



Topical review

How to analyze the Visual Analogue Scale: Myths, truths and clinical relevance

Gillian Z. Heller^{a,*}, Maurizio Manuguerra^a, Roberta Chow^b^a Department of Statistics, Faculty of Science and Engineering, Macquarie University, NSW 2109, Australia^b Brain and Mind Centre, 94 Mallett Street, Camperdown, NSW 2050, Australia

HIGHLIGHTS

- Use of the Visual Analogue Scale (VAS) for pain measurement has grown exponentially.
- Analyses of VAS using suboptimal statistical methods, are frequently reported.
- A new method is demonstrated to have superior power to commonly used methods.

ARTICLE INFO

Article history:

Received 16 March 2016
 Received in revised form 7 June 2016
 Accepted 30 June 2016
 Available online 27 July 2016

Keywords:

Visual Analogue Scale
 Level of measurement
 Ordinal measure
 Ordinal regression analysis
 Wilcoxon test
 Linear regression

ABSTRACT

Background and aims: The Visual Analogue Scale (VAS) is a popular tool for the measurement of pain. A variety of statistical methods are employed for its analysis as an outcome measure, not all of them optimal or appropriate. An issue which has attracted much discussion in the literature is whether VAS is at a ratio or ordinal level of measurement. This decision has an influence on the appropriate method of analysis. The aim of this article is to provide an overview of current practice in the analysis of VAS scores, to propose a method of analysis which avoids the shortcomings of more traditional approaches, and to provide best practice recommendations for the analysis of VAS scores.

Methods: We report on the current usage of statistical methods, which fall broadly into two categories: those that assume a probability distribution for VAS, and those that do not. We give an overview of these methods, and propose continuous ordinal regression, an extension of current ordinal regression methodology, which is appropriate for VAS at an ordinal level of measurement. We demonstrate the analysis of a published data set using a variety of methods, and use simulation to compare the power of the various methods to detect treatment differences, in differing pain situations.

Results: We demonstrate that continuous ordinal regression provides the most powerful statistical analysis under a variety of conditions.

Conclusions and Implications: We recommend that in the situation in which no covariates besides treatment group are included in the analysis, distribution-free methods (Wilcoxon, Mann–Whitney) be used, as their power is indistinguishable from that of the proposed method. In the situation in which there are covariates which affect VAS, the proposed method is optimal. However, in this case, if the VAS scores are not concentrated around either extreme of the scale, normal-distribution methods (*t*-test, linear regression) are almost as powerful, and are recommended as a pragmatic choice. In the case of small sample size and VAS skewed to either extreme of the scale, the proposed method has vastly superior power to other methods.

© 2016 Scandinavian Association for the Study of Pain. Published by Elsevier B.V. All rights reserved.

Contents

1. Introduction.....	68
2. Methods.....	68
2.1. Level of measurement of VAS.....	68

DOI of refers to article: <http://dx.doi.org/10.1016/j.sjpain.2016.08.005>.

* Corresponding author. Tel.: +61 2 98508541.

E-mail address: gillian.heller@mq.edu.au (G.Z. Heller).

<http://dx.doi.org/10.1016/j.sjpain.2016.06.012>

1877-8860/© 2016 Scandinavian Association for the Study of Pain. Published by Elsevier B.V. All rights reserved.

2.2.	Methods of analysis in use	69
2.2.1.	Distribution-based methods	69
2.2.2.	Distribution-free methods	69
2.2.3.	Categorical methods	69
2.2.4.	Ordinal regression analysis	70
2.2.5.	Continuous ordinal regression	70
2.3.	Laser therapy for chronic knee pain study	70
2.4.	Simulation study	70
2.4.1.	Scenario 1	71
2.4.2.	Scenario 2	71
3.	Results	71
3.1.	Laser therapy for chronic knee pain study	71
3.2.	Simulation study	71
4.	Discussion	72
	Ethical issues	74
	Funding	74
	Acknowledgement	74
	References	74

1. Introduction

The Visual Analogue Scale (VAS) has been in use for the measurement of intangible quantities such as pain, quality of life and anxiety since the 1920s [1]. It consists of a line usually 100 mm in length, with anchor descriptors such as (in the pain context) “no pain” and “worst pain imaginable”, depicted in Fig. 1. The patient makes a mark reflecting his or her perception, and the distance from the left endpoint to the mark is measured, in mm. The VAS was initially used in psychology for the measurement of mood disorders, and was used for the measurement of pain from the mid-1960s [2]. The scale can be horizontal or vertical. Alternatives to the VAS are the Verbal Rating Scale (VRS) which involves intermediate descriptors (e.g. “mild”, “moderate”, “severe”), and the Numeric Rating Scale (NRS) [3].

Pain is defined as “an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage” [4]. Breivik et al. [5], in a study that compared the VAS with a four-category VRS (VRS-4) and an 11-point NRS (NRS-11), found similar sensitivity (i.e. power to detect treatment differences) and strong agreement between the VAS and NRS-11, while the VRS-4 was found to be less sensitive than the VAS. Some difficulties with the VAS were reported by [6]: “. . . the Visual Analogue Scale has more practical difficulties than the Verbal Rating Scale or the Numerical Rating Scale”; Hawker et al. [7] and Breivik et al. [5] give a similar assessment; and Jensen et al. [8] found that the group that had the most difficulty with the VAS were older patients. In addition, Dexter and Chestnut [9] raise a caveat on the situation in which many patients report their pain at one of the extremes of the scale.

Despite these concerns, the VAS has become a very popular instrument for the measurement of pain in recent years. The results of a PubMed search using the search terms “pain” and “visual analogue scale” or “visual analog scale”, by calendar year from 1975 to 2014, are shown in Fig. 2. This reveals an exponential increase in the number of articles reporting the use of the VAS in the pain context over this period, from two in 1975 to 2181 in 2014.

There is no uniformity in the literature on statistical methods used for the analysis of VAS, and it is this issue that we seek to address in this review. The central problem is whether VAS is a ratio or ordinal variable; and given this, what the appropriate statistical methods for analysis are.



Fig. 1. The Visual Analogue Scale for pain.

