Dear Author,

Please, note that changes made to the HTML content will be added to the article before publication, but are not reflected in this PDF.

Note also that this file should not be used for submitting corrections.
Basic neuroscience

Reliable estimation of nociceptive withdrawal reflex thresholds

Michael Brun Jensen *, José Biurrun Manresa, Ole Kæseler Andersen

Center for Neuroplasticity and Pain, SMU, Department of Health Science and Technology, Aalborg University, Frederik Bajers Vej 7, Room E1-202, 9220 Aalborg Øst, Denmark

HIGHLIGHTS

• Withdrawal reflex thresholds can often not be estimated using sural nerve stimulation.
• A lower failure-rate may be obtained by eliciting the reflex at the arch of the foot.
• Electrical stimulation at the arch enables more reliable reflex threshold estimation.

ARTICLE INFO

A B S T R A C T

Background: Assessment of the nociceptive withdrawal reflex (NWR) is frequently applied to probe the excitability level of the spinal nociceptive circuitry. In humans, the NWR threshold (NWR-T) is often estimated by applying electrical stimulation over the sural nerve at the lateral malleolus. Such stimulation may be associated with substantial pain and discomfort rendering completion of the assessment infeasible.

New method: As an alternative to sural nerve stimulation, NWR-Ts were also estimated by electrical stimulation at the arch of the foot. Failure-rates and test–retest reliability of these two procedures were evaluated. A fully-automated interleaved up-down staircase procedure was used to estimate the NWR-T for both stimulation sites. NWRs were detected from EMG measured over the biceps femoris and tibialis anterior muscles, respectively. A total of three repeated measures were performed in two different sessions to evaluate the test–retest reliability of the two methods using Bland–Altman agreement analysis.

Results: The failure rate of NWR-T estimation based on electrical stimulation of the sural nerve (29%) was substantially higher than when the NWR was elicited by stimulation at the arch of the foot (5%).

Comparison with existing method: The analysis of test–retest reliability indicated that the two methods for NWR-T estimation were equally reliable for within-session comparisons, but stimulation at the arch of the foot enabled NWR-T estimation with superior between-session reliability.

Conclusions: These results support a paradigm shift within NWR-T estimation favoring stimulation at the arch of the foot.

© 2015 Published by Elsevier B.V.

1. Introduction

The nociceptive withdrawal reflex (NWR) is a polysynaptic spinal reflex responsible for moving the limbs away from potentially harmful stimuli (Sandrini et al., 2005). There exists a close correlation between the NWR and experienced pain (Willer, 1977). Several independent investigations have provided evidence for NWR facilitation in patients with various pain conditions involving central mechanisms leading to increased neuronal responsiveness (Banic et al., 2004; Biurrun Manresa et al., 2013; Courtney et al., 2009; Desmeules et al., 2003; Lim et al., 2011, 2012; Neziri et al., 2010; Sandrini et al., 2006; Sterling et al., 2008). This indicates that the NWR threshold (NWR-T) constitute a valid measure of the excitability level of the spinal nociceptive system.

In humans, the NWR is generally elicited by electrical stimulation of the sural nerve at the lateral malleolus and measured...
with surface EMG recorded over the ipsilateral brevis head of the biceps femoris muscle (Sandrini et al., 2005). However, the electrical stimulation applied directly over the sural nerve is often associated with substantial pain and discomfort. In both healthy volunteers and pain patients, a number of subjects find the painful stimulation intolerable and the assessment cannot be completed (Banic et al., 2004; Biurrun Manresa et al., 2014b; Sterling et al., 2008). In this way, the practical applicability of the NWR-T as a clinical measure of nociceptive excitability is reduced. Another concern related to this assessment methodology regards the reliability of the obtained measurements. Previous studies investigating the reliability of the NWR-T reported excellent within-session reliability but larger variability between sessions (Biurrun Manresa et al., 2011; Micalos et al., 2009) and it has been suggested that the variability may be caused to some extent by slight inconsistencies in electrode placement resulting in variation in neural activation by the electro-cutaneous stimulation over the sural nerve.

Methods for quantification of reflex receptive fields involve elicitation of the NWR by electrical stimulation applied to the skin under the sole of the foot and measurement of the NWR by surface electromyography (EMG) over the tibialis anterior muscle (Neziri et al., 2009). To most subjects, this procedure is more tolerable than sural nerve stimulation, increasing the likelihood of successful NWR-T estimation. Furthermore, it is hypothesized that stimulation of free nerve endings in the skin at the arch of the foot (the most sensitive area) is less prone to variations in electrode placement than transcutaneous stimulation applied directly over the sural nerve trunk. If so, this may result in a more reliable NWR-T assessment.

