

Article type : Original Manuscript

Full title:

Assessing thermal sensitivity using transient heat and cold stimuli combined with a Bayesian adaptive method in a clinical setting: a proof of concept study

Running title:

Thermal sensitivity assessed by the psi method

Authors:

A.S. Courtin^a, S. Maldonado Slootjes^{a, e}, G. Caty^{a, b, c, d}, M.P. Hermans^{d, e, f}, L. Plaghki^{a, b}, A. Mouraux^{a, e}

Authors affiliations :

^a Institute of Neuroscience (IoNS), Université catholique de Louvain (UCLouvain), Brussels, Belgium

^b Faculté des sciences de la motricité, Université catholique de Louvain (UCLouvain), Louvain-la-Neuve, Belgium

^c Service de médecine physique et réadaptation, Cliniques universitaires Saint-Luc, Brussels, Belgium

^d Institut de Recherche expérimentale et clinique (IREC), Université catholique de Louvain (UCLouvain), Brussels, Belgium

^e Faculté de médecine et de médecine dentaire, Université catholique de Louvain (UCLouvain), Brussels, Belgium

^f Unité endocrinologie et nutrition, Cliniques universitaires Saint-Luc, Brussels, Belgium

Corresponding author:

Arthur Courtin

Institute of Neuroscience (IONS), UCLouvain

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.1628](https://doi.org/10.1002/ejp.1628)

This article is protected by copyright. All rights reserved

53 Avenue Mounier (B1.53.04), B1200, Brussels, Belgium

arthur.courtin@uclouvain.be

<https://uclouvain.be/fr/repertoires/arthur.courtin>

+32 2 764 73 86

Category:

Original article

Funding:

This study was supported by an ERC Starting Grant (336130-PROBING-PAIN). The funder was not involved in the study design, data collection, data analysis, manuscript preparation and publication decision.

Conflict of interest:

None of the authors has any conflict of interest to disclose.

Significance:

Current QST protocols provide an incomplete and potentially biased estimate of sensory detection performance. We propose a method that estimates the slope and the threshold of the psychometric function, defining heat and cold sensory detection performance, in only a few minutes. Furthermore, we provide preliminary evidence that combining slope and threshold parameters of cold detection performance leads to a better discriminative ability than relying solely on the threshold.

ABSTRACT

Background: Quantitative sensory testing of thermal detection abilities is used as a clinical tool to assess the function of pain pathways. The most common procedure to assess thermal sensitivity, the “method of limits”, provides a quick but rough estimate of detection thresholds. Here, we investigate the potential of evaluating not only the threshold but also the slope of the psychometric functions for cold and warm detection.

Method: A convenience sample of 15 patients with *diabetes mellitus* (DM) and 15 age-matched healthy controls (HC) was tested. Thirty brief (100 ms) stimuli of each modality were applied to the volar wrist and foot dorsum. Cold and warm stimuli were delivered with a Peltier thermode and a temperature-controlled CO₂ laser, respectively. Stimulus intensities were dynamically selected using an adaptive Bayesian algorithm (psi method) maximizing information gain for threshold and slope estimation. ROC analyses were used to assess the ability of slopes, thresholds and the combination of both to discriminate between groups.

Results: Assessment of the slope and threshold of the psychometric function for thermal detection took about 10 minutes. The ability to detect warmth was not reduced in DM patients as compared to HC. Cold detection performance assessed using slope or threshold parameters separated DM from HC with good discriminative power. Discrimination was further increased when both parameters were used together (93% sensitivity and 87% specificity), indicating that they provide complementary information on patient status.

Conclusion: The psi method may be an interesting alternative to the classical method of limits for thermal QST.

INTRODUCTION

The assessment of thermal detection thresholds allows quantifying negative (hypoesthesia) and positive sensory signs (hyperesthesia) in patients with suspected somatosensory impairment (Mucke et al., 2016). They are increasingly used in the diagnostic workup of small-fiber neuropathies (neuropathies affecting autonomic and thermnociceptive primary afferents)(Devigili et al., 2019; Terkelsen et al., 2017; Tesfaye et al., 2010) and, compared to structural or electrophysiological testing, they can be easily implemented in routine clinical practice because the procedure is non-invasive, inexpensive, and easy to conduct.

There are two main methods to estimate thermal thresholds: the method of limits and the method of levels. In the method of limits, a continuous heating or cooling ramp is applied to the skin at a slow rate (e.g. 1°C/s)(Mucke et al., 2016) and the patient is instructed to press a button as soon as a given sensation is felt (Fruhstorfer et al., 1976; Rolke et al., 2006). The temperature reached at the moment the patient presses the button is taken as measure of the detection threshold. In the method of levels or constant stimuli, a number of stimulus intensities are presented and the subject is asked after each stimulation to report whether it was felt or not (Mucke et al., 2016).

Choosing the method to assess thermal detection performance is determined by two main factors: the precision of the information it provides about the psychometric function (relationship between stimulus intensity and detection probability; Fig 1E) and the time required to perform the assessment. The method of limits requires only a few stimuli but provides a crude estimate of the threshold. In contrast, the classical method of levels requires (many) more stimuli but allows to assess both the threshold and slope (informative on the signal to noise ratio in the neural system) of psychometric functions.

Assessing detection performance with only a few stimuli is important in clinical practice, especially when several sensory modalities are explored such as during quantitative sensory testing. The traditional choice of clinicians has been to favor short testing time over precise estimates (method of limits over classical method of levels).

The development of adaptive “method of levels” algorithms (taking into account responses to previous stimuli to select next stimulus intensity) has led to a drastic reduction in the number of trials necessary to obtain a given degree of precision of the estimates. The psi method is such an algorithm, developed by Kontsevich and Tyler (1999) to optimize assessment of the

slope and threshold of psychometric functions. It selects stimulus intensities to maximize information gain (using a Bayesian approach). According to Kingdom and Prins (2010), it is “arguably the most efficient of the adaptive methods”.

Here, we implement the psi method to assess heat and cold detection performances, evaluate the applicability of this new method in a clinical setting, and assess the added value of estimating slope in addition to threshold parameters of the psychometric function by comparing the estimated parameters in a set of patients with *diabetes mellitus* and a group of age-matched healthy control subjects.

