

WITHIN- AND BETWEEN-SESSION RELIABILITY OF SECONDARY HYPERALGESIA INDUCED BY ELECTRICAL HIGH-FREQUENCY STIMULATION

T. Cayrol¹, J. Lebleu¹, A. Mouraux², N. Roussel³, L. Pitance^{1*}, E. N. van den Broeke^{2*}

1: Institute of Experimental and Clinical Research, Health Sciences division, Université Catholique de Louvain, Neuro-Musculo-Skeletal-Lab (NMSK), Brussels, Belgium.

2: Institute of Neuroscience, System and Cognition division, Université catholique de Louvain, Brussels, Belgium.

3: Department of Rehabilitation Sciences and Physiotherapy (MOVANT), Faculty of Medicine and Health Sciences, University of Antwerp, Wilrijk, Belgium.

*Both authors contributed equally to this work.

Corresponding author:

Timothee Cayrol

Institute of Experimental and Clinical Research (IREC)

Université Catholique de Louvain

B-1200, Brussels

Belgium

Tel: +32 2 764 53 75

E-mail: timothee.cayrol@uclouvain.be

Submission type: Original article

Funding sources: This study was supported by the Université Catholique de Bruxelles. Emanuel N. van den Broeke is supported by the Fonds de Recherche Clinique UCL.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1002/EJP.1613](https://doi.org/10.1002/EJP.1613)

This article is protected by copyright. All rights reserved

Conflict of interest: None declared

What's already known about this topic?

- Secondary hyperalgesia is thought to result mainly from central sensitization, which may be an important feature of persistent pain
- Electrical high-frequency stimulation is a non-invasive, easily applicable, inexpensive and well-controlled method that induces secondary hyperalgesia in humans

What does this study add?

- This study evaluated the within- and between-session reliability of the secondary hyperalgesia induced by electrical high-frequency stimulation

Significance: It is crucial to evaluate central sensitization adequately in humans. This study formally establishes the reliability of secondary hyperalgesia induced by electrical high-frequency stimulation. The results of this study will improve future studies investigating secondary hyperalgesia in humans.

WITHIN- AND BETWEEN-SESSION RELIABILITY OF SECONDARY HYPERALGESIA INDUCED BY ELECTRICAL HIGH-FREQUENCY STIMULATION

ABSTRACT

Background. An increasing number of studies are focusing on secondary hyperalgesia to better understand central sensitization, as this phenomenon may play an important role in persistent pain. Recent studies have shown that, compared to the classical high frequency stimulation protocol (HFS) at 100 Hz, a protocol using 42 Hz stimulation induces a more intense and a larger area of secondary hyperalgesia (SH).

Objectives. The aim of this study was to investigate the within- and between-session reliability of SH induced by this optimized HFS protocol.

Methods. Thirty-two healthy subjects received HFS to their volar forearm in two sessions, separated by at least two weeks. SH was assessed by measuring the area size of increased sensitivity to pinprick stimuli after applying HFS, the sensitivity to pinprick stimuli after applying HFS, and the change in pinprick sensitivity after vs. before HFS. Assessments were made before HFS, and 30, 35 and 40 minutes after HFS. Relative and absolute reliability were analyzed using intra-class correlation coefficients (ICCs), coefficients of variation (CVs), standard error of means (SEMs) and the minimum detectable changes (MDCs).

Results. The area of SH showed good to excellent within-session and between-session *relative* reliability (ICCs > 0.80), except for the change in pinprick sensitivity, which showed close to poor between-session relative reliability (ICC=0.53). Furthermore, measures of *absolute* reliability generally demonstrated large between-subject variability and significant fluctuations across repeated measurements.

Conclusions. HFS-induced hyperalgesia is suitable to discriminate or compare individuals but it may not be sensitive to changes due to an intervention.

1. INTRODUCTION

Tissue injury is generally associated with increased pain sensitivity in the area of actual tissue damage (i.e., primary hyperalgesia) as well as the surrounding uninjured skin (i.e., secondary hyperalgesia) (Treede et al., 1992). Whereas primary hyperalgesia involves increased pain sensitivity to mechanical and heat stimuli, secondary hyperalgesia (SH) predominantly involves increased pain sensitivity to mechanical stimuli (Vollert et al., 2018). Primary hyperalgesia likely results from both peripheral and central sensitization, while SH is thought to result mainly from central sensitization (Baumann et al., 1991; Simone et al., 1991).

A large number of studies have focused on SH to better understand central sensitization, which may be an important feature of persistent pain (Arendt-Nielsen et al., 2018; Woolf, 2011). This is supported by evidence in musculoskeletal disorders and in postsurgical pain, as the area size of SH induced by capsaicin was found to be substantially increased in patients with fibromyalgia (Morris et al., 1998), and a larger area of SH surrounding the surgical incision within the first week after surgery seems associated with the development of persistent postsurgical pain (Lavand'homme et al., 2005; Martinez et al., 2012; Momeni et al., 2010).

Several experimental methods have been developed to induce SH. These include intradermal capsaicin injection or topical capsaicin application (Koltzenburg et al., 1992; LaMotte et al., 1991; Magerl et al., 1998), heat stimulation (Hansen et al., 2016a; Werner et al., 2013), the combination of capsaicin and heat stimulation (Werner et al., 2013), and electrical stimulation at high (Klein, 2004) or low frequencies (Sauerstein et al., 2018). However, some of these methods have drawbacks. Intradermal capsaicin injections are invasive procedures and the after-effects could be dependent on how deep the capsaicin is injected. Moreover, topical application relies on the absorption of capsaicin into the skin, which is variable across subjects (Pershing et al., 2004). In contrast, electrical stimulation, delivered using a specifically designed electrode to preferentially activate free nerve endings in the skin is non-invasive, easily applicable, inexpensive, and stimulates the axon directly resulting in a highly synchronized afferent activity (Staahl and Drewes, 2004). In addition, capsaicin or heat typically induces after-sensations, which is not the case for electrical stimulation. Furthermore, parameters such as the pattern and frequency of electrical

stimulation can be easily controlled, which has recently led to the finding that burst-like 42 Hz stimulation seem the optimal parameters for inducing SH (Van Den Broeke et al., 2019; Gousset et al., 2019).

No study has examined the reliability of SH induced after burst-like 42 Hz stimulation (High frequency stimulation, HFS). However, this is crucial for future studies that aim at investigating central sensitization, manifested as SH, in humans.

Therefore, the aim of this study was to assess the within- and between-session reliability of SH induced by 42 Hz HFS. Furthermore, we assessed the reliability of secondary outcomes including reported pain, stress, anxiety, catastrophizing, and autonomic responses evoked by HFS.

2. MATERIALS AND METHODS

2.1 Subjects

Thirty-two healthy volunteers participated in this study (16 females, 23.5 ± 3.1 years [mean $\pm$ SD]) and 16 males (27.8 ± 6.4 years). This number was based on a sample size calculation analysis (Bonett, 2002), which showed that, based on our design, a random sample of 32 subjects was required to produce a two-sided 95% confidence interval for intraclass correlation coefficients (ICCs), with a distance between the lower and the upper confidence limit of 0.100. Based on a secondary analysis of the area size of SH of a previous within-subject cross-over study (Van Den Broeke et al., 2019, Exp 1), the value of the ICC for the area size of increased pinprick sensitivity was estimated as 0.9.

All subjects were healthy adults recruited by advertisement. Exclusion criteria were: (1) experiencing a pre-existing pain condition; (2) self-reported medication (except contraceptives) consumption in the 48 hours preceding the study and/or self-reported recreational drug use; (3) presenting any self-reported medical conditions, including heart, neurological, dermatological, and psychiatric diseases; (4) exhibiting sign of damage at or near the volar forearm; (5) presence of tattoos on the tested forearm.

The study was approved by the local Ethical Committee and conducted according to the declaration of Helsinki, with preregistration (NCT03966508), at the Institute of Experimental and Clinical Research (IREC), Brussels, Belgium, from April 2019 to August 2019. All subjects provided informed consent and received financial compensation. They were not informed that this study was designed to assess reliability.

2.2 Design

In this within-subject design, subjects received HFS to their volar forearm (6-10 cm distal to the cubital fossa) in two sessions, separated by at least two weeks.

During each session (Fig. 1) pinprick sensitivity was measured before HFS (M0), 30 min after HFS (M1), 35 min after HFS (M2) and 40 min after HFS (M3). At the three post-HFS measurements also the area size of increased pinprick sensitivity was measured. The time points of the post-HFS measurements were chosen based on previous studies showing that the increased pinprick sensitivity induced by HFS reaches its plateau between 20 and 40 minutes (Klein, 2004; Pfau et al., 2011).

INSERT FIGURE 1 HERE

2.3 Procedure

The study was conducted in a quiet secluded room. The two sessions for each subject took place in the afternoon. Subjects were asked to refrain from drinks containing caffeine, nicotine, alcohol beverages, medication (except contraceptives) and intense physical activity the day of testing.

At baseline, demographic characteristics (gender, age, height, weight, hand dominance) were collected. During HFS and mechanical pinprick testing, subjects were seated in a comfortable chair, with their palms up resting on a table. The instructions to the subject were standardized across subjects. All procedures were conducted by the same investigator, who was trained in performing the mechanical pinprick testing and in applying HFS.