The present study investigates the practical applicability of the two methods for NWR-T estimation by evaluating the frequency with which each method fails to yield a successful NWR-T estimation and by examining the test–retest reliability of the two procedures.

### 2. Materials and methods

#### 2.1. Subjects

Twenty one healthy volunteers (age: 18–35 years, 16 females) participated in the study. Written informed consent was obtained from all subjects prior to participation and the Declaration of Helsinki was respected. The study was approved by the local ethical committee (approval number VN–20130006).

#### 2.2. Experimental procedure

Two individual estimations of the NWR-T were performed by applying electrical stimulation over the sural nerve and under the sole of the foot, respectively (Fig. 1). An automated staircase procedure was used for online NWR detection to allow for a completely objective evaluation. This experimental procedure was performed three times on each subject: two times in one double session (1 h pause between estimates, no replacement of electrodes) and one time in a single session. Sessions were 48 ± 1 h apart, and the order was randomized. The order of the two methods for NWR-T estimation was also randomized between subjects but kept constant between both intra- and inter-session reassessments for each individual subject.

#### 2.2.1. Initial setup

Subjects were asked to refrain from caffeine, nicotine, alcohol and strenuous exercise for at least 4 h and from analgesic medication for 24 h prior to the experiment. During the experiment, the subject was placed in supine position with back support inclined 135 degrees relative to the horizontal level. A pillow was placed under the knees, resulting in a knee flexion of 45 degrees. Prior to actual data acquisition, the subject was thoroughly familiarized with electrical stimulation to reduce the effects of arousal or anxiety.

![Fig. 1. Study design and overall methodology. The nociceptive withdrawal reflex threshold (NWR-T) was assessed using two different procedures (A and B) involving two different locations for electrical stimulation and EMG recording.](image-url)
2.3. NWR-T estimation for sural nerve stimulation

2.3.1. Electrical stimulation

Electrical stimulation was delivered through two surface electrodes (20 × 15 mm, type 700, Ambu A/S, Denmark) placed dorsally and dorso-caudally of the lateral malleolus to stimulate the sural nerve at its retromalleolar pathway. Each stimulus consisted of a constant current pulse train of five individual 1 ms pulses (felt as a single stimulus) delivered at 200 Hz by a computer controlled electrical stimulator (Noxitec IES 230, Aalborg, Denmark).

2.3.2. EMG recording

Activity in the ipsilateral BF muscle was recorded using surface electromyography (EMG). Two surface electrodes (type 720, Ambu A/S, Denmark) were placed in parallel on the shaved skin over the ipsilateral brevis head of the BF muscle with an interelectrode distance of 2 cm. A common reference electrode (70 × 100 mm, Pals, Axelgaard, USA) was placed on the ipsilateral knee. The signals were amplified, filtered (10–500 Hz), sampled (10 kHz) and stored (1 s window including 200 ms pre-stimulation).

2.3.3. Automated NWR detection

Automated NWR detection was performed online using custom made software based on evaluation of interval peak z-score (Rhudy and France, 2007). The EMG signals were rectified and their interval peak z-score were calculated as:

$$\text{Interval peak } z\text{-score} = \frac{\text{reflex window peak} - \text{baseline mean}}{\text{baseline standard deviation}} \quad (1)$$

The reflex window was defined as the 80–150 ms post-stimulus interval whereas the baseline mean and standard deviation was calculated from the 70 ms pre-stimulus interval. A NWR was detected if the interval peak z-score exceeded 12 (France et al., 2009; Rhudy and France, 2007).

2.4. NWR-T estimation for foot sole stimulation

2.4.1. Electrical stimulation

A surface stimulation electrode (20 × 15 mm, type 700, Ambu A/S, Denmark) was placed at the arch of the foot to elicit the NWR. One large common anode (70 × 100 mm, Pals, Axelgaard, USA) was placed on the dorsum of the foot to ensure that the stimulus was perceived as coming from the sole of the foot. The electrode was moved slightly in case the evoked sensation indicated direct nerve trunk stimulation (i.e. radiating sensation). Each stimulus consisted of a constant current pulse train of five individual 1 ms pulses (felt as a single stimulus) delivered at 200 Hz by a computer controlled electrical stimulator (Noxitec IES 230, Aalborg, Denmark).