RESEARCH DESIGN AND METHODS

Subjects

Fifteen type 2 *diabetes mellitus* patients (DM; 9 males & 6 women; mean age 55 ± 4 years; mean height 1.69 ± 0.12 m) were selected at random from a large database of outpatients followed at least twice a year at the diabetology clinic of Cliniques universitaires Saint Luc (Brussels, Belgium). Diabetes duration was 13.5 ± 3.0 years (mean $\pm$ SD; range: 10-19 years). HbA1c was $7.4 \pm 1.1\%$ (range: 6.0-10.1%). None had spontaneous complaints of neuropathy, with a Toronto Clinical Neuropathy Score (TCNS) ranging from 0 to 6 (Bril and Perkins 2002). The items evaluating pain symptoms, temperature perception, and pinprick perception were normal in all patients but one (who did not normally feel pinprick). Detailed TCNS are available in Table S1. Fifteen age-matched healthy controls were recruited (nine women; age 53 ± 4 years). No participant had a history of neurological or major psychiatric conditions. Written informed consent was obtained from all participants. The study was carried out in accordance with the latest version of the Declaration of Helsinki (October 2013 revision) and was approved by the local Ethics Committee (CE: 2015/30NOV/652).

Nerve conduction studies

All patients underwent a routine nerve conduction study (NCS) assessing large-fiber function. The recordings were conducted using a Viking™ EDX EMG/EP system (Natus, Pleasanton, The USA). Three motor nerves were studied on the right side of the body: the common peroneal nerve (recorded from the *extensor digitorum brevis*, stimulated at the ankle and at the head of the fibula), the tibial nerve (recorded from the *abductor hallucis*, stimulated at the ankle and the popliteal fossa), and the median nerve (recorded from the *abductor pollicis brevis*, stimulated at the elbow and at the wrist 7 cm proximally to the

recording electrode). Similarly, three sensory nerves were studied on the right side of the body: the sural nerve (stimulated at the calf, recorded at the ankle), the ulnar nerve (stimulated at the wrist, recorded at the little finger), and the radial nerve (stimulated at the forearm, recorded at the anatomical snuffbox). The reference values were taken from Preston and Shapiro (2013) and were not corrected for age (all patients were between 18-65 years of age for which the reference values were established) or height.

Heat and cold stimulation

Brief pulses of radiant heat were delivered to the skin using an infrared CO₂ laser stimulator (LSD, SIFEC, Belgium) under radiometer feedback control allowing to deliver a stepped increase in skin temperature (ramps >1000°C/s) during 100 ms over a surface area of 28 mm². Brief cold stimuli were delivered to the skin using a contact thermode using micro-Peltier elements (TCS, QST.Lab, France) designed to deliver stepped decreases in skin temperature (ramp of 300°C/s) during 100 ms over a surface area of 38 mm².

The psi method

The psi method is a Bayesian adaptive algorithm that associates each potential value of slope and threshold with a probability, updates this probability distribution based on the subject's response, and selects the next stimulus intensity so that the response to its presentation maximizes the reduction of the uncertainty of the estimated slope and threshold (see Fig. 1) (Kontsevich and Tyler 1999).

The method, implemented in the open-source MATLAB toolbox Palamedes (<http://www.palamedestoolbox.org>), was used to estimate the thresholds as well as the slopes of the heat and cold detection psychometric functions. The shape of the psychometric function was assumed to be a logistic (as given by Equation 1).

$$p = \frac{1}{1 + e^{-\beta(x-\alpha)}} \quad (1)$$

p is the probability of detection, α is the threshold (*i.e.* the stimulus intensity associated with a $p=0.5$ probability of detecting the stimulus), β is the slope of the psychometric function, and x is the stimulus intensity in °C.

Because the TCS was not able to heat the skin and cold detection thresholds were expected to be close to baseline skin temperature, the ranges for potential threshold values and cold

stimulus intensity were expressed as deviation from baseline skin temperature (from 0°C to -19°C relative to baseline skin temperature). Baseline skin temperature was obtained by averaging three temperature measurements obtained with an infrared thermistor (Tempett, SENSELab, Sweden) at three different sites in the area of stimulation. Conversely, potential heat detection threshold values and heat stimuli range were expressed in absolute values (from 36°C to 60°C). Heat detection threshold values were subsequently transformed to be expressed as deviation from baseline skin temperature. Potential slope values considered by the algorithm ranged from 0.001 to 6 for heat detection and from 0.2 to 5 for cold detection. The guess and lapse rate were set to 0, given the small number of stimulations.

Each run of the psi method consisted of 30 stimuli. This number was chosen as it appeared to be a good tradeoff between estimation precision, duration, and risk of habituation/sensitization as well as fatigue/distraction due to repeated testing (and therefore experimentally induced change of threshold/slope). Based on our experience with the recording of scalp laser evoked potentials, patients can provide reliable reaction times and intensity ratings when such a number of stimuli is used. Whether this number was appropriate is addressed in the Results and Discussion sections.

Separate runs for laser and cold stimulation were applied in random order at the left or right volar wrist and the contralateral foot dorsum. A run lasted maximum 10 minutes. Between each stimulus, the laser beam or contact thermode was slightly displaced to avoid receptor sensitization and/or habituation.

Data analysis

Estimated values of the slope of the psychometric function were transformed into the more intuitive measure of spread (in °C) using equation 2. The spread of the psychometric function was defined as the difference, in °C, between the intensity of stimulation associated with a $p=0.1$ and $p=0.9$ probability of detecting the stimulus.

$$\sigma = 2 \frac{\ln\left(\frac{0.9}{0.1}\right)}{\beta} \quad (2)$$

σ is the spread, β is the slope

The discrimination performance of sensory detection parameters (threshold, spread and both) was estimated by Receiver Operating Characteristic (ROC) analysis based on classification

with logistic regression. The best cut-off to discriminate between patients and controls was determined based on the optimal operating point of the ROC curve. Area under the ROC curves (AUC) are reported as *mean [95% Confidence Interval]*. These analyses were conducted using MATLAB (The MathWorks, the USA). Additionally, to compare the ability of different models (including threshold, slope or both) to explain the data, Bayes Factor (BF) were obtained to compare these models, using the *glib* function from the *BMA* package in the free statistical software R. As recommended by Raftery (1996) several plausible values (the default 1, 1.65, and 5) were investigated for the ϕ parameter of the prior. The values reported in this paper correspond to $\phi=1.65$.