2.3.1 High-frequency electrical stimulation (HFS)

Before attaching the HFS electrode onto the skin, the skin was first cleaned with ether and alcohol. Then, the electrical detection threshold to a single square wave electrical pulse (pulse-width 2 ms) was determined using the methods of limits. The threshold was defined as the intensity at which a subject detected the electrical stimulus. The intensity of the first stimulus was set at 0.30 mA. Then, the intensity of subsequent stimuli decreased until the stimulus was no longer detected. The threshold was confirmed by at least two ascending and descending stimuli with intensities just below and above the threshold. Subjects were blinded to the intensity and timing of stimulation.

After determining the electrical detection threshold, HFS was applied. HFS consisted of 12 trains of 42 electrical pulses (charge compensated pulse shape: biphasic pulse consisting of a 2-ms square-wave pulse followed, after a 0.1 ms delay, by a 4-ms compensation pulse of opposite polarity having half the intensity of the first pulse) lasting for 1 second each and delivered in a 10 second inter-train interval (Van Den Broeke et al., 2019). The total duration of the stimulation protocol was 2 minutes. The stimulation intensity of HFS was set to 5 mA for all subjects and sessions.

The electrical pulses were triggered by a digital-analogue interface (National Instruments, Austin, USA) controlled by Matlab 2014B (The Mathworks Inc., Natick, USA) and delivered using a constant current electrical stimulator (Digitimer DS5, Digitimer, UK) to the skin via a custom electrode designed and built at the Centre for Sensory-Motor Interaction (Aalborg University, Denmark). The cathode of this electrode consisted of 16 blunt stainless-steel pins with a diameter of 0.2 mm, protruding 1 mm from the base and placed in a circle with a diameter of 10 mm. The anode was a stainless-steel ring with an inner diameter of 22 mm and an outer diameter of 40 mm (Van Den Broeke et al., 2019).

To avoid any confounding effect of handedness, the arm onto which high-frequency stimulation (HFS) was applied (dominant vs nondominant) was counterbalanced across subjects.

2.4 Primary outcomes

2.4.1 Perceived intensity elicited by mechanical pinprick stimuli

To assess the perceived intensity elicited by mechanical pinprick stimuli, a calibrated mechanical pinprick stimulator, exerting a normal force of 128 mN (MRC Systems GmbH, Heidelberg, Germany) was used.

The mechanical pinprick stimuli were delivered before and after HFS within a circular area 15-20 mm from the center of the area at which HFS was delivered. At each measurement a total of three stimuli were delivered. To avoid sensitizing the skin due to the repeated pinprick stimulation, the stimuli were not delivered twice at the same exact location.

Subjects were instructed to rate the average perceived intensity of the three stimuli on a numerical rating scale (NRS) ranging from 0 (no perception) to 100 (maximal pain), with 50 representing the transition from non-painful to painful domains of sensation (Cayrol et al., 2018). We analyzed two different outcomes regarding the perceived intensity elicited by the pinprick stimuli. The first outcome was the rating of perceived intensity after HFS and the second outcome was the *change* in perceived intensity (i.e. after HFS minus before HFS).

2.4.2 Estimated area size of increased pinprick sensitivity

At all three post-HFS measurements, after collecting the perceived intensity ratings, we estimated the area size of the increased pinprick sensitivity. For this, we stimulated the skin with the 128 mN pinprick stimulator along two perpendicular axes originating from the center of the area at which HFS was delivered.

Proximal-distal and medial-lateral axes were drawn on the skin at the beginning of the experiment using a felt-tip pen. For evaluating the proximal and distal borders of the area of increased pinprick sensitivity, mechanical pinprick stimulations started respectively at the lower third of the ventral arm and at the wrist. For the medial and lateral borders, mechanical pinprick stimulations started about 7-10 cm (depending on the size of the forearm) away from the center of the area at which HFS was delivered. Subjects were asked to progressively rotate their forearm so that the mechanical pinprick stimulator was always perpendicular to the skin.

Each stimulus was separated by steps of 0.5 cm, at a pace of 1-second stimulation and 1-second interval, in the direction of the HFS site (Hansen et al., 2016b; Werner et al., 2013). During the mapping process, subjects kept their eyes closed and were instructed to say

“now” when they felt “a clear increase of pinprick sensitivity”. When the subjects identified this increase for the first time, they kept their eyes closed and the pinprick stimulation was delivered again 1 cm away from this point. Then, stimulations moved again towards the center. If this time the subject reported an increase in pinprick sensitivity at the same location, the mapping process stopped. If the border was not confirmed, the mapping process was carried on towards the center.

The total area size of increased pinprick sensitivity was estimated with a custom-written Python script, which determined the area based on a virtual polygon connecting the four borders of the area with the best-fit parabolic segments.

2.5 Secondary outcomes

2.5.1 Borders and extent of the area of increased pinprick sensitivity

As secondary outcomes we analyzed individually the distal (D) lateral (L), proximal (P) and medial (M) borders and the length (D+P) and width (L+M) of the area of increased pinprick sensitivity.

2.5.2 Trait and State Anxiety

Anxiety was determined by the French version of the State-Trait Anxiety Inventory (STAI) (Gauthier and Bouchard, 1993). The STAI-Y2, assessing ‘trait’ anxiety was filled in by the subjects at least one day before the study. Moreover, the STAI-Y1 was used to measure ‘situational’ anxiety immediately before HFS.

2.5.3 Pain catastrophizing

Dispositional pain catastrophizing was assessed at least one day before the first experimental session with the French version of the pain catastrophizing scale (Sullivan et al., 1995). Pain catastrophizing related to HFS was assessed by the Situational Catastrophizing Questionnaire (Campbell et al., 2010) directly after HFS.

2.5.4 Heart-rate variability (HRV)

Variation in the time interval between heartbeats was measured as an index of arousal elicited during HFS. The time elapsing between two normal R-peaks in the electrocardiogram (NN intervals) was collected with a Polar H10 HR monitor with a Pro

Strap (Polar Electro Oy, Kempele, Finland) (Gilgen-Ammann et al., 2019) at rest (seated on a chair without talking) and during the 2 min stimulation procedure.

Abnormal R-peaks (artefacts) were removed with the Elite HRV software (Elite HRV LLC, Asheville, USA). Ultra-short heart rate variability (HRV) indices were analyzed with Kubios HRV Standard 3.2.0 for Macintosh (The Biomedical Signal and Medical Imaging Analysis Group, Kuopio, Finland) (Tarvainen et al., 2014) based on recent recommendations and international guidelines (Castaldo et al., 2019; Malik, 1996).

Time domain HRV indices included the NN mean value (MeanNN) and standard deviation (StdNN); Heart Rate mean (MeanHR) and standard deviation (StdHR); the square root of the mean of the sum of the squares of differences between adjacent NNs, divided by the number of NNs minus 1 (RMSSD); the number of pairs of successive NNs that differed by more than 50 ms (NN50) and their proportion (pNN50). Frequency domain indices extracted from power spectrum estimated with autoregressive model methods included high-frequency power (HF), low frequency power (LF), and the LF/HF ratio. Non-linear analysis indices included the standard deviations of the points perpendicular to (SD1) and along (SD2) the line-of-identity obtained from Poincaré plots, which is a graphical presentation of the correlation between consecutive NN intervals.

2.5.6 Perceived stress

Perceived stress was assessed by asking subjects to rate, directly after HFS, their perceived level of stress during HFS using a numerical rating scale ranging from 0 (no stress) to 100 (maximal stress).

2.6 Statistical analysis

2.6.1 Primary outcomes

Multiple steps were followed to examine the reliability of the primary outcomes.

First, scatterplots were created to visualize the within- and between-session reliability, checking for extreme outliers and general trends. Mean values and standard deviations (SDs) were reported.

Second, we checked for potential within- and between-session systematic errors. Within-session differences for each outcome were tested either with one way repeated-measures ANOVA or the Friedman test depending on the normality of the data distribution (Dontje et al., 2018). Between-session differences for the average of Session 1 versus Session 2 were tested using either the paired T-test or the Wilcoxon signed rank test.

Third, the relative reliability was assessed by estimating ICCs, which is an index related to the relative contribution of between-subject variance to the total variance (Portney, 2020). The distribution of the data for all outcome variables was checked for normality using histograms and Q-Q plots. There were no major deviations from normality and the homogeneity of variance criterion was met in all cases, therefore we calculated ICCs using parametric statistical tests. Within-session relative reliability was calculated with a single-measurement, absolute agreement, two-way mixed-effects model (Koo and Li, 2016). Between session relative reliability was assessed with a mean-rating for each session ($k=3$), absolute agreement, two-way mixed-effects model. Significance level was set to 0.05. We also compared the first, second and third measure of the first session with the first, second and third measure of the second session with a single-measurement, absolute agreement, two-way mixed-effects model. ICCs were interpreted as such: $ICC \leq 0.50$ poor; $0.50-0.75$ poor to moderate; $>0.75-0.90$ good; and >0.90 excellent relative reliability (Portney, 2020).