2.4.2. EMG recording

Activity in the ipsilateral tibialis anterior muscle was recorded using surface electromyography (EMG). Three surface electrodes (type 720, Ambu A/S, Denmark) were placed in parallel on the shaved skin along the overall orientation of the muscle with an interelectrode distance of 2 cm. A common reference electrode (70 × 100 mm, Pals, Axelgaard, USA) was placed on the ipsilateral knee. The tri-polar electrode configuration fed three separate amplifiers for simultaneous recording of one double differential and two single-differential EMG signals. The signals were amplified, filtered (10–500 Hz), sampled (10 kHz) and stored (1 s window including 200 ms pre-stimulation).

2.4.3. Automated NWR detection

Automated NWR detection was performed online using custom made software based on evaluation of interval peak z-score (Rhudy and France, 2007) (see Section 2.3.3) and a recently published method denoted Conduction Velocity Analysis (CVA) (Jensen et al., 2013). CVA was specifically designed to distinguish genuine reflexes from EMG crosstalk resulting in reflex detection with an improved specificity.

To perform the CVA, the two single differential EMG signals were high pass filtered (10th order Butterworth with zero phase delay) with cut-frequency at 80 Hz, whereupon cross-correlation of the two recordings were performed. The cross-correlations were normalized by the product of the norm of each of the two signals. From the resulting cross-correlogram, two features were extracted; the conduction velocity (CV) defined as the temporal displacement of the peak and the maximal value of the normalized cross-correlation, (see Jensen et al., 2013 for detailed description of the method).

A NWR was detected if the interval peak z-scores from both single and double differential EMG signals exceeded 12 and both the CV and the maximal value of the normalized cross-correlation did not exceed a set of fixed thresholds of respectively 34 m/s and 0.80 indicative of EMG crosstalk (Jensen et al., 2013).

2.5. Staircase method for NWR-T estimation

Both methods for NWR-T estimation involved the same up-down staircase method: the stimulation intensity was initially set to 1 mA and then increased in steps of 2 mA until a NWR was detected, whereupon the intensity was reduced in steps of 1 mA until a NWR was no longer detected. The intensity was subsequently increased and decreased in steps of 0.5 mA until a total of 3 ascending and 3 descending estimations of the NWR-T had been performed. The final NWR-T was then calculated as the mean value of the last 2 ascending and 2 descending estimations. Stimulations were applied with a 3–5 s inter-stimulus interval via a computer controlled electrical relay. If the stimulation intensity reached 50 mA or the stimulation intensity became intolerable for the subject, the stimulation was discontinued and the NWR-T assessment was considered unsuccessful.

2.6. Data analysis and statistics

For both methods, the percentage of unsuccessful NWR-T estimations was calculated and the number of subjects in which any of the three repeated assessments did not result in a successful NWR-T estimation was noted. A Wilcoxon signed rank test was applied to compare the estimated NWR-Ts. In the following within- and between-session reliability analysis, only the comparisons directly involving a missing value were excluded from the analysis.

Reliability was assessed using Bland–Altman analysis (Bland and Altman, 1999). Bland–Altman analysis compares the average and the difference between data measured in two different sessions, from which the limits of agreement (LOA) can be derived as the average difference (bias) ± 1.96 times the standard deviation of the differences. The LOA delimits the range within which 95% of the differences between two repeated measurements may be expected to lie. The results of the Bland–Altman agreement analysis is expressed as lower LOA, upper LOA and bias. The 95% confidence intervals (CI) were calculated for the lower and upper LOA. Paired t-tests were performed to determine if the biases were significantly different from zero in order to evaluate if there exists any significant difference in the means from the paired measures.

Finally, to test the level of association between the two methods for NWR-T estimation, a Pearson correlation was performed.
combining all assessments from the three repeated measures but excluding estimates for both methods in the case of missing data.

3. Results

In 6 out the 21 subjects (29%), the NWR-T could not be estimated in at least one of the three repeated measures when the electrical stimulation was applied over the sural nerve (the NWR-T estimation was unsuccessful in 16% of all assessments). For these subjects, the level of electrical stimulation became intolerable and the assessment was terminated. This was only the case in one subject (5%) when applying the stimulation at the arch of the foot (the NWR-T estimation was unsuccessful in less than 2% of all assessments). The NWR-Ts (median [lower quartile, upper quartile] estimated involving electrical stimulation over the sural nerve (13.8 [9.5, 21.8] mA) were significantly higher than those estimated by stimulation at the arch of the foot (7.4 [5.3, 11.4] mA).

