ที่รายงานความปวดแบบ ceiling effect มาวิเคราะห์ พบว่ามารดาที่ได้รับดนตรีมีความปวดน้อยกว่า กลุ่มที่ไม่ได้รับดนตรี และผลทางสถิติทั้ง 2 ครั้ง ไม่มีความแตกต่างกัน และการไม่นำข้อมูลของมารดารายที่มี ceiling effect มาวิเคราะห์ ทำให้ค่าความแปรปรวนของทั้ง 2 กลุ่มแตกต่างกันน้อยลง ซึ่งพิสูจน์ได้ว่าการออกแบบการวิจัยที่ดีและกลุ่มตัวอย่างขนาดใหญ่ช่วยให้ ผลการวิจัยน่าเชื่อถือ   ถึงแม้ข้อตกลงเบื้องต้นเรื่องความแปรปรวนของข้อมูลจะไม่เป็นไปตามข้อตกลงเบื้องต้นอย่างสมบูรณ์ก็ตาม

คำสำคัญ: ความปวด, การเจ็บครรภ์, ปัญหาเพดาน, ความไม่เที่ยง, การวัดซ้ำ, VAS

Introduction

Visual analogue scales (VAS) are single-item measures that have been used to study postoperative pain,[1-3] temporomandibular joint pain[4] and labor pain.[5,6] They have shown a high correlation (r = .70) when used in reporting the pain of patients with back pain and in the physician's ratings of patient improvement[7] supporting the concurrent validity. The VAS is quick and easy to use and does not require any reading ability;[8] from explanation time to sample is less than five minutes, the rating (scoring) time is less than two minutes[9] and only 5% of people with chronic pain report pain incorrectly.[10] The VAS is more sensitive to change in pain than numerical or simple descriptive scales.[11-16] The mean difference between current and preceding VAS scores has been shown to be small, 13 mm (95% CI, 10 to 17 mm),[17] supporting the sensitivity of VAS.

The main disadvantage of VAS is the "ceiling effect" that occurs when people rate pain at the top of the scale, but then their pain increases and their rating cannot be changed or rated accurately,[18] thus, an invalid rating of pain will be reported. Repeated measures of transition scores using VAS have raised concerns about such invalidity.[19] In repeated measures of temporomandibular joint pain, in one study the error variance was 31% for VAS while the error variance for the McGill Pain Questionnaire (MPQ) was only 24%.[4] The use of VAS in repeated measures can result in inconsistency and highlight the ceiling effect. Inconsistencies in the data and the ceiling effect can both lead to questionable results.

Few studies have reported on the inconsistency and no studies have reported on the ceiling effect in the use of VAS in increasing pain. When acute pain was measured repeatedly in 48 patients with trauma level I, there were some inconsistencies in the rating and experience of changing pain.[17] They measured pain using a dual horizontal 100 mm VAS with descriptors "least possible pain" and "worst possible pain". Pain was measured every 20 minutes for 2 hours. From 248 data point comparisons, 139 were categorized as about the same pain, 51 as increasing pain, and 52 as decreasing pain. The inconsistency of the reported experience of pain and the pain rating on the VAS was 15 (6%) of 248 data points. The researchers also determined that the ceiling effect occurred when subjects experienced severe pain, and this was reflected as a skewing of the distribution, but they did not provide the exact number of cases who reached the ceiling effect. In another study (N = 25) of repeated measurements

of temporomandibular joint pain, the researchers, who used VAS to repeat the measurements 4 times, showed that the smallest detectable difference decreased by increasing the number of repetitions.[4] They suggested that this showed an inconsistency in the measurement results. The researchers did not report the number of inconsistencies or the ceiling effect of the reported pain.

Despite the disadvantages outlined, VAS is still a suitable instrument for measuring labor pain, especially when women experience severe pain, as it is difficult to use questionnaires that are long and complicated, such as the MPQ of Melzack.[20] However, as labor pain generally increases over time,[21] there is the possibility that a ceiling effect will occur, but there are no published reports for labor pain studies on the ceiling effect or on the inconsistency of repeated measures of pain. Thus, it is important to study the inconsistencies and ceiling effect of using VAS in measuring increasing pain. This study is a secondary data analysis from a randomized controlled trial with an experimental pretest-posttest repeated measures design in testing the effect of soft music without lyrics on the sensation of labor pain during the first three hours of the active phase of labor.[5] In the original study, primiparas were studied and, as they did not have previous experience of labor pain, the ceiling effect may have occurred because primiparas often report their experience of pain at a high level, leaving no room[22] for a later rating of higher pain. The purposes of this study were to investigate the number of women who rated their sensation of pain levels inconsistently between the experiencing changing of pain and pain rating on the VAS and who experienced a ceiling effect in rating pain. In addition, this study tested whether the ceiling effect in rating pain affected the results of the original study or not.