Fourth, we determined absolute reliability by calculating the coefficients of variation (CV), the standard error of means (SEMs), the minimum detectable changes (MDCs) and Bland-Altman plots (Furlan and Sterr, 2018; Portney, 2020). The CV, which is a measure of variability, was calculated as the ratio of the pooled standard deviation to the mean. The SEM, which provides an estimate for interpreting how much error is likely to be present in a

single measurement, was estimated as the square root of the mean square error term (MSE) from the repeated-measures ANOVA (Portney, 2020; Weir, 2005). The MDC, defined as the smallest change that can be detected beyond measurement error, was calculated as follows: $MDC = 1.96 \times \sqrt{2} \times SEM$. Bland-Altman plots were used to visually assess the spread of differences and the limits of agreement (mean differences $\pm 1.96 \times SD$) for repeated measurements of Sessions 1 and 2 and for the mean of each session. The bias (the average of the differences) was also calculated.

2.6.2 Secondary outcomes

The statistical analysis for the borders, the length and the width characterizing the area of increased pinprick sensitivity were the same as for the primary outcomes.

Between-session systematic errors for pinprick sensitivity, electrical detection thresholds, pain during the stimulation procedure, stress ratings, 'situational' anxiety, pain catastrophizing and HRV indices were assessed with paired T-tests or Wilcoxon matched-pair signed rank tests. ICCs were calculated with a single-measurement, absolute agreement, two-way mixed-effects model for all secondary outcomes, except for HRV indices, which were calculated with a mean-rating, absolute agreement, two-way mixed-effects model. ICCs were based on logarithmically transformed data for Mean NN, StdNN, RMSSD, NN50, pNN50, LF, HF, LF/HF and SD1 (Ellis et al., 2008).

3. RESULTS

All 32 healthy subjects completed both sessions. Table 1 presents the demographic characteristics of subjects and their psychological profile, which were in the normal range. On average, the time that elapsed from the first to the second session was 20 ± 6 days. On average the pinprick stimuli applied before HFS were not perceived as painful for most subjects (91%). After HFS, pinprick stimuli were perceived as more intense in all subjects but two, and 69% rated them as painful.

INSERT TABLE 1 HERE

3.1 Within-session reliability

Means and standard deviation of primary outcomes at all three post-HFS measurements for the first testing session, as well as within-session ICCs, CVs, SEMs, and MDCs, are shown in Table 2. Individual data for session 1 are shown in Fig. 2.

INSERT TABLE 2 HERE

INSERT FIGURE 2 HERE

Regarding the first testing session, the CVs ranged from 26.63% to 56.09%, indicating large between-subject variability.

The repeated measure one-way ANOVA revealed a significant difference across the three measurements for the estimated area size of increased pinprick sensitivity ($F(31, 62) = 38.10$; $p < 0.001$). Post hoc multiple comparisons showed that there was a significant decrease (11.2%) from the first to the second measure.

Excellent relative reliability was observed across the three post-HFS measurements for the estimated area size ($ICC = 0.91$), pinprick sensitivity after HFS ($ICC = 0.92$) and the change in pinprick sensitivity ($ICC = 0.90$).

SEMs and MDCs expressed as percentages of the mean for pinprick sensitivity after HFS (respectively 7.71% and 21.37%) were lower than for the change in pinprick sensitivity (respectively 17.75% and 49.20%) and for the estimated area size (respectively 14.56% and 40.36%).

INSERT FIGURE 3 HERE

When comparing measurement 1 with measurement 2 (M1 vs M2), and measurement 2 with measurement 3 (M2 vs M3) with Bland-Altman plots (Fig. 3), a random scattering of points distributed evenly around the average mean difference for the estimated area size (bias ranging from 1.57 to 9.44), the pinprick sensitivity after HFS (bias ranging from 0.00 to 1.19), and for the change in pinprick sensitivity (bias ranging from 0.00 to 1.19). For the estimated area size, a larger pattern of disagreement and more outliers were identified among subjects with high mean scores. Limits of agreement (LoAs) were large for the estimated area size (from -18.92 to 37.82 for M1 vs V2 ; from -26.3 to 29.5 for M2 vs M3), but less so for the pinprick sensitivity after HFS (from -11.06 to 11.06 for M1 vs M2 ; from -

8.03 to 10.40 for M2 vs M3) and for the change in pinprick sensitivity after HFS (from -11.06 to 11.06 for M1 vs M2 ; from -08.03 to 10.40).

Regarding the second session, similar trends were observed (Fig. S1 and Table S1).

3.2 Between-session reliability

Means and standard deviation of primary outcomes at each session, as well as the between-session ICCs, CVs, SEMs and MDCs, are shown in Table 3. Individual data of the average obtained for the three measurements for each session are shown in Fig. 4.

INSERT TABLE 3 HERE

INSERT FIGURE 4 HERE

Large CVs were observed, ranging from 30.54% to 58.06%, indicating large between-subject variability.

The Wilcoxon signed rank test revealed a significant decrease from Session 1 to Session 2 for the estimated area size ($p=0.033$, 5.9%). No significant difference was found for the pinprick sensitivity after HFS and the change in pinprick sensitivity.

Relative reliability was excellent for the estimated area size (ICC=0.93) and good for pinprick sensitivity after HFS (ICC=0.82). However, the relative reliability of the change in pinprick sensitivity was close to poor (ICC=0.53).

When respectively comparing the first, second and third measures of session one with session two, the relative reliability declined. Only the estimated area size showed good relative reliability (≥ 0.83). The pinprick sensitivity after HFS showed moderate relative reliability (≥ 0.65) and the reliability of change in pinprick sensitivity was poor (< 0.40).

SEMs and MDCs expressed as percentages of the mean were lower for the pinprick sensitivity after HFS (respectively 16.81% and 46.61) and for the estimated area size of increased pinprick sensitivity (respectively 20.26 % and 56.16%) than for the change in pinprick sensitivity (respectively 46.12% and 127.84%).

INSERT FIGURE 5 HERE

The Bland-Altman plots (Fig. 5), illustrated a random scattering of points distributed evenly around the average mean difference for the estimated area size (bias = 4.54), the pinprick sensitivity after HFS (bias = -1.79), and for the change in pinprick sensitivity (bias = 0.56). For the estimated area size, a larger pattern of disagreement and more outliers were identified among subjects with high mean scores. Limits of agreement (LoAs) were large for the

estimated area size (from -27.65 to 46.74), for the pinprick sensitivity after HFS (from 27.78 to 24.19) and for the change in pinprick sensitivity (from -29.62 to 30.75).

Secondary outcomes

Means and standard deviation for the borders, the length and the width characterizing the area of increased pinprick sensitivity for the first testing session, as well as the within-session ICCs, CVs, SEMs, and MDCs, are shown in Table S2. For session 1, excellent relative reliability (ranging from 0.92 to 0.94) was observed across the three post-HFS measurements for the distal border, the proximal border, and the length of the area of increased pinprick sensitivity after HFS. Good relative reliability (ICC=0.84) was observed for the width of the area and the lateral border. The medial border showed moderate relative reliability (ICC=0.72). The lowest SEMs and MDCs were found for the distal border (respectively 10.22% and 28.33%) and for the length (respectively 8.57% and 23.76%). Regarding the second testing session, similar trends were observed (Table S3).

Means and standard deviation for the borders, the length and the width characterizing the area of increased pinprick sensitivity at each session, as well as the between-session ICCs, CVs, SEMs and MDCs, are shown in Table S4. Regarding between-session reliability, excellent relative reliability (ranging from 0.90 to 0.97) was observed the distal border, the proximal border, the length and the width of the area of increased pinprick sensitivity after HFS. Good relative reliability was observed for the lateral border (ICC=0.86) and the medial border (ICC=0.79). The lowest SEMs and MDCs were found for the distal border (respectively 13.72% and 38.02%), the length (respectively 9.26% and 25.68%) and the width (respectively 14.22% and 39.41%).

Means and standard deviation of baseline pinprick sensitivity, electrical detection thresholds, pain during HFS, stress ratings, 'situational' anxiety and pain catastrophizing related to HFS, as well as within-session ICCs, CVs, SEMs, and MDCs, are presented in Table 4. Only perceived stress level during HFS significantly ($p=0.02$) decreased from session 1 to session 2 (18%).

INSERT TABLE 4 HERE

Relative reliability was poor to moderate (ranging from 0.50 to 0.72) for baseline pinprick sensitivity, electrical detection thresholds, pain during HFS, stress, 'situational' anxiety, and catastrophizing.

From resting state to the HFS, mean HR, StdNN, StdHR, RMSSD, NN50, LF, HF, SD1, SD2 significantly increased, conversely, mean NN significantly decreased. The pNN50 and LF/HF did not change.

Means and standard deviation of HRV indices during HFS, as well as within-session ICCs and CVs, are presented in Table S5. The HRV indices recorded during HFS (MeanNN, StdNN, MeanHR, StdHR, RMSSD, NN50, pNN50, LF, HF, LF/HF, SD1, and SD2) did not significantly differ between the first and second session. The relative reliability of HRV indices according to ICCs was good (ranging from 0.81 to 0.86) for all variables except StdHR, NN50, pNN50 and LF/HF, for which it was poor to moderate (ranging from 0.65 to 0.71).