To better characterize the collected data, the effect of the group (patient vs control) and the recording site (wrist vs foot dorsum) were investigated. Given that not all the data followed a normal or log-normal distribution, we decided to use a non-parametric Bayesian independent sample t-test (based on Mann-Whitney test) to assess the effect of the group (patients with DM vs. HC) on the different variables and a non-parametric Bayesian paired sample t-test (based on Wilcoxon signed rank-test) to assess the effect of the recording site (volar wrist vs. foot dorsum), separately for each group or site. These analyses were conducted using the free open source software JASP (v0.12.2.) (Wagenmakers et al., 2018a).

The relationship between (1) age, diabetes duration, BMI, HbA1c, Toronto clinical score and number of abnormal NCV (0 to 6 abnormal NCVs) and (2) the slope and thresholds for cold and heat detection at the forearm and foot dorsum were assessed using Pearson's rho (with 95% credible interval, the Bayesian counterpart of frequentist 95% confidence intervals). These analyses were conducted using JASP.

The Bayes factor (BF) is the ratio between the likelihood of two hypothesis and therefore quantifies the support given by the data to H_1 compared to that given to H_0 (in the case of BF_{10}). A $BF_{10}>0$ indicates more evidence in favor of H_1 whereas a $BF_{10}<0$ includes more evidence for H_0 . Jeffreys (1998) proposed the following way, based on 10 to the power of multiples of 0.5, to interpret BF in favor of H_1 : $0<BF<3.2$ barely worth mentioning, $BF>3.2$ moderate support, $BF>10$ strong support, $BF>32$ very strong support, $BF>100$ decisive support. A practical introduction to Bayes Factor in Wagenmakers et al., (2018b) and Wagenmakers et al., (2018a).

RESULTS

Group differences

The average baseline skin temperatures were similar for both groups (wrist: 33.1°C [31.6°C; 34.3°C], foot: 31.0°C [29.6°C; 32.4°C]; *wrist* $BF_{10} = 0.360$; *foot* $BF_{10} = 0.404$) and lower at the foot than at the wrist (*control* $BF_{10} = 125.374$; *patient* $BF_{10} = 23.735$).

Warm detection thresholds were similar between patients and controls (*wrist* $BF_{10} = 0.707$; *foot* $BF_{10} = 0.902$) but larger at the foot (+15.6°C [+13.8°C; +18.0°C] relative to baseline skin temperature) than at the wrist (+9.7°C [+8.8°C; +11.4°C]; *control* $BF_{10} = 211.831$; *patient* $BF_{10} = 761.295$). The spread of the psychometric function for warm detection (3.1°C [2.5°C; 5.1°C]) did not differ between groups (*wrist* $BF_{10} = 0.467$; *foot* $BF_{10} = 0.344$), or between stimulated limbs (*control* $BF_{10} = 0.647$; *patient* $BF_{10} = 0.420$).

In contrast, cold detection thresholds were increased in patients (wrist: -2.4°C [-1.4°C; -2.9°C], foot: -8.0°C [-5.2°C; -10.1°C]) compared to controls (wrist: -1.0°C [-1.0°C; -2.0°C], foot: -3.7°C [-2.1°C; -5.7°C]; *wrist* $BF_{10} = 11.091$; *foot* $BF_{10} = 5.744$) and were greater at the foot dorsum than at the wrist (*control* $BF_{10} = 266.858$; *patient* $BF_{10} = 157.461$). The spread of the corresponding function was also larger in patients (wrist: 3.9°C [2.8°C; 6.1], foot: 11.9°C [5.0; 16.6]) than in controls (wrist: 2.1°C [1.7; 2.6], foot: 4.3°C [2.5°C; 4.6°C]; *wrist* $BF_{10} = 4.671$; *foot* $BF_{10} = 10.587$) and larger at the foot dorsum than at the wrist (*control* $BF_{10} = 4.878$; *patient* $BF_{10} = 62.004$).

Discrimination performance

Based on the Receiver Operating Characteristic (ROC) analysis, warm detection thresholds did not discriminate between patient and control groups, neither when stimulating the wrist (AUC: 0.65 [0.41; 0.85]), nor when stimulating the foot (AUC: 0.67 [0.46; 0.84]). The spread of the psychometric function for warm detection was also uninformative (AUC wrist: 0.59 [0.36; 0.80]; AUC foot: 0.50 [0.29; 0.73]). In contrast, both the cold detection threshold (AUC wrist: 0.83 [0.62; 0.94], AUC foot: 0.81 [0.61; 0.93]) and the spread of the psychometric function for cold detection (AUC wrist: 0.82 [0.59; 0.94]; AUC foot: 0.84 [0.62; 0.96]) displayed good discriminative properties. When including both the slope and threshold in the ROC analysis, warm detection was still uninformative (AUC wrist: 0.67 [0.40; 0.84], AUC foot: 0.66 [0.44; 0.84]) whereas cold detection discrimination performance between patients and controls was further increased (AUC wrist: 0.89 [0.60; 0.99], AUC foot: 0.94 [0.78; 1.00]).

Comparing the regression models using BF yielded similar results. It appeared that the models including both the slope and threshold were best at explaining the data for cold detection at the wrist (*threshold* $BF_{10}=10.432$, *slope* $BF_{10}=65.772$, *slope & threshold* $BF_{10}=209.937$) and at the foot dorsum (*threshold* $BF_{10}=37.070$, *slope* $BF_{10}=171.350$, *slope & threshold* $BF_{10}=3686.104$). Conversely, the null model appeared to be the best model to explain the data for warm detection at the wrist (*threshold* $BF_{10}=0.310$, *slope* $BF_{10}=0.104$, *slope & threshold* $BF_{10}=0.035$) and at the foot dorsum (*threshold* $BF_{10}=0.583$, *slope* $BF_{10}=0.104$, *slope & threshold* $BF_{10}=0.062$). In the case of cold detection, the slope appeared to be more efficient at explaining the data (*wrist* $BF_{\text{slope / threshold}}=6.305$, *foot* $BF_{\text{slope / threshold}}=4.622$) and the model including both parameters outperformed both models including only one parameter (*wrist* $BF_{\text{slope & threshold / threshold}}=20.125$, *foot* $BF_{\text{slope & threshold / threshold}}=99.435$, *wrist* $BF_{\text{slope & threshold / slope}}=3.192$, *foot* $BF_{\text{slope & threshold / slope}}=21.512$).