As shown in Fig. 2, Bland–Altman analysis indicated that NWR-T estimation based on sural nerve stimulation showed better within-session reliability than between-session reliability. A corresponding difference in within-versus between-session reliability was not observed for the NWR-Ts estimated by stimulation at the arch of the foot. Accordingly, whereas the reliability of the two methods was similar for within-session comparisons, assessment

Fig. 2. Bland–Altman plots for within- (left) and between- session comparison of reflex thresholds estimated by electrical stimulation respectively over the sural nerve (top) and at the arch of the foot. The dotted lines represent the bias—the mean differences between the two measurements (measurement 2–measurement 1), whereas the slashed lines and the grey areas indicate respectively the mean values and the 95% confidence intervals for the limits of agreement.

Please cite this article in press as: Jensen MB, et al. Reliable estimation of nociceptive withdrawal reflex thresholds. J Neurosci Methods (2015). http://dx.doi.org/10.1016/j.jneumeth.2015.06.014
of the NWR-T by stimulation over the sural nerve trunk was associated with inferior between-session reliability. However, only the CIs for the upper LOAs for the two methods did not overlap. None of the calculated biases were significantly different from zero.

Notably, as depicted in Fig. 3, only a negligible and non-significant correlation was found between the NWR-Ts estimated using the two different methods. However, by excluding two subjects with consistent but incongruous NWR-Ts as outliers, a clearly significant linear correlation was identified.

4. Discussion

The present study investigated the practical applicability of two methods for NWR-T estimation eliciting the NWR by electrical stimulation applied over the sural nerve and at the arch of the foot, respectively. A focal point of the investigation was the subjective sensation experienced by the test subjects, because stimulation intensities that reach individual tolerance thresholds result in unsuccessful NWR-T estimation. Furthermore, the evaluation was based on examination of test–retest reliability of the two methods. The results showed that the success rate of NWR-T estimation based on electrical stimulation of the sural nerve (71%) was substantially lower than when the NWR was elicited by electrical stimulation at the arch of the foot (95%). Analysis of test–retest reliability indicated that the two methods for NWR-T estimation are more or less equally reliable for within-session comparisons, but stimulation at the arch of the foot results in more reliable estimations of the NWR-T between individual sessions.

4.1. Subjective experience and unsuccessful NWR-T estimation

Over the years, numerous studies have estimated the NWR-T using sural nerve stimulation. However, except from a very limited number of papers (Banić, 2004; Biurrun Manres et al., 2014a; Sterling et al., 2008), these studies do not report failure rates of the assessment. Electrical stimulation applied over the sural nerve trunk are often perceived as rather painful and both personal communication with various labs using the method on a regular basis and experience from our own lab, suggest that more often than indicated by the literature, reliable NWR-T estimation is not obtainable. To our knowledge, the present study is the first one to address this issue. The results indicate that a substantial number of subjects find the assessment intolerable and that failure to estimate the NWR-T estimation constitutes a genuine methodological concern. The problem may however be reduced by changing the stimulation site to the arch of the foot, which resulted in a much improved success rate. This could likely be related to the fact that the NWR-Ts estimated by sural nerve stimulation generally are higher, which intuitively may cause more pain and an elevated risk of reaching the tolerance level of the subjects.

The NWR response may be modulated by the position of the subject (Sandrini et al., 2005). To control for this, the two different NWR-T assessment methods were performed with the subjects lying in the same supine position. However, the biomechanical response to electrical stimulation at the two sites may differ. Whereas electrical stimulation at the arch of the foot may result in a small isolated NWR response predominantly involving dorsal flexion, stimulation over the sural nerve trunk will generally trigger a more widespread response involving ankle, knee and hip flexion, which to some degree may be restricted by the bed in this supine position. To elaborate on this issue, additional studies investigating the feasibility and reliability of NWR-T assessment during different positions of the subjects may be of great practical importance.

4.2. Reliability

The reliability of the two NWR-T estimation methods was assessed using Bland–Altman analysis. The resulting LOAs (Fig. 2) showed that NWR-Ts estimation involving sural nerve stimulation demonstrated much better within-session reliability than between-session reliability. A corresponding difference was not present for the method involving stimulation at the arch of the foot, indicating an equally good within- and between-session reliability.