4. DISCUSSION

The present study shows that the within- and between-session *relative* reliability of HFS-induced SH measured by estimating the area size of increased pinprick sensitivity and the pinprick sensitivity after applying HFS was good to excellent (ICCs > 0.80). However, between-session *relative* reliability of the change (post minus pre HFS) in pinprick sensitivity was close to poor (ICC=0.53). Moreover, measures of *absolute* reliability (CV, SEM and MDC) generally demonstrated large between-subject variability and significant fluctuations across repeated measurements of SH. Regarding secondary outcomes, large between-subject variability was also observed. Relative reliability was good for most HRV indices evaluating autonomic arousal, and poor to moderate for pain, anxiety, stress and catastrophizing related to HFS.

The relative reliability of the estimated area size of increased pinprick sensitivity was higher than the relative reliability found in studies with other experimental pain models, including heat/capsaicin sensitization (0.30 to 0.76), brief thermal sensitization (0.48 to 0.84) and burn injury sensitization models (0.58) (Hansen et al., 2016a; Werner et al., 2013). Our results are in agreement with a previous study using 10 Hz continuous stimulation where within- and between-session relative reliability was good (ICC>0.80) for pinprick sensitivity assessed with the same mechanical pinprick stimulator (Xia et al., 2016). Therefore, it

appears that the increased pinprick sensitivity induced by electrical stimulation is more reliable than the increased pinprick sensitivity induced by the other methods. One possible explanation could be that the afferent neural activity generated by electrical stimulation is more reproducible as compared to, for example, capsaicin, which is strongly dependent on skin permeability when applied topically and the localization of the injection site when delivered intradermal (Handwerker and Kobal, 1993; Staahl and Drewes, 2004).

We observed a large between-subject variability in the estimated area size. At present the specific factors contributing to this variability are unknown, but may include stress, diet, hormone levels, skin receptor density, anatomical and functional brain differences, and genetics (Hansen et al., 2016b; Werner et al., 2013). Moreover, there is evidence that stress-like doses of hydrocortisone attenuates increased pinprick sensitivity (Michaux et al., 2012), and that increased spatial attention (Filbrich et al., 2019) and negative expectations (Van Den Broeke et al., 2014) can increase increased pinprick sensitivity.

It could be that the poor between-session relative reliability for the change in pinprick sensitivity is due to the fact that this outcome variable is calculated as a difference score, which inexorably diminishes ICCs because it restricts the range of possible values (Rankin and Stokes, 1998). Nonetheless, large SEMs, large MDCs and Bland-Altman plots demonstrated the poor absolute reliability for this outcome variable. Additionally, the medial border of the area of increased pinprick sensitivity was less reliable than the other borders. This could be due to the fact that our pinprick stimulator had to be applied perpendicularly to the skin, which, for the medial border, required subjects to externally rotate their forearm and tilt their chest. This was sometimes uncomfortable for the subjects and may have affected the rating.

Interestingly, we observed a systematic decrease of the estimated area size across measurements in both sessions. This could reflect the phenomenon of habituation of sensitization (Thompson and Spencer, 1966). In contrast, the pain ratings remained consistent, which suggest that habituation is less prominent in the perceived intensity. Alternatively, subjects may have been tempted to repeat their previous ratings rather than carefully re-assessing the sensitivity to pinprick stimuli, even though they were not informed that this study was designed to assess reliability.

In our study, all outcome variables, except the change in pinprick sensitivity, showed good to excellent relative between-session reliability. However, this contrasted with measures of absolute reliability, as large SEMs and MDCs were observed, indicating that features characterizing the area of increased pinprick sensitivity in individual subjects fluctuate across repeated measurements. Therefore, these features seem suited to discriminate between subjects with respect to their response to nociceptive input, but may not be very sensitive to detect change in a subject (Portney, 2020; de Vet et al., 2006). Minimal detectable changes calculated in this study will help determine sample sizes of future studies with the aim to identify clinically meaningful changes (Dontje et al., 2018; Marcuzzi et al., 2017).

A potential limitation of the present study is that our results cannot not be generalized as the subjects were all young adults, which may represent a meaningful sample bias. Moreover, other reliability studies have used eight directions to assess the extent of SH, which may provide a more accurate estimation of the area size of increased pinprick sensitivity as compared to assessing extent along four directions (Hughes et al., 2002; Ringsted et al., 2015). While this study was not specifically designed to compare the reliability between the different outcomes, the change in pinprick sensitivity (after HFS minus before HFS) appears unreliable between sessions. This was not the case for the length of the area of increased pinprick sensitivity along the proximal-distal axis, or the pinprick sensitivity reported after HFS. Furthermore, a larger disagreement was identified on Bland-Altman plots among subjects with high mean scores for the estimated area size, suggesting that larger area sizes of increased pinprick sensitivity are less reliable.

Strictly speaking, the HFS induced increase in sensitivity to the pinprick stimuli delivered in the present study does not satisfy the definition of “hyperalgesia”, as proposed by the International Association for the Study of Pain, because pinprick stimuli applied before HFS were not perceived as painful. Nevertheless, the punctate probe that we used preferentially activates mechano-sensitive nociceptors (Garell et al., 1996; Nagi et al., 2019; Slugg et al., 2000, 2004). Therefore the increase in pinprick sensitivity that was reported by the participants is most likely related to SH and the central sensitization of mechanical nociceptive pathways (Baumann et al., 1991; Simone et al., 1991; Treede and Cole, 1993).

To conclude, pinprick sensitivity after HFS and most features characterizing the area of increased pinprick sensitivity show good to excellent between sessions *relative* reliability. Therefore, the HFS-mechanical hyperalgesia model is suitable to discriminate or compare subjects. However, the model may not be ideal for studies interested in between-session changes in individual subjects, for example, due to an intervention, because measures of *absolute* reliability demonstrated fluctuations across repeated measurements.

ACKNOWLEDGEMENTS

We would like to thank Louis de Viron, Prof. Christine Detrembleur and Prof. Céline Bugli for their technical help and material support.

REFERENCES

- Arendt-Nielsen, L., Morlion, B., Perrot, S., Dahan, A., Dickenson, A., Kress, H.G., Wells, C., Bouhassira, D., Mohr Drewes, A. (2018). Assessment and manifestation of central sensitisation across different chronic pain conditions. *Eur J Pain (United Kingdom)* 22, 216–241.
- Baumann, T.K., Simone, D.A., Shain, C.N., RH, L. (1991). Neurogenic hyperalgesia: the search for the primary cutaneous afferent fibers that contribute to capsaicin-induced pain and hyperalgesia. *J Neurophysiol* 66, 212–227.
- Bonett, D.G. (2002). Sample size requirements for estimating intraclass correlations with desired precision. *Stat Med* 21, 1331–1335.
- Campbell, C.M., Kronfli, T., Buenaver, L.F., Smith, M.T., Berna, C., Haythornthwaite, J.A., Edwards, R.R. (2010). Situational Versus Dispositional Measurement of Catastrophizing: Associations With Pain Responses in Multiple Samples. *J Pain* 11, 443-453.e2.
- Castaldo, R., Montesinos, L., Melillo, P., James, C., Pecchia, L. (2019). Ultra-short term HRV features as surrogates of short term HRV: A case study on mental stress detection in real life. *BMC Med Inform Decis Mak* 19, 1–13.
- Cayrol, T., Pitance, L., Roussel, N., Mouraux, A., Van Den Broeke, E.N. (2018). No evidence of widespread mechanical pressure hyperalgesia after experimentally induced central sensitization through skin nociceptors. *Pain Reports* 3, 1–6.
- Dontje, M.L., Dall, P.M., Skelton, D.A., Gill, J.M.R., Chastin, S.F.M. (2018). Reliability, minimal detectable change and responsiveness to change: Indicators to select the best method to measure sedentary behaviour in older adults in different study designs. *PLoS One* 13, 1–16.

Ellis, R.J., Sollers, J.J., Edelstein, E.A., Thayer, J.F. (2008). Data transforms for spectral analyses of heart rate variability. *Tech Pap ISA* 472, 371–376.

Filbrich, L., van den Broeke, E.N., Legrain, V., Mouraux, A. (2019). The focus of spatial attention during the induction of central sensitization can modulate the subsequent development of secondary hyperalgesia. *Cortex*.

Furlan, L., Sterr, A. (2018). The applicability of standard error of measurement and minimal detectable change to motor learning research - A behavioral study. *Front Hum Neurosci* 12, 1–10.

Garell, P.C., McGillis, S.L.B., Greenspan, J.D. (1996). Mechanical response properties of nociceptors innervating feline hairy skin. *J Neurophysiol* 75, 1177–1189.

Gauthier, J., Bouchard, S. (1993). Adaptation Canadienne-Française de la forme révisée du State-Trait Anxiety Inventory de Spielberger. *Can J Behav Sci* 25, 559–578.

Gilgen-Ammann, R., Schweizer, T., Wyss, T. (2019). RR interval signal quality of a heart rate monitor and an ECG Holter at rest and during exercise. *Eur J Appl Physiol* 119, 1525–1532.