Association between patient characteristics and psychophysical variables

There was no clear linear relationship between indicators of diabetes severity (DM duration, Toronto Clinical Neuropathy score, HbA1C, number of abnormal nerve conduction velocity and BMI) and the values of threshold and slope estimated with the psi method (all Pearson's Rho $BF_{10}<10$; see also Figure 3).

Asymptotic convergence of the psychometric estimates

The psi method provides estimates of the slope and the threshold after each stimulation. Once the threshold estimate had converged (usually within 5 to 10 trials), the change of the estimate value and the associated standard error tended to decrease as a function of the number of applied stimuli (see Figure 4).

DISCUSSION

The present study shows that the psi method can estimate both parameters (threshold and slope) of the psychometric function for cold and warm detection in a time-effective way (a few minutes *per* test) and suggests that the threshold and slope parameters of the psychometric function for cold detection are both affected by the preclinical neuropathic changes in DM patients.

Relevance of estimating the slope of the psychometric function

In order to assess the clinical relevance of estimating the slope/spread of the psychometric function in addition to its threshold, we compared the values obtained for cold and warm detection at the wrist and foot dorsum in a group of DM patients without clinically evident symptoms of small fiber neuropathy and in a group of age-matched healthy volunteers. Threshold and slope parameters of warm detection performance did not differ between groups. In contrast, both the threshold and the slope of the psychometric function of cold detection were altered in DM patients as compared to HC. ROC analysis showed that both slope and threshold estimates for cold detection were good markers to distinguish patients from controls and that the discrimination performance was even better when using both estimates rather than only one, which suggests that they provide complimentary information. These observations were confirmed when analyzing the BF for the logistic regression corresponding to the ROC analysis.

Therefore, our study provides preliminary evidence that the slope of the psychometric function for cold detection provides valuable information on the pathological/physiological status of the thermonociceptive system and that estimating this parameter in patients could be clinically relevant.

Interpreting alterations of the slope

An increased spread of the psychometric function could be interpreted as a decreased signal-to-noise ratio in the sensory system (the signal being the stimulus-evoked activity in the sensory system). In our group of patients with DM, alterations of the threshold and slope could be the consequence of a rarefaction of intra-epidermal free nerve endings (Shun et al., 2004). The exact reason why cold detection performance was more discriminating than warm detection performance in our sample remains an open question but is consistent with previous findings in diabetic and HIV-associated neuropathies (Loseth et al., 2008; Sorensen et al., 2006; Zhou et al., 2007). One explanation could be that cold and heat detection are mediated by different types of primary afferents (respectively small myelinated and unmyelinated fibers)(Schepers and Ringkamp 2010) which may react differently to the pathological changes leading to neuropathy. Another possible explanation would be that, given that skin nerve fibers expressing the cold sensitive channel TRPM8 are fewer than that expressing the heat sensitive TRPV1 (Axelsson et al., 2009; Jeon and Caterina 2018; Weyer-Menkhoff et al., 2019), a progressive reduction of intra-epidermal nerve fiber density could more rapidly affect the ability to detect information conveyed by cold vs heat sensitive afferents.

In a clinical setting, sensory detection performance is most often assessed using the method of limits because it allows estimating detection threshold in a very time-efficient manner. This comes at the cost of precision, as this method only provides limited information (rough threshold estimate) on the underlying sensory process. Another limitation of this method is its reliance on reaction times, leading to an overestimation of the threshold due to the time elapsing between the detection and its signaling by a motor response. This would be inconsequential if reaction times were approximately the same in all individuals. However, such assumption is unlikely to be true. First, because age increases reaction times and reaction time variability (Gorus, De Raedt, & Mets, 2006). Furthermore, age and diabetes increase the probability of presenting comorbidities, some of which can lead to a slowing of reaction times due to cognitive or motor impairment. This argument against the method of limits may, however, be anecdotal, as a relatively large increase in reaction time would be needed to significantly affect the results.

In this study, we propose a relatively time-effective method to assess both threshold and slope of the psychometric function. Furthermore, we provide preliminary evidence that the information provided by the slope parameter is complementary to the information provided by the threshold parameter. However, future studies are needed to determine the clinical relevance of this information. First, studies must be conducted to obtain reference values for the parameters obtained using our method. Second, studies should directly compare this method with existing methods in patient samples, to be able to properly assess their respective diagnostic accuracy.

In the absence of such studies, one can only speculate on the relative accuracy of the different methods. Conflicting results have been reported on the diagnostic accuracy of thermal QST conducted using the method of limits and such comparison would be hazardous (Backonja et al., 2013). Importantly, one must keep in mind that the accuracy of a method is dependent upon how the two groups that need to be sorted are defined. In a previous study, some of the authors found a relatively good AUC value (0.840) for cold detection threshold obtained with the method of limits (Rage et al., 2011). However, their case definition criteria (abnormal intraepidermal nerve fiber density) was much stricter than the one used in this study (suffering of DM) and the level of pathological change of cold detection threshold was probably higher than in our patient group. We can therefore reasonably expect that the AUC

for the combination of slope and threshold would have been further increased if it had had to sort these patients (as a reminder, we obtained an AUC for cold detection of 0.89 at the wrist and 0.94 at the foot in the present study).

Finally, two points should be considered before its implementation in daily practice: whether a testing time of a few minutes as compared to a few seconds (method of limits) is compatible with daily clinical practice, and whether the algorithm could be implemented in the software controlling currently available thermal stimulators. If the first one is dependent on local (and therefore various) organizational considerations, the second point seems easy to overcome. *Limitations*

Post-hoc, our results indicate that 30 stimuli is enough to obtain satisfying estimates of the slope and threshold of the psychometric function for heat and cold detection. However, as shown in Figure 4, the slope estimate was not as stable as the threshold estimate. Increasing the number of stimuli could lead to a more precise estimate of the slope. However, it would come at the cost of an increased assessment time, and with the risk that repeated stimulations could affect thermal sensitivity through habituation and/or sensitization. The results of Kontsevich and Tyler (1999) suggest that about 10 times more stimuli are necessary to estimate the slope than the threshold, for a given degree of precision (which would translate to 50 to 100 stimuli in the case of this study).