2. Methods

2.1. Level of measurement of VAS

The VAS is usually 100 mm in length. When measured to 1-mm accuracy, this results in a 101-point scale: (0, 1, . . . , 100). In this article we will consider VAS to be reported in centimetres, i.e. on the scale 0–10. There is robust discussion in the literature on how best to categorize the level of VAS measurement. Stevens [10] introduced a typology of levels of measurement, viz. nominal, ordinal, interval and ratio. In the context of VAS, ratio measurement means that differences are meaningful, i.e. a difference between two VAS measurements of e.g. 2 represents the same pain difference at any part of the VAS scale. However in clinical practice there are important qualitative differences between a decrease of VAS from 8 to 6 from that of 3 to 1, though the “difference” in both cases is 2. This is reflected in the types of analgesic agents used where disabling and distressing pain scores of 7 and above may require opioid management compared with simple analgesics at the low end. This is clearly a difficulty, and calls into question the use of change scores in analyses. The issue is resolved by considering VAS scores to be at an ordinal level of measurement, where lower numerical values of VAS reflect less pain than higher levels but they cannot be regarded as absolute. This issue was considered by Philip [11], who

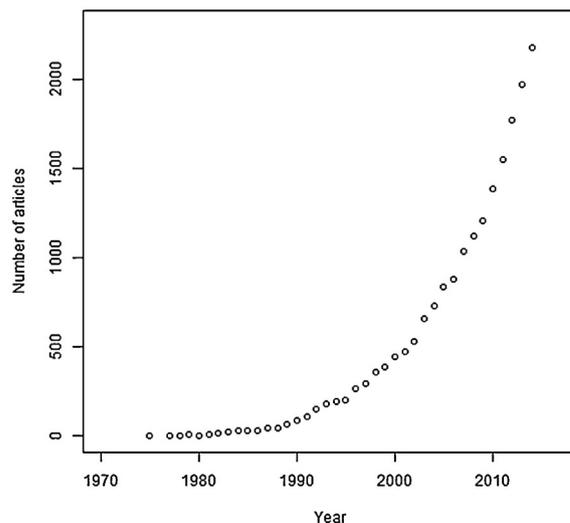


Fig. 2. Number of articles 1975–2014, PubMed search on pain and visual analogue/analog scale.

Table 1
Statistical methods used for analysis of VAS, December 2014.

Method	Number of articles
Distribution-based methods:	
Normal distribution	32
Non-normal distribution	1
Distribution-free methods	25
Categorical methods	4
Descriptive only	8
Unclear	10
Total	80

concluded that VAS “lies somewhere between ordinal and interval values”; Wewers and Lowe [12] favour the ordinal interpretation (“ipsative”); other authors [13–15] argue strongly in favour of the ordinal interpretation. The opposing viewpoint, i.e. that VAS should be treated as a ratio variable, is presented by Price et al. [16,17].

Whether one considers VAS to be ratio or ordinal dictates which statistical methods are appropriate for analysis. If VAS is ratio, distribution-based methods are applicable. If, however, the level of measurement is ordinal, distribution-based methods are inappropriate and either distribution-free (nonparametric) methods, or ordinal regression models, should be used. The use of distribution-based methods on ordinal data is strongly criticized by several authors [18,12–15]. Svensson [13] states that “There is now an increased awareness that different statistical approaches are appropriate for data on different measurement levels . . . However, tradition and the preference for established statistical methods often have a delaying effect on the acceptance of new statistical approaches in applied sciences. Inappropriate statistical methods . . . are still common.”

2.2. Methods of analysis in use

To inform us on the current usage of statistical methods for the analysis of pain as measured by the VAS, we conducted a PubMed search on the terms “pain” and “visual analogue scale” or “visual analog scale”, restricted to December 2014. A total of 191 articles were found, of which we randomly selected 100. Of these, 20 were not relevant for the following reasons: two were in foreign languages; nine were meta-analyses, case reports, systematic reviews or protocols; nine were not relevant for other reasons. This left 80 articles which used VAS for pain measurement and analyzed it as an outcome measure. Table 1 lists the statistical methods used in these articles.

2.2.1. Distribution-based methods

When VAS is treated as the outcome measure in an analysis, we refer to it as the response variable. Distribution-based methods are statistical methods which assume a probability distribution for the response variable, and are appropriate for response variables which are at a ratio level of measurement. These methods can be broadly classified into two groups, depending on which distribution is assumed: (a) the normal distribution and (b) non-normal distributions. The normal (or Gaussian) distribution has the features of being symmetric, bell-shaped and unbounded, i.e. a normally distributed random variable can assume any value from minus infinity to infinity.

2.2.1.1. Methods based on the normal distribution. Methods which assume the normal distribution for the response variable include the Student’s *t*-test, analysis of variance (ANOVA), repeated measures ANOVA and linear regression. As the VAS is bounded between 0 and 10 cm, it is theoretically impossible for the VAS to have the normal distribution, which is unbounded. However in practice we may find VAS scores to be approximately normally distributed, as long

Table 2
Distribution-free tests and their normal-distribution equivalents.

Distribution-free test	Normal-distribution equivalent
Wilcoxon signed-rank	Paired <i>t</i> -test
Mann–Whitney <i>U</i> /Wilcoxon rank sum	Two-sample <i>t</i> -test
Kruskal–Wallis	One-way ANOVA
Friedman	Two-way ANOVA

as their distribution is not concentrated near the extremes of the scale. Of the 80 eligible articles that we examined, 32 used methods based on the normal distribution. Of these, only six reported statistical testing for normality of the VAS scores; in the remaining 26 articles no comment was made on the validity of the normality assumption. Only one article reported applying a transformation to VAS scores to achieve normality. For reasonably large sample sizes and distributions not too far from the normal, the consequences of non-normality on statistical inference are not severe; but if the sample size is small to moderate (as is frequently the case in clinical studies) and the distribution severely skewed at one extreme (if most patients experience either no pain or extreme pain) in, for example, a population of patients in the 24 h post-operative period, the statistical analysis may deliver misleading results, i.e. biased estimates and incorrect *p*-values.