Gousset, S., Mouraux, A., Van Den Broeke, E.N. (2019). Burst-like conditioning electrical stimulation is more efficacious than continuous stimulation for inducing secondary hyperalgesia in humans. *J Neurophysiol*.

Handwerker, H.O., Kobal, G. (1993). Psychophysiology of experimentally induced pain. *Physiol Rev* 73, 639–671.

Hansen, M.S., Wetterslev, J., Pipper, C.B., Østervig, R., Asghar, M.S., Dahl, J.B. (2016a). The area of secondary hyperalgesia following heat stimulation in healthy male volunteers: Inter- and intra-individual variance and reproducibility. *PLoS One* 11, e0155284.

Hansen, M.S., Wetterslev, J., Pipper, C.B., Østervig, R., Asghar, M.S., Dahl, J.B. (2016b). The area of secondary hyperalgesia following heat stimulation in healthy male volunteers: Inter- and intra-individual variance and reproducibility. *PLoS One* 11, 1–17.

Hughes, A., Macleod, A., Growcott, J., Thomas, I. (2002). Assessment of the reproducibility of intradermal administration of capsaicin as a model for inducing human pain. *Pain* 99, 323–331.

Klein, T. (2004). Perceptual Correlates of Nociceptive Long-Term Potentiation and Long-Term Depression in Humans. *J Neurosci* 24, 964–971.

Koltzenburg, M., Lundberg, L., Torebjörk, H.E. (1992). Dynamic and static components of mechanical hyperalgesia in human hairy skin. *Pain* 51, 207–219.

Koo, T.K., Li, M.Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med* 15, 155–163.

LaMotte, R.H., Shain, C.N., Simone, D.A., Tsai, E.F.P. (1991). Neurogenic hyperalgesia: Psychophysical studies of underlying mechanisms. *J Neurophysiol* 66, 190–211.

Lavand'homme, P., De Kock, M., Waterloos, H. (2005). Intraoperative epidural analgesia combined with ketamine provides effective preventive analgesia in patients undergoing major digestive surgery. *Anesthesiology* 103, 813–820.

Loeser, J.D., Treede, R.-D. (2008). The Kyoto protocol of {IASP} Basic Pain Terminology. *Pain* 137, 473–477.

Magerl, W., Wilk, S.H., Treede, R.D. (1998). Secondary hyperalgesia and perceptual wind-up following intradermal injection of capsaicin in humans. *Pain* 74, 257–268.

Malik, M. (1996). Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *Circulation* 93, 1043–1065.

Marcuzzi, A., Wrigley, P.J., Dean, C.M., Adams, R., Hush, J.M. (2017). The long-term reliability of static and dynamic quantitative sensory testing in healthy individuals. *Pain* 158, 1217–1223.

Martinez, V., Ben Ammar, S., Judet, T., Bouhassira, D., Chauvin, M., Fletcher, D. (2012). Risk factors predictive of chronic postsurgical neuropathic pain: The value of the iliac crest bone harvest model. *Pain* 153, 1478–1483.

Michaux, G.P.N., Magerl, W., Anton, F., Treede, R.D. (2012). Experimental characterization of the effects of acute stresslike doses of hydrocortisone in human neurogenic hyperalgesia models. *Pain* 153, 420–428.

Momeni, M., De Kock, M., Lavand'Homme, P., Watremez, C., Van Dyck, M., Baele, P. (2010). Abnormal sensations evoked over the chest and persistent peri-incisional chest pain after cardiac surgery. *Acta Anaesthesiol Belg* 61, 55–62.

Morris, V., Cruwys, S., Kidd, B. (1998). Increased capsaicin-induced secondary hyperalgesia as a marker of abnormal sensory activity in patients with fibromyalgia. *Neurosci Lett* 250, 205–207.

Nagi, S.S., Marshall, A.G., Makdani, A., Jarocka, E., Liljencrantz, J., Ridderström, M., Shaikh, S., O'Neill, F., Saade, D., Donkervoort, S., Reghan Foley, A., Minde, J., Trulsson, M., Cole, J., Bönnemann, C.G., Chesler, A.T., Catherine Bushnell, M., McGlone, F., Olausson, H. (2019). An ultrafast system for signaling mechanical pain in human skin. *Sci Adv* 5, 1–10.

Pershing, L.K., Reilly, C.A., Corlett, J.L., Crouch, D.J. (2004). Effects of vehicle on the uptake and elimination kinetics of capsaicinoids in human skin in vivo. *Toxicol Appl Pharmacol* 200, 73–81.

Pfau, D.B., Klein, T., Putzer, D., Pogatzki-Zahn, E.M., Treede, R.D., Magerl, W. (2011). Analysis of hyperalgesia time courses in humans after painful electrical high-frequency stimulation identifies a possible transition from early to late LTP-like pain plasticity. *Pain* 152, 1532–1539.

Portney, L.G. (2020). *Foundations of clinical research : applications to evidence-based practice* (Philadelphia: F.A DAVIS).

Rankin, G., Stokes, M. (1998). Reliability of assessment tools in rehabilitation: An illustration of appropriate statistical analyses. *Clin Rehabil* 12, 187–199.

Ringsted, T.K., Enghuus, C., Petersen, M.A., Werner, M.U. (2015). Demarcation of secondary hyperalgesia zones: Punctate stimulation pressure matters. *J Neurosci Methods* 256, 74–81.

Sauerstein, K., Liebelt, J., Namer, B., Schmidt, R., Rukwied, R., Schmelz, M. (2018). Low-Frequency Stimulation of Silent Nociceptors Induces Secondary Mechanical Hyperalgesia in Human Skin. *Neuroscience* 387, 4–12.

Simone, D.A., Sorkin, L.S., Oh, U., Chung, J.M., Owens, C., RH, L., Willis, W.D. (1991). Neurogenic hyperalgesia: central neural correlates in responses of spinothalamic tract neurons. *J Neurophysiol* 66, 228–246.

Slugg, R.M., Campbell, J.N., Meyer, R.A. (2004). The population response of A- and C-fiber nociceptors in monkey encodes high-intensity mechanical stimuli. *J Neurosci* 24, 4649–4656.

Slugg, R.M., Meyer, R.A., Campbell, J.N. (2000). Response of cutaneous A- and C-fiber nociceptors in the monkey to controlled-force stimuli. *J Neurophysiol* 83, 2179–2191.

Staahl, C., Drewes, A. (2004). Experimental Human Pain Models: A Review of Standardised Methods for Preclinical Testing of Analgesics. *Basic Clin Pharmacol Toxicol* 95, 97–111.

Sullivan, M.J.L., Bishop, S.R., Pivik, J. (1995). The Pain Catastrophizing Scale: Development and validation. *Psychol Assess* 7, 524–532.

Tarvainen, M.P., Niskanen, J.P., Lipponen, J.A., Ranta-aho, P.O., Karjalainen, P.A. (2014). Kubios HRV - Heart rate variability analysis software. *Comput Methods Programs Biomed* 113, 210–220.

Thompson, R.F., Spencer, W.A. (1966). Habituation: A model phenomenon for the study of neuronal substrates of behavior. *Psychol Rev* 73, 16–43.

Treede, R.D., Cole, J.D. (1993). Dissociated secondary hyperalgesia in a subject with a large-fibre sensory neuropathy. *Pain* 53, 169–174.

Treede, R.D., Meyer, R.A., Raja, S.N., Campbell, J.N. (1992). Peripheral and central mechanisms of cutaneous hyperalgesia. *Prog Neurobiol* 38, 397–421.

Van Den Broeke, E.N., Geene, N., Rijn, C.M., O.H.G., W., Oosterman, J. (2014). Negative expectations facilitate mechanical hyperalgesia after high-frequency electrical stimulation of human skin. *Eur J Pain* 18, 86–91.

Van Den Broeke, E.N., Gousset, S., Bouvy, J., Stouffs, A., Lebrun, L., Van Neerven, S.G.A., Mouraux, A. (2019). Heterosynaptic facilitation of mechanical nociceptive input is dependent on the frequency of conditioning stimulation. *J Neurophysiol* 122, 994–1001.

de Vet, H.C.W., Terwee, C.B., Knol, D.L., Bouter, L.M. (2006). When to use agreement versus reliability measures. *J Clin Epidemiol* 59, 1033–1039.

Vollert, J., Magerl, W., Baron, R., Binder, A., Elena, E.-K., Geisslinger, G., Gierthmühlen, J., Henrich, F., Hüllemann, P., Klein, T., Lötsch, J., Maier, C., Oertel, B., Sigrid, S.-H., Tölle, T., Treede, R.-D. (2018). Pathophysiological mechanisms of neuropathic pain: comparison of sensory phenotypes in patients and human surrogate pain models. *Pain Articles* i, 1.

Weir, J.P. (2005). Quantifying Test-Retest Reliability Using the Intraclass Correlation Coefficient And The SEM. *J Strength Cond Res*.

Werner, M.U., Petersen, K.L., Rowbotham, M.C., Dahl, J.B. (2013). Healthy Volunteers Can Be Phenotyped Using Cutaneous Sensitization Pain Models. *PLoS One* 8.