This study was designed as a proof of concept. Therefore, it assessed only a limited number of subjects and it did not directly compare the different methods to assess thermal sensitivity. Furthermore, in-depth diagnostic work-up was not conducted to fully characterize the potential presence of neuropathy in these patients. The alterations observed in this study could reflect subclinical peripheral diabetic neuropathy. Further studies relating threshold and slope estimates with direct assessments of small fiber innervation density are required to confirm this hypothesis.

Conclusions

Our study proposes the evaluation of heat and cold thermal sensitivity in patients using an adaptive algorithm that allows estimating both the threshold and the slope of the psychometric function with a limited number of stimuli. We show that the method can be implemented in patients and that the information provided by the slope parameter could be complementary to the information provided by the threshold parameter, as combining both

slope and threshold information allows a better discrimination between DM patients without clear symptoms of small fiber neuropathy and healthy volunteers. This may suggest that the method is very sensitive to early preclinical sensory alterations in DM. The results of this proof of concept study need to be confirmed in larger patient studies directly comparing the psi method to the established thermal assessment methods used for quantitative sensory testing.

Author Contributions: A.S.C. performed the data analysis and data interpretation, wrote the manuscript and made the figures. S.M.S., L.P., G.C. and M.P.H. assisted in the study design, the data acquisitions, data analysis, data interpretation, and edited and critically reviewed the manuscript. A.M. obtained funding, assisted in the study design, data analysis, data interpretation, and edited and critically reviewed the manuscript. All authors approved the final manuscript as submitted. A.S.C. is the guarantor of this work and, as such, had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. All authors approved the final version of the manuscript.

REFERENCES

- Axelsson HE, Minde JK, Sonesson A, Toolanen G, Hogestatt ED, Zygmunt PM. Transient receptor potential vanilloid 1, vanilloid 2 and melastatin 8 immunoreactive nerve fibers in human skin from individuals with and without Norrbottnian congenital insensitivity to pain. *Neuroscience* 2009;162: 1322-1332.
- Backonja MM, Attal N, Baron R, Bouhassira D, Drangholt M, Dyck PJ, Edwards RR, Freeman R, Gracely R, Haanpaa MH, Hansson P, Hatem SM, Krumova EK, Jensen TS, Maier C, Mick G, Rice AS, Rolke R, Treede RD, Serra J, Toelle T, Tugnoli V, Walk D, Walace MS, Ware M, Yarnitsky D, Ziegler D. Value of quantitative sensory testing in neurological and pain disorders: NeuPSIG consensus. *Pain* 2013;154: 1807-1819.
- Bril V and Perkins BA. Validation of the Toronto Clinical Scoring System for diabetic polyneuropathy. *Diabetes care* 2002;25: 2048-2052.
- Devigili G, Rinaldo S, Lombardi R, Cazzato D, Marchi M, Salvi E, Eleopra R, Lauria G. Diagnostic criteria for small fibre neuropathy in clinical practice and research. *Brain : a journal of neurology* 2019.
- Fruhstorfer H, Lindblom U, Schmidt WC. Method for quantitative estimation of thermal thresholds in patients. *Journal of neurology, neurosurgery, and psychiatry* 1976;39: 1071-1075.
- Jeffreys H. *The theory of probability* OUP Oxford. 1998.
- Jeon S and Caterina MJ. Molecular basis of peripheral innocuous warmth sensitivity. *Handbook of clinical neurology* 2018;156: 69-82.
- Kingdom FA and Prins N. *Psychophysics: A practical introduction*. Academic Press London; 2010.
- Kontsevich LL and Tyler CW. Bayesian adaptive estimation of psychometric slope and threshold. *Vision research* 1999;39: 2729-2737.
- Loseth S, Stalberg E, Jorde R, Mellgren SI. Early diabetic neuropathy: thermal thresholds and intraepidermal nerve fibre density in patients with normal nerve conduction studies. *Journal of neurology* 2008;255: 1197-1202.
- Mucke M, Cuhls H, Radbruch L, Baron R, Maier C, Tolle T, Treede RD, Rolke R. Quantitative sensory testing (QST). English version. *Schmerz (Berlin, Germany)* 2016.

- Preston DC and Shapiro BE. 26 - Polyneuropathy. In: Electromyography and Neuromuscular Disorders (Third Edition). London: W.B. Saunders; 2013; 384-416.
- Raftery AE. Approximate Bayes factors and accounting for model uncertainty in generalised linear models. *Biometrika* 1996;83: 251-266.
- Rage M, Van Acker N, Knaapen MW, Timmers M, Streffer J, Hermans MP, Sindic C, Meert T, Plaghki L. Asymptomatic small fiber neuropathy in diabetes mellitus: investigations with intraepidermal nerve fiber density, quantitative sensory testing and laser-evoked potentials. *Journal of neurology* 2011;258: 1852-1864.
- Rolke R, Magerl W, Campbell KA, Schalber C, Caspari S, Birklein F, Treede RD. Quantitative sensory testing: a comprehensive protocol for clinical trials. *European journal of pain (London, England)* 2006;10: 77-88.
- Schepers RJ and Ringkamp M. Thermoreceptors and thermosensitive afferents. *Neuroscience and biobehavioral reviews* 2010;34: 177-184.
- Shun CT, Chang YC, Wu HP, Hsieh SC, Lin WM, Lin YH, Tai TY, Hsieh ST. Skin denervation in type 2 diabetes: correlations with diabetic duration and functional impairments. *Brain : a journal of neurology* 2004;127: 1593-1605.
- Sorensen L, Molyneaux L, Yue DK. The relationship among pain, sensory loss, and small nerve fibers in diabetes. *Diabetes care* 2006;29: 883-887.
- Terkelsen AJ, Karlsson P, Lauria G, Freeman R, Finnerup NB, Jensen TS. The diagnostic challenge of small fibre neuropathy: clinical presentations, evaluations, and causes. *The Lancet Neurology* 2017;16: 934-944.
- Tesfaye S, Boulton AJ, Dyck PJ, Freeman R, Horowitz M, Kempler P, Lauria G, Malik RA, Spallone V, Vinik A, Bernardi L, Valensi P. Diabetic neuropathies: update on definitions, diagnostic criteria, estimation of severity, and treatments. *Diabetes care* 2010;33: 2285-2293.
- Wagenmakers EJ, Love J, Marsman M, Jamil T, Ly A, Verhagen J, Selker R, Gronau QF, Dropmann D, Boutin B, Meerhoff F, Knight P, Raj A, van Kesteren EJ, van Doorn J, Smira M, Epskamp S, Etz A, Matzke D, de Jong T, van den Bergh D, Sarafoglou A, Steingroever H, Derks K, Rouder JN, Morey RD. Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic bulletin & review* 2018a;25: 58-76.
- Wagenmakers EJ, Marsman M, Jamil T, Ly A, Verhagen J, Love J, Selker R, Gronau QF, Smira M, Epskamp S, Matzke D, Rouder JN, Morey RD. Bayesian inference for

psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic bulletin & review* 2018b;25: 35-57.

Weyer-Menkhoff I, Pinter A, Schlierbach H, Schänzer A, Lötsch J. Epidermal expression of human TRPM8, but not of TRPA1 ion channels, is associated with sensory responses to local skin cooling. *Pain* 2019;160: 2699-2709.

Zhou L, Kitch DW, Evans SR, Hauer P, Raman S, Ebenezer GJ, Gerschenson M, Marra CM, Valcour V, Diaz-Arrastia R, Goodkin K, Millar L, Shriver S, Asmuth DM, Clifford DB, Simpson DM, McArthur JC. Correlates of epidermal nerve fiber densities in HIV-associated distal sensory polyneuropathy. *Neurology* 2007;68: 2113-2119.

Figure 1- Example of a run of the method for cold detection assessment

a. Stimulus placement and response for each trial. The stimulus intensity (expressed in °C relative to baseline skin temperature) is chosen by the algorithm, based on the current probability distribution for the threshold and slope, in order to maximize entropy (i.e. uncertainty) reduction. **b.** Evolution across trials of the probability density associated with each potential threshold and slope value (stimulus presentation). The probability distribution is updated after each trial based on the response of the subject (detected/not detected). The first colormap corresponds to the probability distribution before the first trial, the last colormap corresponds to the probability distribution after the last trial. **c.** Marginalized (*i.e.* reduced to one dimension) probability distribution for the threshold before the first trial (blue) and after the last trial (red). **d.** Marginalized probability distribution for the slope before the first trial (blue) and after the last trial (red). **e.** The initial (starting point of the algorithm) and final fitted psychometric functions.

Figure 2– Results obtained in a convenience sample of *diabetes mellitus* patients and age-matched healthy controls

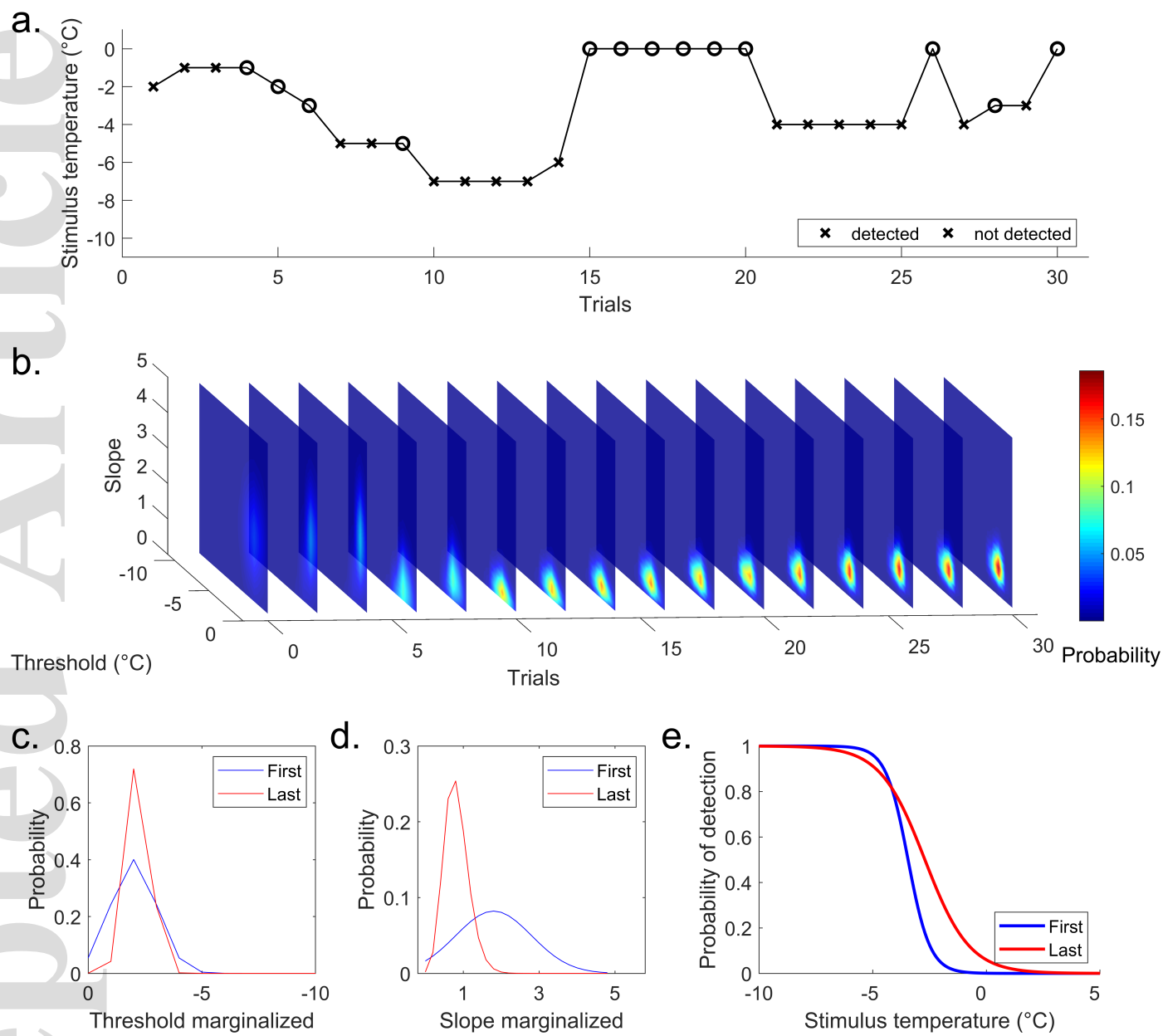
a. & c. Group level average psychometric functions (PF) for the detection of phasic contact cold stimuli and radiant laser heat stimuli applied to the volar wrist (first and second graphs) and foot dorsum (third and last graph) in healthy controls (dotted line) and diabetic patients (full line). The thresholds (vertical arrows) represent the skin temperature at which the subject detects 50% of the stimuli. The spread of the PF corresponds to the width of the transition from a detection probability of 0.1 to a detection probability of 0.9. **b. & d.** Scatter plots of the spread and threshold (expressed relative to baseline skin temperature) estimated in each healthy control (open circles) and diabetic patient (filled circles). The larger crossed markers represent the average values for each group. The black dotted lines represent the best cut-off for discriminating diabetic patients (upper right) from healthy controls (lower left) according to the ROC analysis based on the bivariate (slope and threshold) logistic regression.