2.2.1.2. Methods based on other distributions. The Generalized Linear Model (GLM) was introduced in the 1980s and became popular in the 1990s [19]. GLMs broadened statistical modelling to accommodate response variables having a wide class of distributions other than the normal. In particular, the distribution that is relevant to VAS is the beta distribution, which is bounded to a finite interval and may be symmetric or skewed to either extreme. Some examples are shown in Fig. 3: in a palliative care population, pain levels will be aggregated to the “high” end (Fig. 3(c)); in a healthy population, most of the population will have no pain and will be concentrated at the lower end (Fig. 3(a)). It is possible to employ GLM methodology for the modelling of VAS, using the beta distribution for the response variable. As this method implies a ratio level of measurement and is not widely used in the analysis of pain outcomes, we do not consider it further in this article.

2.2.2. Distribution-free methods

Distribution-free (or nonparametric) methods do not assume any probability distribution for the response variable. Instead they are based on the orderings or ranks of the VAS scores, rather than the VAS scores themselves. Therefore they are appropriate for ordinal outcomes. Table 2 lists the commonly-used distribution-free tests, and their normal-distribution equivalents. If the assumption can be made that the distributions of the scores have the same shape across groups, then the null hypotheses being tested are that the medians of the groups are equal.

While distribution-free methods are a correct approach for the analysis of ordinal outcomes, they have the disadvantage that they cannot incorporate continuous covariates (such as blood pressure), or multiple categorical covariates (such as gender or disease type). This is overcome by using ordinal regression, which we introduce with discussion of the categorization of VAS.

2.2.3. Categorical methods

Another approach to analysis of VAS scores is to categorize it into a small number of levels. It is useful to conceptualize this categorization as in Fig. 3.

The horizontal axis is VAS, and its probability distribution is plotted on the vertical axis. We illustrate the VAS distribution for three hypothetical populations, with VAS categorized into five levels: none, mild, moderate, severe, and unbearable pain. Fig. 3(a)

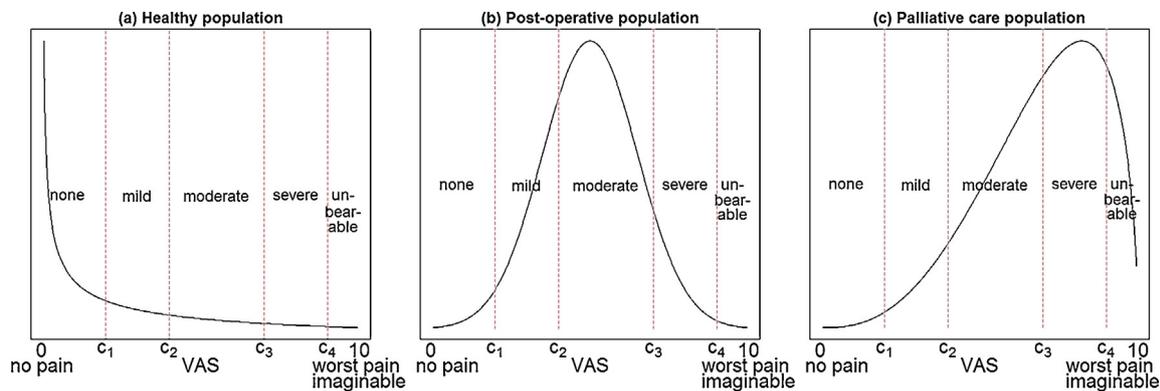


Fig. 3. Categorization of VAS to five levels: none, mild, moderate, severe, unbearable. Shown for three different patient populations: (a) healthy, (b) post-operative and (c) palliative care.

depicts what we would expect in a healthy population, with most subjects experiencing no pain. Fig. 3(b) is what we would expect in a postoperative population, in which most patients are experiencing moderate levels of pain. Fig. 3(c) shows a population with mostly very high levels of pain, for example patients in palliative care. We use the notation K for the number of categories (in this case $K=5$). The cutpoints c_1, c_2, \dots, c_{K-1} delineate the categories, and we can extend this to include the endpoints, with $c_0=0$ and $c_K=10$. Note that the choices of K , and the cutpoints c_1, c_2, \dots, c_{K-1} , are at the researcher's discretion and quite arbitrary.

The case $K=2$ leads to dichotomization (“pain”, “no pain”), which arises from the use of just one cutpoint c_1 . If only one categorical predictor (such as treatment/placebo group) is being used, the appropriate analysis method is the chi-square or Fisher's exact test; if more than one predictor, and/or continuous covariates (such as patient's age) are to be included in the statistical model, logistic regression is the correct method.

When $K>2$, the levels of the resulting categorical pain variable are ordinal in nature, and need to be treated as such in statistical analysis. In some cases (for example [20]) this has been analyzed using the Chi-square or Fisher's exact test. This approach ignores the ordering in the levels and would give the same result if, for example, “none” and “severe” were switched around. On the other hand, ordinal regression analysis is a method which respects the ordering of the categories. This is discussed in the next section.

2.2.4. Ordinal regression analysis

Ordinal regression analysis [21] provides a statistical framework for the analysis of ordinal response variables, and allows both categorical and continuous covariates to be included in the model. This represents a distinct advantage over distribution-free methods.

We consider the pain that the patient experiences as an intangible, continuous, unbounded variable which we cannot observe or measure. What we do observe is the VAS score, and there is a one-to-one relationship between these two quantities. The left panel of Fig. 4 shows an example of this relationship. As in Fig. 3, the horizontal axis is the observed VAS score. The vertical axis is the pain experience. The α_j s are the cutpoints on the scale of pain, corresponding to the cutpoints c_j on the VAS scale.

In general, the ordinal regression model has response variable with K ordinal levels, and p covariates x_1, x_2, \dots, x_p , which may be either categorical or continuous. The model produces estimates of the coefficients $\beta_1, \beta_2, \dots, \beta_p$ corresponding to each of the covariates, as well as estimates of the $K-1$ cutpoints on the pain scale $\alpha_1, \alpha_2, \dots, \alpha_{K-1}$.