Woolf, C.J. (2011). Central sensitization: Implications for the diagnosis and treatment of pain. *Pain* 152, S2–S15.

Xia, W., Mørch, C.D., Andersen, O.K. (2016). Test-retest reliability of 10 Hz conditioning electrical stimulation inducing long-Term potentiation (LTP)-Like pain amplification in humans. *PLoS One* 11, 1–18.

Female, n (%)	16 (50)
Age, years (range)	25.7 ± 5.3 (19-42)
Height , cm	172.6 ± 8.0
Weight, kg	66.7 ± 12.6
BMI, kg/m ²	22.3 ± 3.7
Pain Catastrophising ('dispositional'), PCS (0-52)	14.5 ± 8.2
Anxiety (trait), STAI-Y2 (20-80)	38.2 ± 7.8

Table 1. Demographic characteristics of the participants and their psychological profile. Data are presented in Mean ± SD unless otherwise specified. BMI, body mass index; PCS, Pain Catastrophizing Scale French-Canadian; STAI-Y2, Trait subscale of the State-Trait Anxiety Inventory.

Assessment measure	Measure 1 (M1)	Measure 2 (M2)	Measure 3 (M3)	ICC	CV	SEM	MDC
	<i>Mean ± SD</i>	<i>Mean ± SD</i>	<i>Mean ± SD</i>	<i>ICC (95% CI)</i>	<i>%</i>	<i>Absolute vs %</i>	<i>Absolute vs %</i>
Estimated area size, cm ²	84.25 ± 43.65	74.80 ± 39.45	73.23 ± 40.41	0.91 (0.82 to 0.95)	53.22	11.27 (14.56)	31.25 (40.36)
Pinprick sensitivity after HFS, 0-100	55.25 ± 13.28	55.25 ± 14.35	54.06 ± 16.06	0.92 (0.86 to 0.96)	26.63	4.23 (7.71)	11.72 (21.37)
Change in pin-prick sensitivity, 0-100	24.22 ± 12.92	24.22 ± 13.30	23.03 ± 13.86	0.90 (0.83 to 0.95)	56.09	4.23 (17.75)	11.72 (49.20)

Table 2. Within-session comparisons of the primary outcomes. The mean ± SD of the first, second and third measurement of session 1 for each primary outcome are shown. Intraclass correlation coefficients (ICCs) are reported with their 95% confident intervals and range from 0 to 1. The coefficient of variation (CV) is reported. In addition, standard error of mean (SEM) and minimal detectable change (MDC) are reported in the same unit as the measurements and in percentages relative to the mean. HFS, high frequency stimulation.

Assessment measure	Session 1	Session 2	ICC	CV	SEM	MDC
	<i>Mean ± SD</i>	<i>Mean ± SD</i>	<i>ICC (95% CI)</i>	<i>%</i>	<i>Absolute vs %</i>	<i>Absolute vs %</i>
Estimated area size, cm ²	77.42 ± 40.17	72.88 ± 46.85	0.93 (0.87 to 0.97)	58.06	15.23 (20.26)	42.21 (56.16)
Pinprick sensitivity after HFS, 0-100	54.85 ± 14.20	56.65 ± 19.45	0.82 (0.64 to 0.91)	30.54	9.37 (16.81)	25.98 (46.61)
Change in pin-prick sensitivity, 0-100	23.82 ± 12.91	23.30 ± 14.03	0.53 (0.02 to 0.77)	57.23	10.87 (46.12)	30.12 (127.84)

Table 3. Between-session comparisons of the primary outcomes. The mean ± SD of the first and second session for each primary outcome are shown. Intraclass correlation coefficients (ICCs) are reported with their 95% confident intervals and range from 0 to 1. The coefficient of variation (CV) is reported. In addition, standard error of mean (SEM) and minimal detectable change (MDC) are reported in the same unit as the measurements and in percentages relative to the means. HFS, high frequency stimulation.

Assessment measure	Session 1	Session 2	ICC	CV	SEM	MDC
	<i>Mean ± SD</i>	<i>Mean ± SD</i>	<i>ICC (95% CI)</i>	<i>%</i>	<i>Absolute vs %</i>	<i>Absolute vs %</i>
Baseline pinprick sensitivity, NRS (0-100)	31.03 ± 13.45	33.34 ± 12.90	0.62 (0.35 to 0.88)	40.95	8.14 (25.30)	22.57 (70.12)
Electrical detection threshold (mA)	0.19 ± 0.08	0.20 ± 0.08	0.50 (0.19 to 0.72)	42.17	0.06 (29.97)	0.16 (83.07)
Pain during HFS, NRS (0-100)	71.44 ± 10.67	72.91 ± 12.67	0.62 (0.35 to 0.79)	16.23	7.28 (10.08)	20.17 (27.95)
Stress, NRS (0-100)	50.78 ± 23.66	41.44 ± 23.74	0.55 (0.25 to 0.75)	51.40	15.26 (33.09)	42.29 (91.71)
Anxiety, STAI-Y1 (20-80)	32.91 ± 8.83	32.34 ± 8.01	0.66 (0.41 to 0.82)	25.84	4.93 (15.13)	13.68 (41.93)
Catastrophizing SCQ (0-24)	7.13 ± 4.79	7.47 ± 5.76	0.72 (0.49 to 0.85)	72.63	2.85 (39.04)	7.90 (108.22)

Table 4. Between-session comparisons of the secondary outcomes (except heart rate variability indices). The mean ± SD of the first and second session for each of the primary outcomes are shown. Intraclass correlation coefficients (ICC) are reported with their 95% confident intervals and range from 0 to 1. The coefficient of variation (CV) is reported. In addition, standard error of mean (SEM) and minimal detectable change (MDC) are reported in the same unit as the measurements and in percentages relative to the means. HFS, high frequency stimulation; NRS, numerical rating scale; STAI-Y1, State subscale of the State-Trait Anxiety Inventory; SCQ, Situational Catastrophizing Questionnaire.

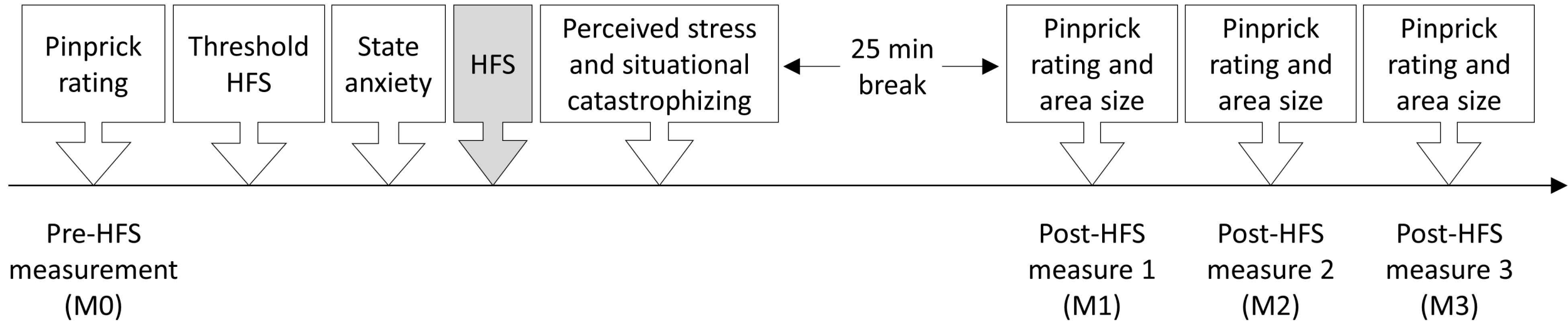
Figure 1. Time-line of the session. HFS, high-frequency electrical stimulation.