Figure 3 – Scatter plot of the cold detection estimates of threshold and spread/slope and clinical severity markers of DM.

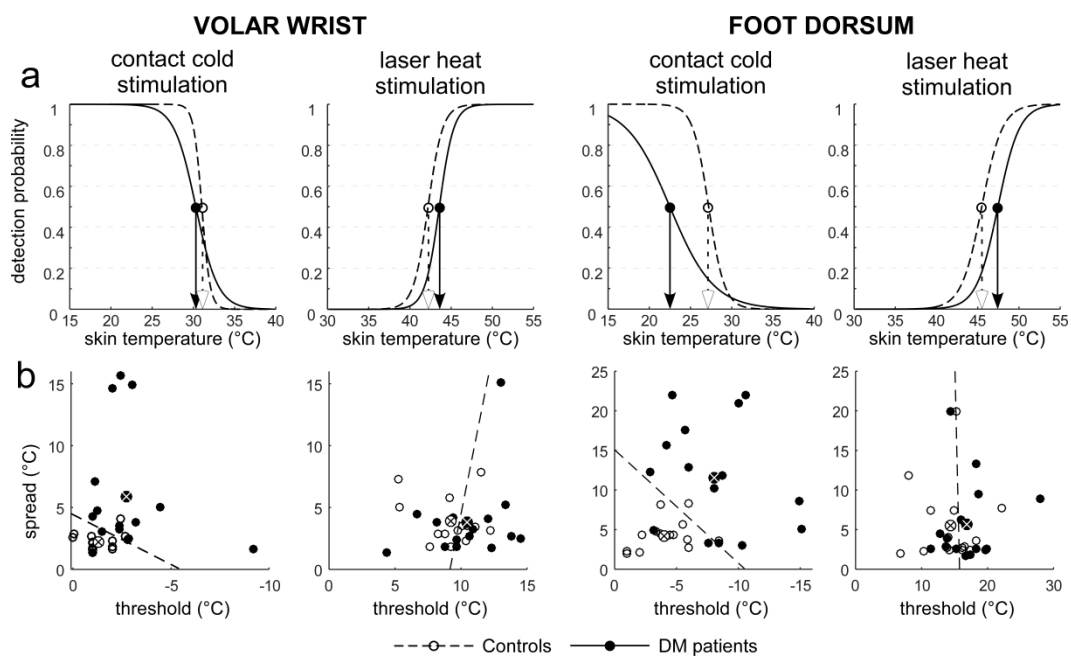
DM Duration: time since diabetes mellitus diagnosis, y: years, BMI: body mass index, TCNS: Toronto Clinical Neuropathy Score, NCV: nerve conduction velocity. Three motor (common peroneal, tibial, median) and three sensory nerves (sural, ulnar, radial) were tested. No clear linear relationship appeared between cold detection parameters and markers of DM severity.

Figure 4 – Asymptotic convergence in the estimation procedure of the threshold and the slope