If VAS is measured to 1-mm accuracy, it is a 101-point scale and we could in theory treat these as $K=101$ levels of an ordinal variable. However, it is doubtful that patients can reliably distinguish

101 levels of pain [22]. From a statistical point of view this model is not attractive as it produces estimates of 100 cutpoints ($\alpha_1, \alpha_2, \dots, \alpha_{100}$). A reasonable approach is to group the VAS scores into the finest gradation which the clinician feels that patients are able to distinguish, and to analyze these levels using ordinal regression analysis. However, this is not the most satisfactory solution. There is a degree of arbitrariness in the choice of the number of levels K , and the VAS cutpoints c_1, c_2, \dots, c_{K-1} . One may obtain different results for different choices of these parameters, and the parameters could be chosen in order to deliver the most favourable results for the researcher, which is not a good feature. In addition, categorization involves loss of information, which is not desirable. In the next section we present an alternative version of ordinal regression analysis for VAS scores that does not involve grouping of the scores into categories.

2.2.5. Continuous ordinal regression

It is possible to avoid the pitfalls of categorization and retain the original VAS scores for analysis in an ordinal regression framework. The “continuous ordinal” regression model [23] preserves the ordinality of VAS scores while avoiding the estimation of 100 cutpoints.

Fig. 4 depicts the relationship between VAS score and pain. In Fig. 4(a) and (b), VAS is categorized into five and ten levels, respectively. In Fig. 4(c), VAS is treated as continuous and points for each category boundary are replaced by the smooth curve of the relationship between VAS and pain. In the continuous ordinal regression model, the cutpoints $\alpha_1, \alpha_2, \dots, \alpha_{K-1}$ are replaced by this smooth curve (the “g function”), whose shape depends on only three parameters. The shape of the g function is informative of the change in perception of pain at different levels, particularly at the extremes, and is a model output. As for ordinal regression, parameter estimates for the β s are also a model output, enabling hypothesis testing for all covariates. Ordinal continuous regression is available in the in the statistical language R [24] (ordinalCont package [25]).

2.3. Laser therapy for chronic knee pain study

We illustrate the methods on a study [26] in which $n=35$ patients were given a course of eight sessions of low-level laser therapy for chronic knee pain. VAS was measured at baseline and at the end of treatment. Patient age is also available for analysis.

2.4. Simulation study

In order to demonstrate the performance of the different statistical methods, we have also conducted a simulation study. We consider two equally-sized groups of patients (treatment and

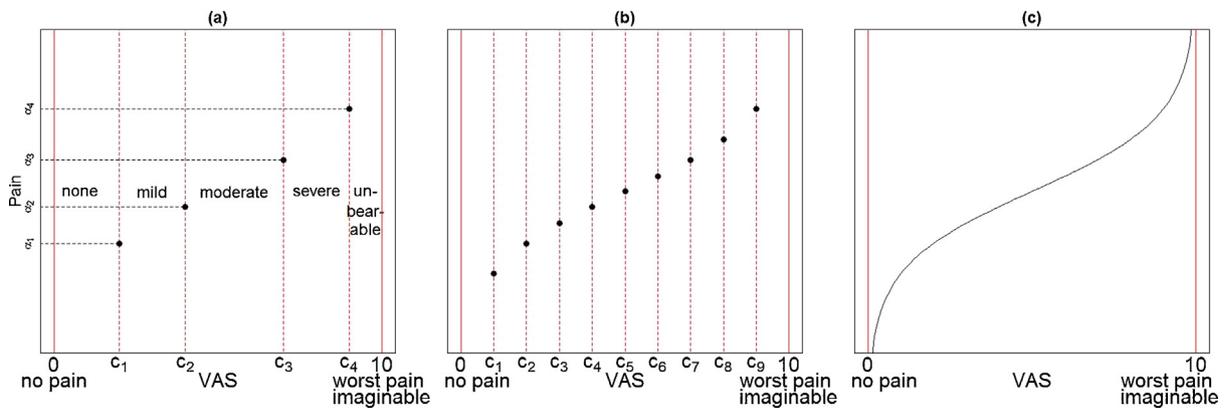


Fig. 4. Relationship between pain and VAS score.

control groups), as well as a continuous covariate, which may be a demographic variable such as age or body mass index, or a clinical biomarker such as C-reactive protein. We assume a beneficial treatment effect, varying from “small” to “large”. Under these assumptions, we randomly simulate VAS readings for patients, and repeat the procedure 1000 times for each combination of assumptions. For each random sample generated, we estimate the treatment effect using: the Wilcoxon rank sum test (equivalent to the Mann–Whitney test); the (unpaired) *t*-test; normal linear regression; and continuous ordinal regression. We compare these four methods by computing the power of the tests, i.e. the percentage of times that the null hypothesis of no treatment effect is rejected. When the treatment effect is small, we expect power to be quite low, increasing steadily as the treatment effect increases.

We have chosen the following scenarios for our simulations:

2.4.1. Scenario 1

The control group has a high level of pain (mean VAS 7.9), as in Fig. 3(c). We consider eight different assumptions for the mean VAS of the treatment group: 7.9 (no treatment effect); 7.5; 7.0; 6.6; 6.0; 5.5; 5.0; and 4.5 (strong treatment effect). In this case, where the VAS scores are highly skewed, it is well known that the use of *t*-tests, linear regression and other normal distribution-based methods is suboptimal and that an appropriate transformation should be applied to the response variable in order to achieve normality. However, our PubMed review revealed that, in the VAS literature, normal distribution-based methods are in the majority of cases being used without transformation or even testing for normality. Our intention with the simulation is to reflect current practice, and we therefore have not transformed the simulated response before applying the *t*-test and linear regression. If data are transformed to normality, that puts us in Scenario 2, described in the next paragraph.

2.4.2. Scenario 2

In this scenario the control group has a moderate level of pain (mean VAS 5.0), as in Fig. 3(b). We consider eight different assumptions for the mean VAS of the treatment group: 5.0 (no treatment effect); 4.5; 3.9; 3.4; 3.0; 2.5; 2.1; and 1.8 (strong treatment effect).