When it comes to comparing which of the two methods is most reliable, the slightly different range of the two types of NWR-Ts does challenge the interpretation of the analysis. However, whereas the LOAs indicate that the two methods are equally reliable for within-session comparisons, the method involving sural nerve stimulation clearly displays inferior reliability when comparing measurements performed in two different sessions. Indeed, 4 subjects in this study (19%) showed a between-session estimation difference of around 10 mA or more when the NWR-T was estimated using sural nerve stimulation. Considering the median NWR-T of 13.8 (9.5, 21.8) mA, this seems to constitute a genuine reliability issue.

4.3. Sources of variability and future recommendations

In line with previous reports (Biurrun Manresa et al., 2011; Micalos et al., 2009), the results of this study show that NWR-Ts estimated using electrical stimulation over the sural nerve suffer from a large inter-session variability. The most probable explanation for this observation is minor variations in electrode positioning at the retromalleolar pathway. Despite using anatomical references to position the stimulation electrodes in a consistent manner, slight differences in electrode positioning are hard to avoid, which may likely change the electrocutaneous stimulation of the well-defined underlying nerve trunk. The free nerve endings in the skin at the arch of the foot constitute a smoother continuum of excitable tissue rendering stimulation of this site more robust with respect to minor variation in electrode position. This is supported by the similar within- and between-session reliability experienced in this study when the NWR is elicited by electrical stimulation at the arch.

The two methods for NWR-T estimation are hypothesized to measure the same phenomenon and were therefore expected to be correlated. This could not readily be confirmed by the present
results. Only when two subjects had been excluded as outliers, a significant linear correlation between the two estimation methods could be demonstrated. Visual inspection of the graphical representation of paired NWR-Ts from the two methods (Fig. 3) shows that for both subjects these extreme measurements are characterized by a situation where disproportionate high NWR-Ts were consistently estimated using one method compared to the other. The consistency of these incongruous estimates of the NWR-T indicates that they are not likely the result of reduced reliability. Instead they seem to draw attention to a methodological imperfection rendering the estimation of invalid high NWR-Ts possible. Concurrently with increasing stimulation intensities in the applied staircase method, the discomfort for the subject also increases. At some point it starts to be difficult for the subject to keep the prestimulus EMG signal at resting level due to anxiety/anticipation of the next painful stimuli. Hereby the sensitivity of the NWR detection is reduced and artificial elevated NWR-Ts may be estimated because only very strong NWRs can counterbalance the increased prestimulus EMG activity to arrive at a sufficient high interval peak z-score. This seems likely to have occurred for both estimation methods, which together with the similar level of within-session reliability of the two methods, indicate that minor differences in NWR detection procedures cannot be a decisive factor for the observed difference in between-session reliability. Assuming that the considered outliers may be ascribed to decreased NWR detection sensitivity (applying to both methods), a linear correlation between the two methods seems to be established.

The NWR-Ts and the interpersonal variability estimated in this study using sural nerve stimulation is in line with previous reports on healthy subjects from different labs using comparable stimulation protocols (Banic et al., 2004; Rhudy et al., 2013; Sandrini et al., 2006). To improve the reliability of NWR-T estimation for research and clinical use, it is therefore recommended to base future NWR-T estimations on NWRs elicited by electrical stimulation at the arch of the foot.  

5. Conclusion
The present study investigates the practical applicability of two methods for NWR-T estimation involving electrical stimulation applied respectively over the sural nerve trunk and at the arch of the foot. In 29% of the subjects the NWR-T could not consistently be estimated using sural nerve stimulation due to intolerable pain sensations. This was only the case for one subject (5%) when the NWRs were elicited by stimulation at the arch of the foot. Furthermore, NWR-T estimation involving sural nerve stimulation showed inferior between-session reliability, likely because this type of stimulation is very sensitive to changes in electrode positioning. Taken together, these results support a paradigm shift within NWR-T estimation. The practical applicability of NWR-T estimation based on sural nerve stimulation seems to entail increase between-session variability, thus a change of stimulation site to the arch of the foot should be strongly considered in future studies to ensure more reliable NWR-T estimation.

Conflict of interest statement
The authors report no conflict of interest.

Acknowledgment
This work was supported by Danish National Research Foundation grant DNF121.