Material and methods

The sample was 110 women who completed the three hour study; each group (a music group and a control group) had 55 primiparas. The original study controlled for the effect of confounding variables on pain by selecting the inclusion and exclusion criteria, and recording receiving oxytocin, presence of relative or nurse at the bedside and assistance from nurse or relative in coping with pain. The inclusion criteria were married women, aged 20-30 years, primiparas, single fetus, those who received antenatal care from the second trimester, and who were in the latent phase of labor no more than ten hours. The fetal criteria were normal fetal heart rate, cephalic presentation and 38-42 weeks of gestation with an estimated fetal weight of 2,500-4,000 grams. The excluded women were those who had received analgesic medication before entering the study, labor induced before entering the study by receiving oxytocin or artificial rupture of membranes, membranes ruptured more than 20 hours before recruitment, a history of psychiatric problems or on major antipsychotic medications, difficulty hearing the spoken word, HIV, infections or asthma. In addition, minimized randomization was used to assign the subject to each group to balance for age, education, time in the latent phase, painful menstruation, and spontaneous rupture of membranes. Minimized randomization was used to control and balance potentially confounding variables and balance the number of the samples.[23]

In this study, sensation of labor pain was measured using the 100 mm horizontal VAS. The scale ranged from "no sensation" to "the most sensation imaginable". Sensation of labor pain was described as the sensation or amount of pain the subjects felt in the abdomen and back from a uterine contraction that just ended. Pain was measured four times when the women were not having uterine contractions. First, at the start of the study just before the treatment period when the cervix was 3-4 cm dilated and uterine contractions lasted 30-60 seconds; then every hour during the three hour study. To examine the consistency of the VAS pain scores, the subjects were asked at the end of the first, second, and third hours whether or not their sensation of pain had increased, decreased, or had not changed from the last time it was measured. Each response was compared with the previous and current VAS score and was scored according to whether they were in the same direction. If the direction of change in the current VAS score compared to the previous VAS score was parallel (in the same direction) to the rating of change (increased, decreased or did not change), it was coded as consistent. If the direction was not parallel, it was considered inconsistent.

Cases of ceiling effect were identified when the VAS score was rated from 95 to 100 mm, and that person then reported that her pain had increased. All individual scores from 95 to 100 mm were examined because the mean increase in sensation of pain at each hour in the control group was about 10 mm; so, increases that were half of the mean increase (5 mm) were identified as having insufficient room at the end of the scale for the next pain rating. If the woman experienced increased pain in comparison to the previous rating of 95-100 VAS, that case was counted as having a ceiling effect.

Results

The average age of both groups (music and control) was 24.30 years and SD = 3.07 years. The majority of the subjects had high school education or higher (n = 64, 58%), were Buddhist (n = 94, 85.5%) and were at 38-40 weeks' gestation (n = 95, 86%).

As the contractions increased during labor, the distribution of sensation of pain scores was characterized by tightly clustered high scores (skewness) at the second and third posttests, especially in the control group (Figure 1 and Table 1). In the control group, the ceiling effect occurred in the second hour for 4 (3.6%) women and in the third hour for 8 (7.2%) women. The ceiling effect occurred in the music group only in the third hour for 5 (4.5%) women.

Over the three hourly measures in both groups there was an increase in inconsistency between the direction of VAS changes and the reports of sensation of pain changes (Table 2). In the control group, in the first hour, 52 (94.5%) women rated the change in their pain consistently but 3 (5.5%) did not; in the second hour, 46 (83.6%) rated the change consistently but 5 (9.1%) did not and 4 more (7.3%) rated consistently but were identified as having a ceiling effect; and in the third hour, 40 (72.7%) rated consistently but 7 (12.7%) did not and 8 more (14.5%) rated consistently but had a ceiling effect. All those who had a ceiling effect also reported an increase in their sensation indicating the consistency of their response. In the music group there was also inconsistent reporting of pain: in both the first and second hours there were 3 (5.5%) women and in the third hour 4 (7.3%) women who rated their changes in pain inconsistently with the VAS score. The 5 (4.5%) women who had reached the ceiling effect were consistent with reporting increased pain.
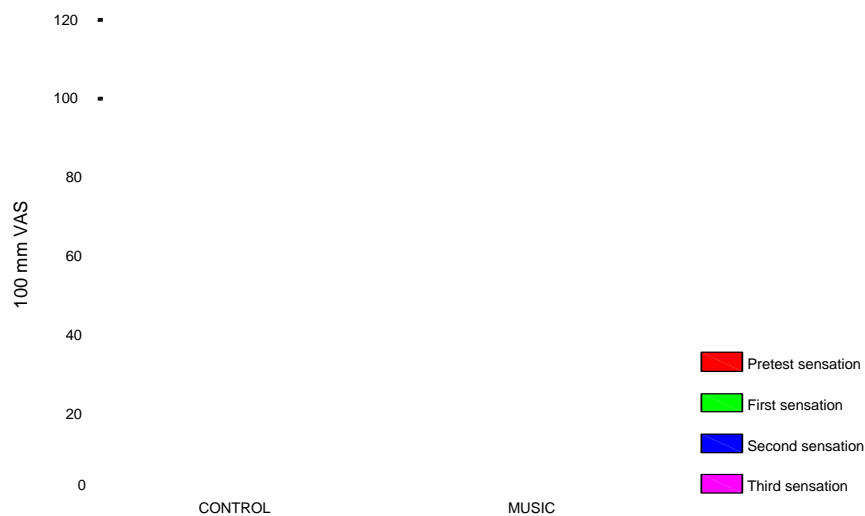