We also consider two alternatives within each of these scenarios:

Scenarios 1a and 2a. There is no covariate affecting VAS.

Scenarios 1b and 2b. There is a covariate affecting VAS, which is balanced across the two groups, and the covariate effect is unrelated to the treatment effect. An example of this in the pain context is the situation in which age is the covariate and

- older people experience more pain;

- the treatment benefit is the same for all ages; and
- the treatment and control groups are not significantly different with respect to age.

Sample sizes of 25 and 50 patients in each group were simulated. We also simulated unbalanced group sample sizes, and found very similar results, so have not reported these.

3. Results

3.1. Laser therapy for chronic knee pain study

Fig. 5 shows the distribution and ladder plot of VAS scores at these two time points, which demonstrate a marked improvement in pain after treatment. Fig. 6 is a scatterplot of VAS scores against patient age, with time (baseline/after treatment) indicated. There appears to be a positive relationship of VAS with age, with older patients reporting higher levels of pain.

As these are before–after measurements, a paired analysis is appropriate. Ref. [26] reported a significant improvement in pain after treatment ($p < 0.001$), based on the Wilcoxon signed rank test. This analysis is correct but not optimal, as it does not adjust for patient age. If age were incorporated into the analysis this would answer the question as to whether laser therapy is more effective in older patients, important because drug side effects for osteoarthritis cause significantly more problems in an older population. The inclusion of a continuous covariate may be accomplished using the continuous ordinal regression model, without categorization of the VAS scores. The repeated-measures nature of the data is modelled by using a random effect for patient. Using the ordinalCont package in R, a significant improvement in VAS is found ($p < 0.001$), with patient age a significant covariate ($p = 0.009$) and sex ($p = 0.817$) and an age–treatment interaction ($p = 0.493$) both not significant. The unadjusted effect size is 4.38 (95% CI: 2.93–5.82); and the effect size adjusted for age is 5.17 (95% CI: 3.61–6.73). This shows that (1) a lowering of VAS after treatment has been demonstrated, unadjusted and after adjusting for patient age; (2) older patients reported worse pain; (3) the lowering of VAS was the same across the age range; and (4) there was no gender effect on the lowering of VAS. **Note that, because there was no control group in this study, the observed effect of the lowering of VAS cannot be attributed to the treatment, even though it is statistically significant.**

3.2. Simulation study

The results of the simulation study are shown in Figs. 7 and 8. We interpret these as follows:

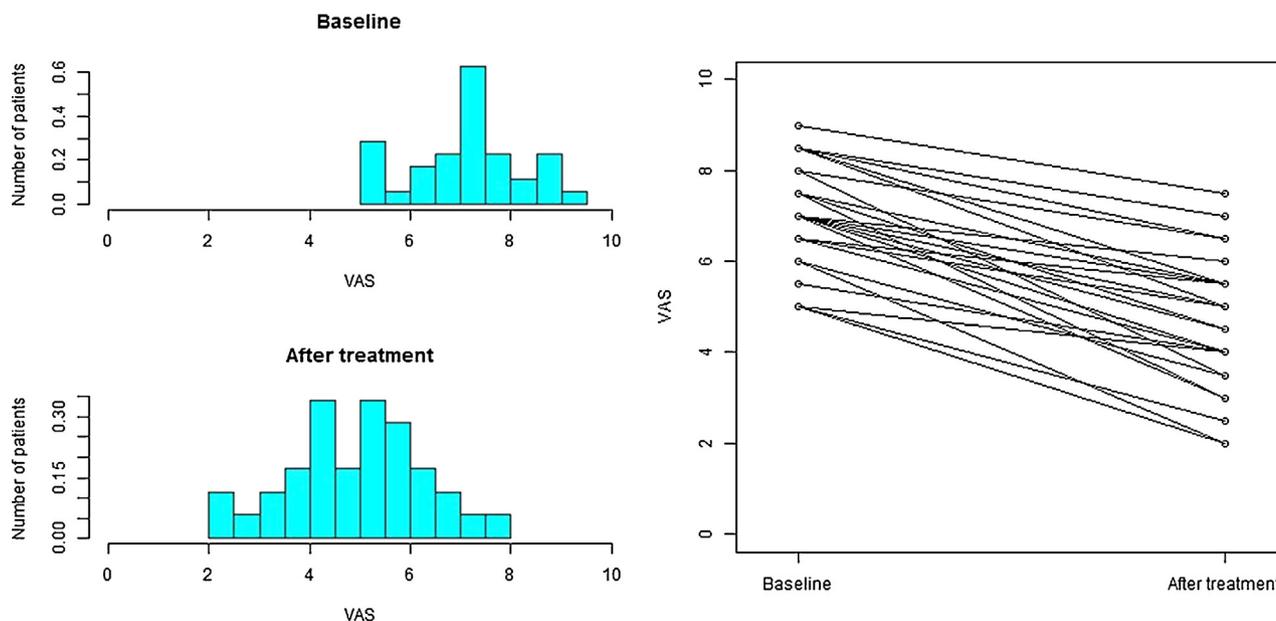


Fig. 5. Histograms (left panel) and ladder plot (right panel) of VAS scores at baseline and after treatment, Nakamura study [26].

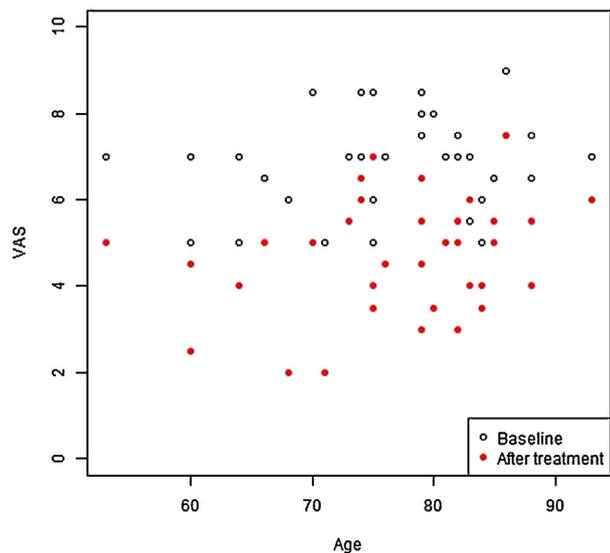


Fig. 6. Scatterplot of VAS against patient age, Nakamura study [26].