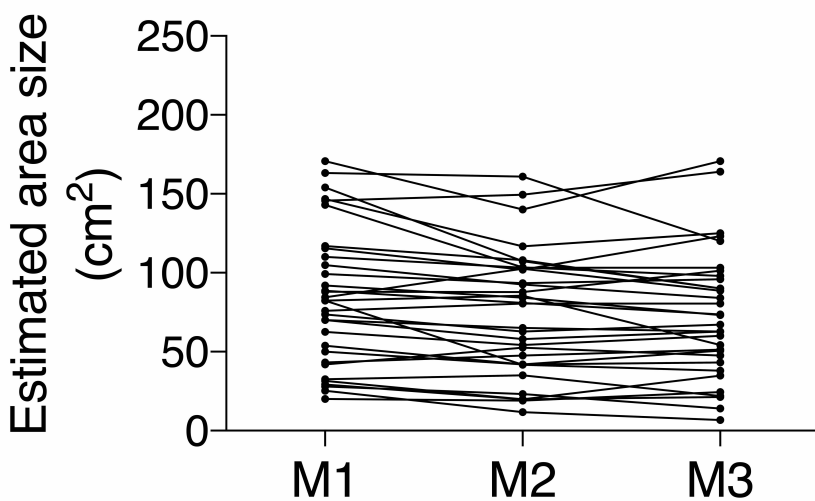
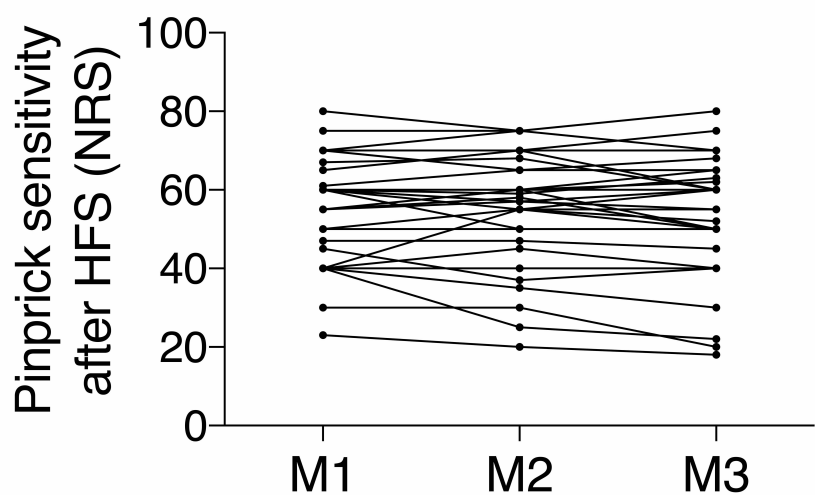
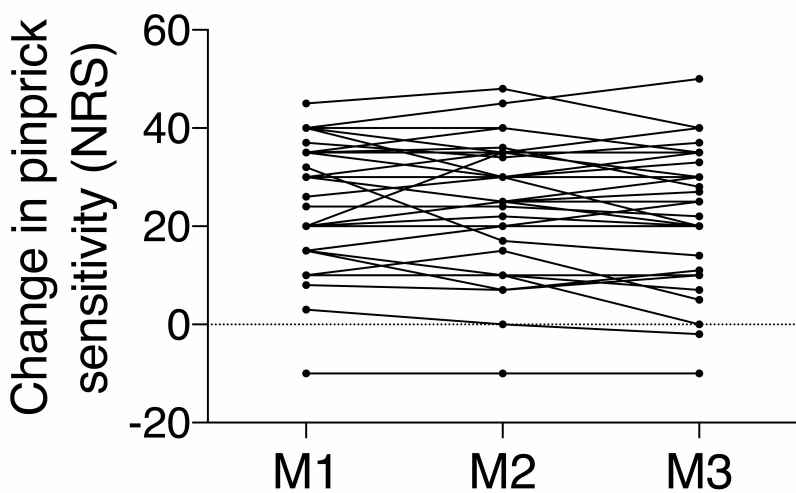
Figure 2. Individual data for the primary outcomes of session 1. The labels on the X-axes (M1, M2 and M3) refer to the first, second and third post-HFS measurements. HFS, high-frequency stimulation; NRS, numerical rating scale.

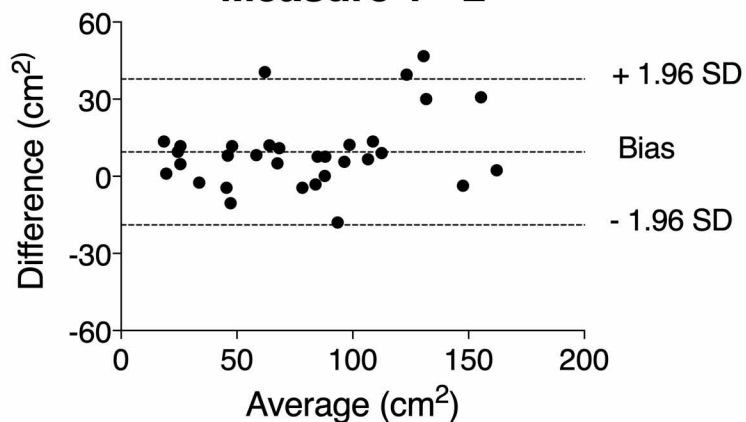
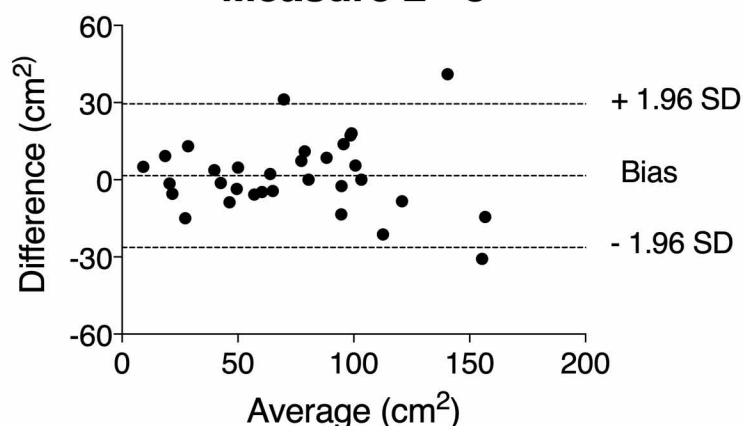
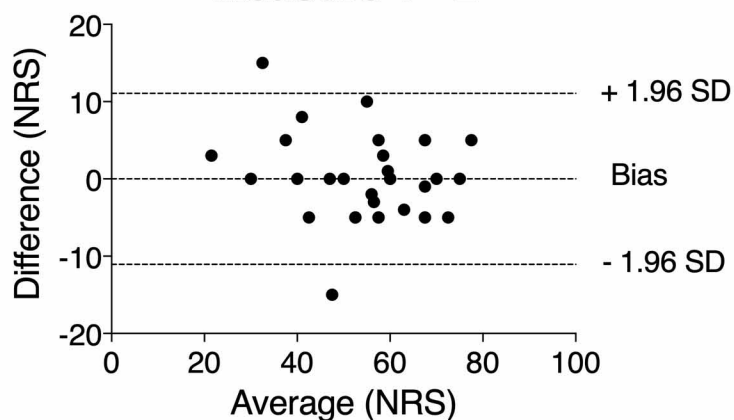
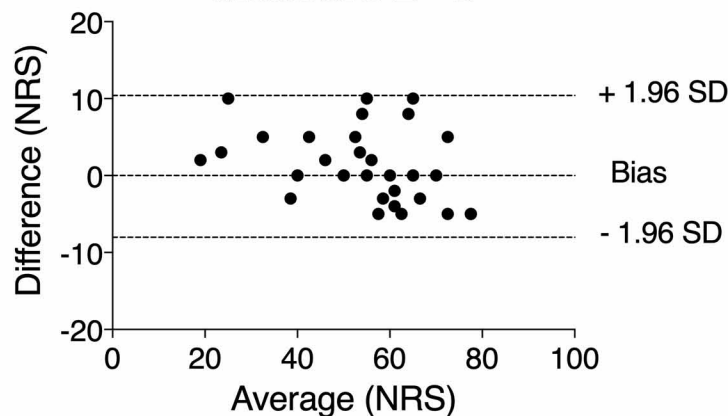
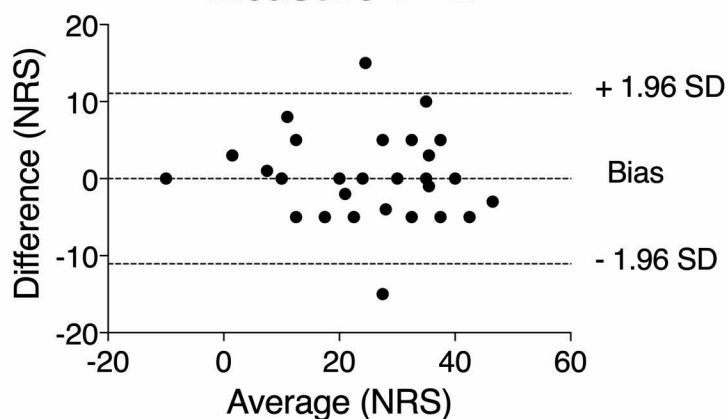
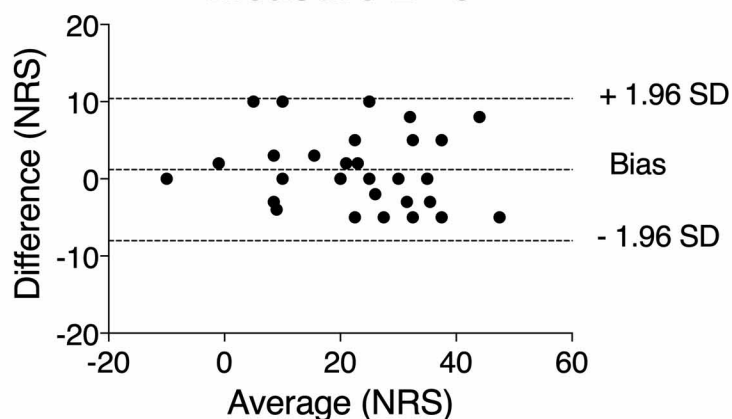
Figure 3. Bland-Altman plots for the within-session comparison of the primary outcomes. The Y-axis shows the difference scores for each subject (left column: measure 1 minus measure 2; right column: measure 2 minus measure 3). Values on the X-axis represent the mean of the two measures for each subject. The dashed line indicates the bias between sessions (the average of the differences), and the limits of agreement, calculated as 1.96 times the standard deviation (SD) of the differences in measurements. HFS, high-frequency stimulation; NRS, numerical rating scale.