Figure 1 The distribution of sensation of pain scores at each data point by groups

Table 1 Means and standard deviations of pain scores at each data point

Data Points	Pain scores					
	Total (N = 110)		Control (n = 55)		Music (n = 55)	
	Mean	SD	Mean	SD	Mean	SD
Pretest	59.12	15.73	59.09	14.84	59.15	16.70
1st hr posttest	67.84	16.01	70.95	15.68	64.73	15.87
2nd hr posttest	75.29	18.15	80.05	14.42	70.53	20.27
3rd hr posttest	83.21	16.34	88.29	11.67	78.13	18.72

Table 2 Pain scores and their inconsistency and differences of each pair for the four data points

Group	Inconsistency pain scores				
	Pretest	1st hr	2nd hr	3rd hr	Difference (mm)*
control	32	31			−1
	43	42			−1
	56	54			−2
		51	51		0
		68	60		−8
		75	74		−1
		91	90		−1
		92	91		−1
			65	64	−1
			87	82	−5
			90	74	−16
			90	86	−6
			90	88	−2
			91	91	0
			94	85	−9
Music					
	63	58			−5
	84	68			−16
	94	90			−4
		36	34		−2
		61	59		−2
		76	85		+9
			59	59	0
			81	81	0
			90	84	−6
			90	86	−4

* − = Women experienced increased pain but they rated pain on VAS lower than the last hour pain. + = Women experienced lowered pain but they rated pain higher than the last hour pain.

Whether the cases of ceiling effect in both groups were included or not, the results of pain differences during the first three hours between both groups were still the same. While controlling for the pretests, one-way repeated measures ANCOVA showed that the music group had significantly less pain than the control group over the three-hourly posttests, F $(1, 107) = 18.96$, $p < .001$, Partial Eta squared (effect size) = .15, and a power of .99. However, there was a violation in the assumption of homogeneity of variance because the Box test was significant ($F_{max} = 3.75$, $p < .01$) but the ANCOVA was considered robust to unequal variance because the groups were randomized and equal in size.[24, 25] This was supported by testing with and without the 13 cases that were considered ceiling effect for pain in the second and the third hours, as similar results were found, $F (1,94) = 17.68$, $p < .001$, Partial Eta squared = .16, and a power of .99; Box's M (F_{max}) was 2.85 (when cases of ceiling effect were not included in the analysis) and significant but considered lowering the violation of homogeneity of variance.

Discussion

Inconsistency in the reporting of pain was evident, 3-7%, especially in the second and third hours when pain increased more in the control group. This inconsistency of reporting may have occurred because those who rated pain inconsistently experienced only small pain changes resulting in difficulty in reporting it accurately: while pain difference scores ranged from 0 to 16, the mode of pain difference was

1 mm (6 cases, 24%) (Table 2). The inconsistency in reported pain is similar to the results of Todd and colleague[17] who found 5% inconsistency in reporting acute pain from accidents.

The ceiling effect, which increased in the second and third hours in both control and music groups, was not totally unexpected because the women were first time mothers who had never experienced labor pain before, a group which has been known to report pain as very high. When their pain increased, and there was little room for them to rate the pain as higher, the ceiling effect, a limitation in using the VAS in repeated measures of increasing pain, was encountered. The number of occurrences of this ceiling effect on rating pain was higher in the control group than in the music group because women in the control group experienced higher pain than the women in the music group.