No covariates (1a, 2a): When the base level of pain is moderate and the distribution of VAS is therefore not highly skewed (2a), the four statistical tests have indistinguishable power. When the base level of pain is high (i.e. VAS has a highly skewed distribution, 1a), continuous ordinal regression and the Wilcoxon test have indistinguishable power, which is superior to that of linear regression and the *t*-test (which are in fact the same test in this case).

Recommendation: If there are no covariates affecting VAS (or none that are available), use the Wilcoxon (or other appropriate distribution-free) test.

Covariates (1b, 2b): In all cases continuous ordinal regression has superior power. When the distribution of VAS is not highly skewed (2b), linear regression has power only slightly lower than that of continuous ordinal regression; however, in the case of highly skewed VAS (1b), the power advantage of continuous ordinal regression is more marked, particularly for smaller sample size. Note that the Wilcoxon and *t*-test are not regression methods, i.e. they do not incorporate covariates. These results show the loss of

power by these two tests when a significant covariate is omitted from an analysis.

Recommendation: When there are significant covariates, continuous ordinal regression has superior power and is the method of choice. However, if the distribution of VAS is not highly skewed, linear regression (and its variants) is a pragmatic alternative. The Wilcoxon and *t*-test have inferior power and should not be considered.

4. Discussion

The aim of this paper is to present a comprehensive approach to the statistical analysis of VAS scores in a research setting, to more accurately reflect potential differences in treatment outcomes compared with placebo or other treatments. Measuring the subjective experience of pain is a continuing challenge in pain medicine. It is one of the original six core outcome domains proposed by Dworkin et al., which should be considered when designing clinical trials of chronic pain [27]. Reflecting the continuing developments in pain medicine assessment, these core domains have now increased to nine and will continue to be refined by future research [28]. In most clinical trials a measure of change in self-reported pain intensity will provide the primary outcome measure and is the “gold standard” for measurement according to Dworkin et al. [27]. A variety of measurement tools have been developed from unidimensional rating scales to complex questionnaires, each with its specific strengths and limitations. The aim of the researcher is to capture as accurately as possible a measure of the intangible, to make appropriate clinical assessments which guide the decision-making process and evidence based practice. The aim for the patient is to convey to the doctor or nurse who is delivering pain relief an accurate reflection of his or her experience of pain so that analgesia commensurate with pain intensity can be administered.

Defined by the IASP as an emotional experience pain, like pleasure or happiness, is an intangible variable which cannot be directly measured, unlike other physiological measures such as blood pressure or heart rate. Technological developments using functional magnetic imaging of the brain demonstrate activation of several cortical regions in response to nociceptive stimuli but as yet such techniques have not translated to an objective measure of what a patient experiences as “pain”. This remains within the realm of