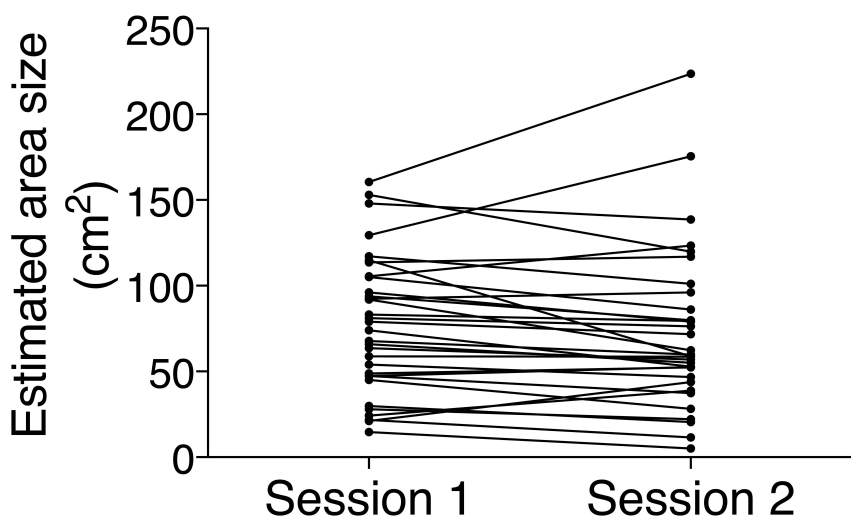
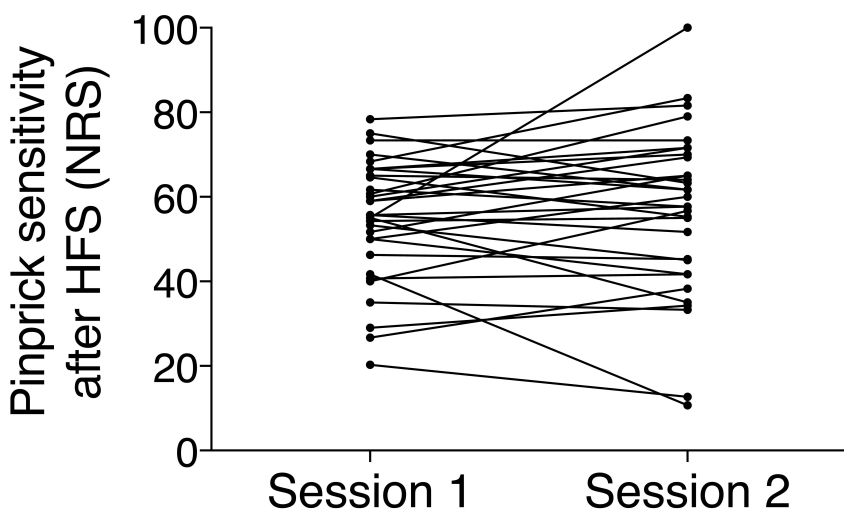
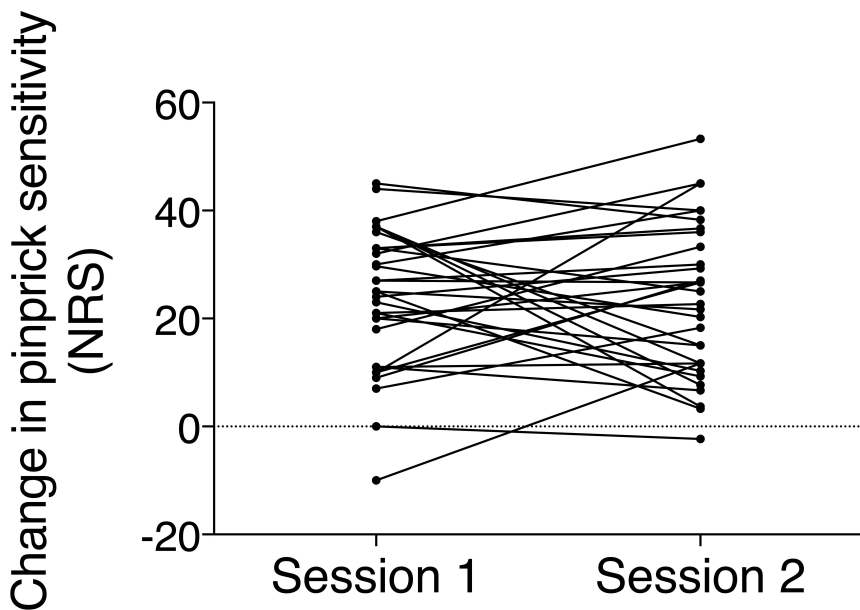
Figure 4. Individual data for the primary outcomes of session 1 and session 2. Each score represents the average of the three post-HFS measurements. HFS, high-frequency stimulation; NRS, numerical rating scale.

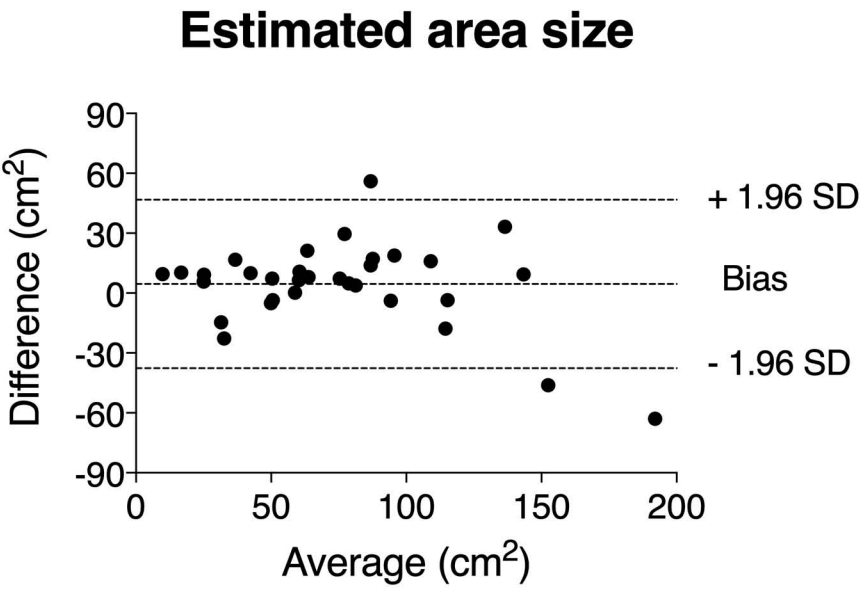
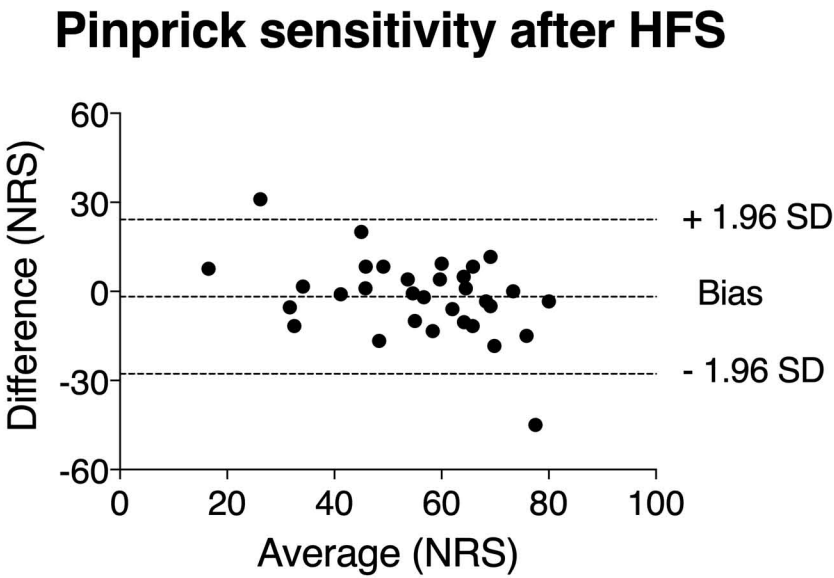
Figure 5. Bland-Altman plots for the between-session comparison of the primary outcomes. The Y-axis shows the difference scores (session 1 minus session 2) for each subject. Values on the X-axis represent the mean of the two measures for each subject. The dashed line indicates the bias between sessions (the average of the differences), and the limits of agreement, calculated as 1.96 times the standard deviation (SD) of the differences in measurements between sessions. HFS, high-frequency stimulation; NRS, numerical rating scale.



A**B****C**

A**Estimated area size
Measure 1 - 2****Estimated area size
Measure 2 - 3****B****Pinprick sensitivity after HFS
Measure 1 - 2****Pinprick sensitivity after HFS
Measure 2 - 3****C****Change in pinprick sensitivity
Measure 1 - 2****Change in pinprick sensitivity
Measure 2 - 3**

A**B****C**

A**B****C**