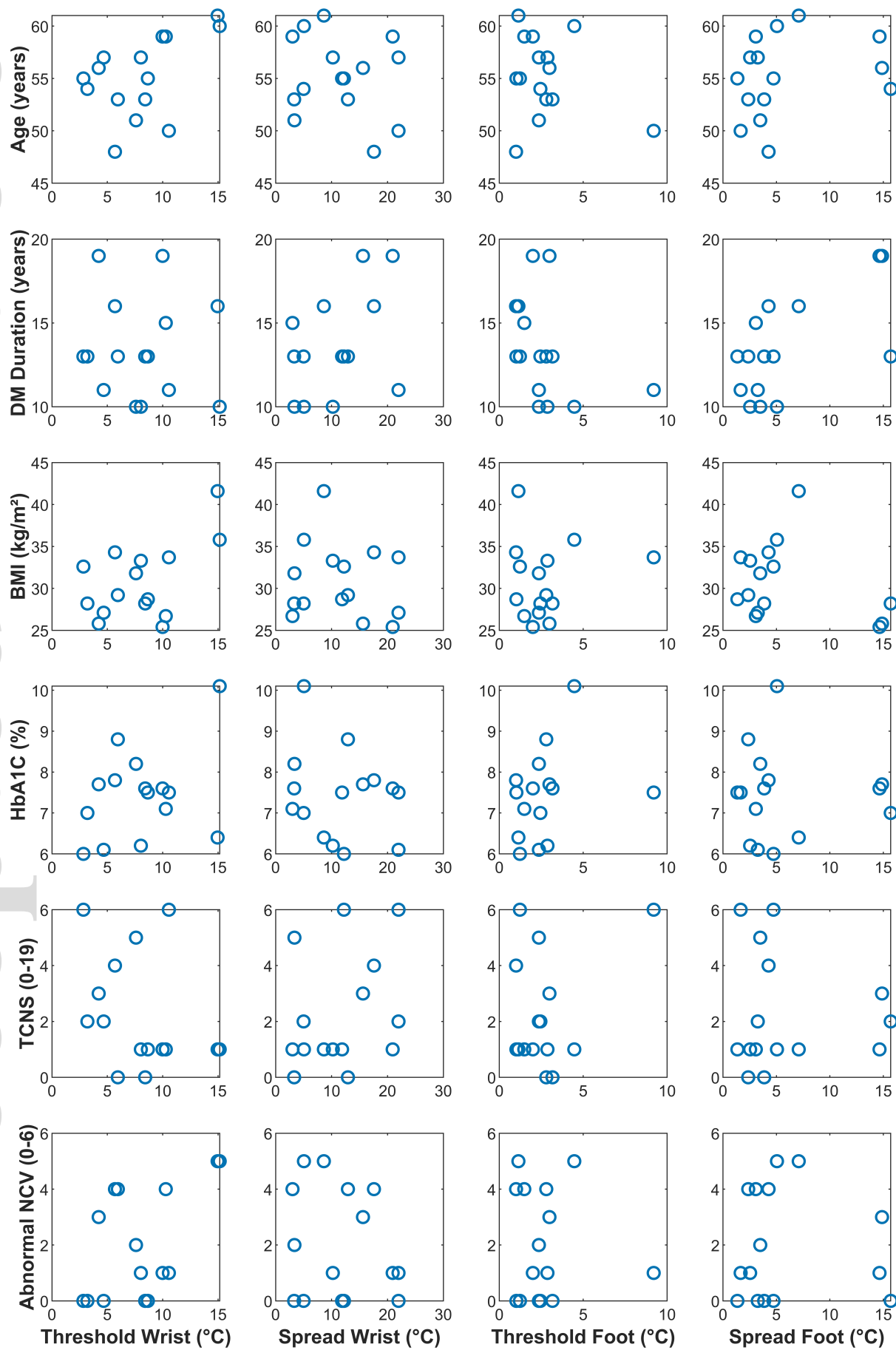
Median ($n = 30$) of the differences between the successive estimates of the threshold (a.; in $^{\circ}\text{C}$) and the slope (b.; in $^{\circ}\text{C}^{-1}$) and the median of the differences of their successive standard errors (SE; c. & d.) obtained by subtracting the current estimate from the previous one. The threshold estimates converged within the 5 – 10 first trials and was then refined (reduction of the SE). Once the threshold is estimated with sufficient accuracy (after 5-10 trials), the transition range was sampled to estimate the slope of the PF (change of the estimate and subsequent decrease in the associated SEs)

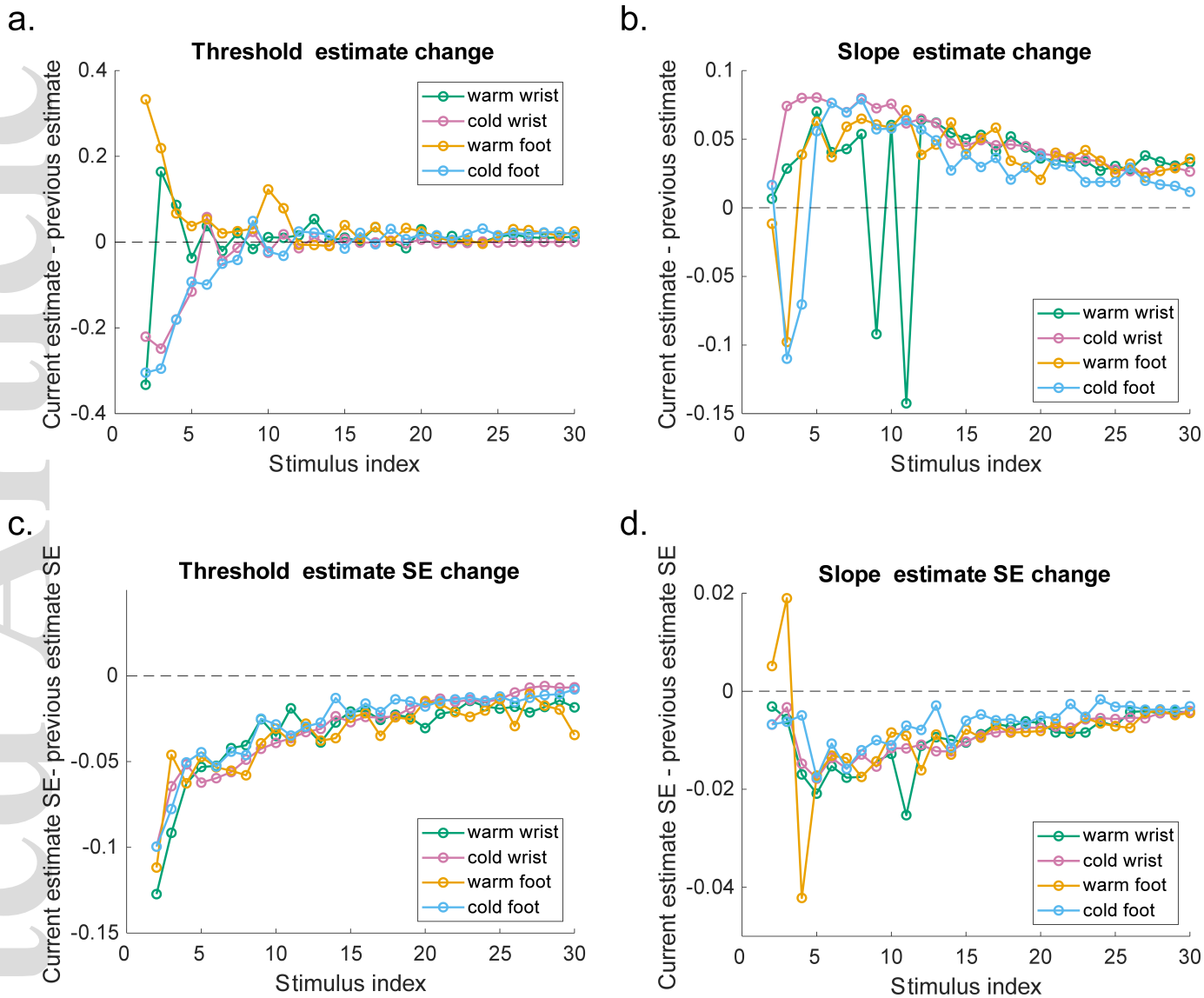


ejp_1628_f1.tif



ejp_1628_f2.tif





ejp_1628_f4.tif