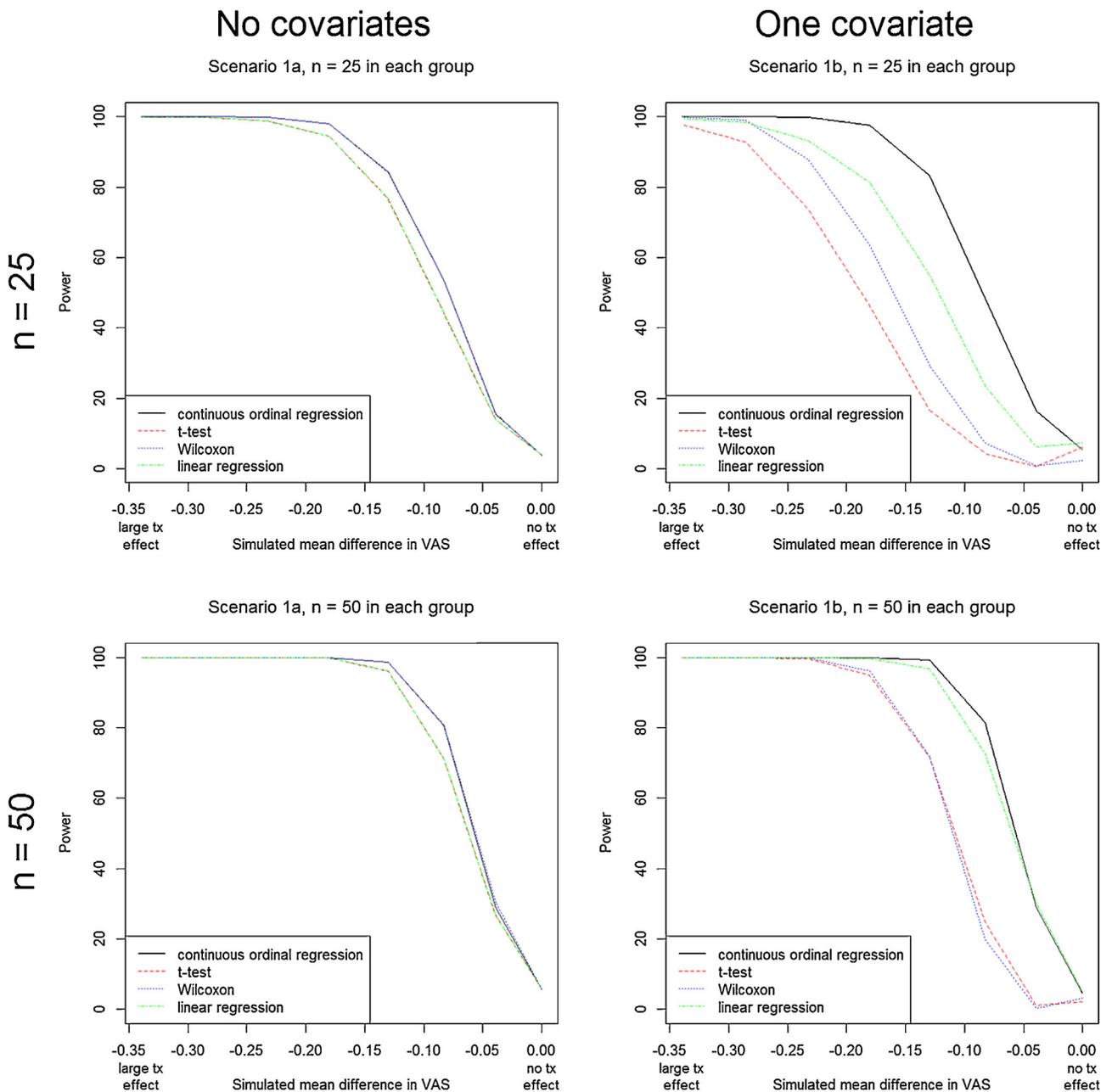


Fig. 7. Power (percentage of times treatment effect was detected) of the four statistical tests, under scenarios 1a and 1b, and sample sizes 25 and 50 per group.

research. At the other end of the technological spectrum, VAS in spite of its limitations is a validated primary outcome measure for pain intensity. It is widely and increasingly used both as a research tool and pragmatically in daily clinical practice. Not only must the VAS accurately reflect pain intensity for a patient on any single occasion such as in the management of acute post-operative pain or labour pain, but it must also be consistent over repeated measures for management of chronic pain, in research settings and guiding development of evidence based practice.

The question we address in this paper is what is the most appropriate statistical method for VAS analysis and under what conditions. As we have demonstrated, this question has been the subject of debate by other authors. However we provide evidence by way of our simulated data analysis that continuous ordinal regression has important advantages over commonly used methods, especially where there are significant covariates and highly skewed VAS scores. Moreover, continuous ordinal regression for

small sample sizes has greater power than more traditional methods. This means that fewer participants are required to achieve a desired level of power, always an advantage in clinical trials. This is a particular advantage where conditions are rare or treatment is expensive or difficult to administer. The method is accessible to researchers, in the R statistical package “ordinalCont” [25]. In addition to standard continuous ordinal regression modelling, ordinalCont also includes sophisticated features such as varying coefficients [29] for the situation in which proportionality, an assumption of ordinal regression models, is not present, and smoothing functions for covariates [30].

It is unquestionable that reliable data underpin evidence-based decisions and that appropriate statistical methods should be used [31]. The statistical approach we present here has important implications for patients, researchers and health costs. Prescribing of opioids, for example, with all the attendant problems is an area of great clinical concern where VAS scores directly influence

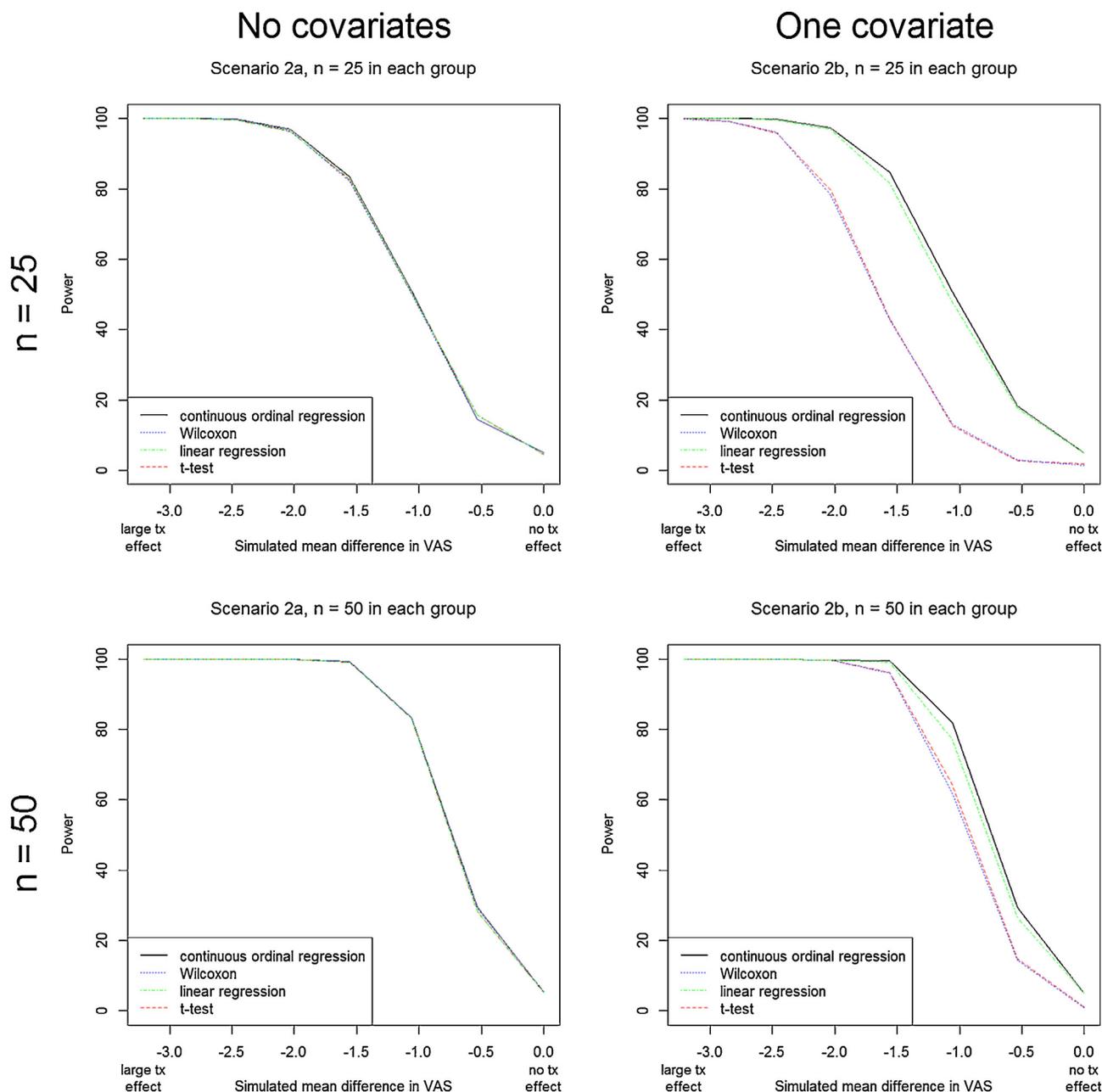


Fig. 8. Power (percentage of times treatment effect was detected) of the four statistical tests, under scenarios 2a and 2b, and sample sizes 25 and 50 per group.

prescribing patterns [32]. Indeed this reinforces the ethical requirement to ensure that appropriate methods are used for data analysis to provide the best available evidence for effective clinical decision-making [33].