When comparing the difference of the pain between the music and the control groups, either with or without cases of ceiling effect, and controlling for the pretest pain scores, the results were still the same. This supports the fact that a strong design, equal and large sample size results in statistics that are robust to violation of the assumption of homogeneity of variance.[24, 25] The large sample size also resulted in high observed power for intervention testing.[26] These results suggest that researchers can use VAS in repeated measures of increasing pain.

Conclusion

The use of VAS in repeated measures of labor pain in primiparas results in some inconsistency and the ceiling effect. However, with a strong design and large sample size, the results of the study can be trusted even though there is a violation of the assumption of homogeneity of variance.

Acknowledgement

References

1. Good M, Stanton-Hicks M, Grass JA, Anderson G, Choi C, Schoolmeesters L, et al. Relief of postoperative pain with jaw relaxation, music and their combination. Pain 1999;81:163-72.

2. Good M, Stanton-Hicks M, Grass JA, Anderson GC, Lai H, Roykulcharoen V, et al. Relaxation and music to reduce postsurgical pain. J Adv Nurs 2001;33:208-15.

3. Good M, Stiller C, Zauszniewski JA, Anderson GC, Stanton-Hicks M, Grass JA, et al. Sensation and distress of pain scales: reliability, validity, and sensitivity. J Nurs Meas 2001;9:210-37.

4. Kropmans TJB, Dijkstra PU, Stegenga B, Stewart R, de Bont LGM. Repeated assessment of temporomandibular joint pain: Reasoned decision-making with use of unidimensional and multidimensional pain scales. Clin J Pain 2002;18:107-15.

5. Phumdoung S, Good M. Music reduces labor pain. Pain Manage Nurs 2003;4: 54-61.

6. Ryding AE, Cox DN. Psychosocial predictors of labor pain. Pain 1985;22:309-15.

7. Price DD, McGrath PA, Rafii A, Buckingham B. The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. Pain 1983; 17:45-56.

8. Lee KA, Kieckhefer GM. Active coping reduces reports of pain from childbirth. J Consult Clin Psychol 1989; 57:365-71.

9. McGuire DB. The measurement of clinical pain. Nurs Res 1984;33:152-6.

10. Jensen MP, Karoly P, Braver S. The measurement of clinical pain intensity: a comparison of six methods. Pain 1986;27:117-26.

11. Briggs M, Close JS. A descriptive study of the use of visual analogue scales and verbal rating scales for the

assessment of postoperative pain in orthopedic patients. J Pain Symptom Manage 1999;18:438-46.

12. Huskisson C. Measurement of pain. Lancet 1974;2: 1127-31.

13. Joyce CRB, Zutshi DW, Hrubes V, Mason RM. Comparison of fixed interval and visual analogue scales for rating chronic pain. Eur J Clin Pharmacol 1975;8:415-20.

14. Ohnhaus E, Adler R. Methodological problems in the measurement of pain: a comparison between the verbal rating scale and the visual analogue scale. Pain 1975; 1:379-84.

15. Scott J, Huskisson EC. Graphic representation of pain. Pain 1976;2:175-84.

16. Sriwatanakul K, Kelvie W, Lasagna L, Calimlim JF, Weis OF, Mehta G. Studies with different types of visual analogue scales for measurement of pain. Clin Pharmacol Ther 1983;34:234-9.

17. Todd KH, Funk KG, Funk JP, Bonacci R. Clinical significance of reported changes in pain severity. Ann Emerg Med 1996;27:485-9.

18. Huskisson EC. Visual analogue scales. In: Melzack R, editor. Pain measurement and assessment. New York: Raven Press, 1983:33-37.

19. Guyatt GH, Norman GR, Juniper EF, Griffith LE. A critical look at transition ratings. J Crit Epidemiol 2002; 55:900-8.

20. Melzack R. The short-form McGill Pain Questionnaire. Pain 1987;30:191-7.

21. Leventhal EA, Leventhal H, Shacham S, Eastering DV. Active coping reduces reports of pain from children. J Consult Clin Psychol 1989;57:365-71.

22. Nunnally JC, Bernstein IH. Psychometric theory. 3rd ed. New York: McGraw-Hill; 1994;542-94.

23. Zeller R, Good M, Anderson GC, Zeller DL. Strengthening experimental design by balancing potentially confounding variables across treatment groups. Nurs Res 1997;46:345-9.

24. Green SB, Salkind NJ, Akey TM. Using SPSS for windows: analyzing and understanding data. 2nd ed. Upper-Saddle River: Prentice Hall; 2000;208-17.

25. Stevens J. Applied multivariate statistics for the social sciences. 3rd ed. Hillsdale: Lawrence Erlbaum Assoicates; 1996;237-60.

26. Cohen J. Statistical power analysis for the behavioral science. 2nd ed. Hillsdale: Lawrence Erlbaum Associates; 1988;273-406.