Ethical issues

The article uses previously published data and a simulation study, and therefore was not subject to ethics approval.

Funding and conflict of Interest

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgement

The authors thank Dr Greg Glazov for his assistance.

References

- [1] Freyd M. The graphic rating scale. *J Educ Psychol* 1923;14:83–102.
- [2] Bond MR, Pilowsky I. Subjective assessment of pain and its relationship to the administration of analgesics in patients with advanced cancer. *J Psychosom Res* 1966;10:203–8.
- [3] Scott J, Huskisson EC. Graphic representation of pain. *Pain* 1976;2:175–84.
- [4] Merskey H, Bogduk N. Classification of chronic pain, IASP Task Force on Taxonomy. Seattle, WA: International Association for the Study of Pain Press; 1994.
- [5] Breivik EK, Bjørnsson GA, Skovlund E. A comparison of pain rating scales by sampling from clinical data. *Clin J Pain* 2000;16:22–8.
- [6] Williamson A, Hoggart B. Pain: a review of three commonly used pain rating scales. *J Clin Nurs* 2005;14:798–804.
- [7] Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for pain (VAS pain), Numeric Rating Scale for pain (NRS pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). *Arthritis Care Res* 2011;63:S240–52.
- [8] Jensen MP, Karoly P, Braver S. The measurement of clinical pain intensity: a comparison of six methods. *Pain* 1986;27:117–26.
- [9] Dexter F, Chestnut DH. Analysis of statistical tests to compare visual analog scale measurements among groups. *Anesthesiology* 1995;82:896–902.

- [10] Stevens SS. On the theory of scales of measurement. *Science* 1946;103:677–80.
- [11] Philip BK. Parametric statistics for evaluation of the visual analog scale. *Anesth Analg* 1990;71:710.
- [12] Wewers ME, Lowe NK. A critical review of visual analogue scales in the measurement of clinical phenomena. *Res Nurs Health* 1990;13:227–36.
- [13] Svensson E. Comparison of the quality of assessments using continuous and discrete ordinal rating scales. *Biom J* 2000;42:417–34.
- [14] Kersten P, Küçükdeveci AA, Tennant A. The use of the Visual Analogue Scale (VAS) in rehabilitation outcomes. *J Rehabil Med* 2012;44:609–10.
- [15] Kersten P, Küçükdeveci AA, Tennant A. How should we use the Visual Analogue Scale (VAS) in rehabilitation outcomes? IV: Reply on How should we use the Visual Analogue Scale (VAS) in rehabilitation outcomes. *J Rehabil Med* 2012;44:803–4.
- [16] 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.
- [17] Price DD, Staud R, Robinson ME. How should we use the visual analogue scale (VAS) in rehabilitation outcomes? II: visual analogue scales as ratio scales: an alternative to the view of Kersten. *J Rehabil Med* 2012;44:800–1.
- [18] Forrest M, Andersen B. Ordinal scale and statistics in medical research. *BMJ* 1986;292:537–8.
- [19] Lindsey JK. Applying generalized linear models. Springer Science & Business Media; 1997.
- [20] Fodstad K, Staff AC, Laine K. Effect of different episiotomy techniques on perineal pain and sexual activity 3 months after delivery. *Int Urogynecol J* 2014;25:1629–37.
- [21] Ananth CV, Kleinbaum DG. Regression models for ordinal responses: a review of methods and applications. *Int J Epidemiol* 1997;26:1323–33.
- [22] Butler PV. Linear analogue self-assessment and procrustean measurement: a critical review of visual analogue scaling in pain assessment. *J Clin Psychol Med Settings* 1997;4:111–29.
- [23] Manuguerra M, Heller GZ. Ordinal regression models for continuous scales. *Int J Biostat* 2010;6.
- [24] R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2015 <https://www.R-project.org/>.
- [25] Manuguerra M, Heller GZ. ordinalCont: ordinal regression analysis for continuous scales. R package version 0.3; 2015 <https://cran.r-project.org/web/packages/ordinalCont/>.
- [26] Nakamura T, Ebihara S, Ohkuni I, Izukura H, Harada T, Ushigome N, Ohshiro T, Musha Y, Takajashi H, Tsuchiya K, Kubota A. Low level laser therapy for chronic knee joint pain patients. *Laser Ther* 2014;23:273–7.
- [27] Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, Kerns RD, Stucki G, Allen RR, Bellamy N, Carr DB. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. *Pain* 2005;113:9–19.
- [28] Turk DC, Dworkin RH, Revicki D, Harding G, Burke LB, Cella D, Cleeland CS, Cowan P, Farrar JT, Hertz S, Max MB. Identifying important outcome domains for chronic pain clinical trials: an IMMPACT survey of people with pain. *Pain* 2008;137:276–85.
- [29] Park BU, Mammen E, Lee YK, Lee ER. Varying coefficient regression models: a review and new developments. *Int Stat Rev* 2015;83:36–64.
- [30] Wood S. Generalized additive models: an introduction with R. CRC Press; 2006.
- [31] Altman DG. Statistical reviewing for medical journals. *Stat Med* 1998;17:2661–74.
- [32] Reddy BS. The epidemic of unrelieved chronic pain: the ethical, societal, and regulatory barriers facing opioid prescribing physicians. *J Leg Med* 2006;27:427–42.
- [33] Svensson E. Different ranking approaches defining association and agreement measures of paired ordinal data. *Stat Med* 2012;31:3104–17.