Master's Thesis

Master in Physiotherapy
May 2020

What's the difference?

Normal limits for within-subject comparisons of thermal thresholds

Name: Øystein Dunker

Course code: MAFYS5900

Word count: 9851

Faculty of Health Sciences

OSLO METROPOLITAN UNIVERSITY
STORBYUNIVERSITETET

Forord

Denne oppgaven er mitt første lille skritt inn i forskningsverdenen. Den ble skrevet i en spesiell tid, og er sådan tidvis et produkt av relativt strenge prioriteringer. Kanskje var det heldig at jeg fortsatt var en forholdsvis selvisolert student da SARS-COV-2 kom til Norge og samfunnet stengte ned – på mange måter toget livet videre i samme velkjente rutine. Likevel har det vært noen utfordringer, som på én side har stjålet både tid og oppmerksomhet, men på den annen side har manet til fokus og effektivitet da vinduet for jobbing var åpent. Når det er sagt, skal jeg forsøksvis unngå samtidig syklisk søvndeprivasjon, globale pandemier og totalrenovasjon av en hel etasje *neste gang* jeg skriver en lengre oppgave.

Jeg vil gjerne takke FORMI ved Ullevål Sykehus for at de tok i mot en nybegynner i forskningsverdenen med åpne armer. Marie Skovli Pettersen, Sara Maria Allen, og spesielt biveileder Marie Udnesseter Lie har hjulpet meg ekstra mye, med opplæring på smertelaboratoriet, inklusjon og rekruttering til studiet, tips, råd og gode svar da behovene meldte seg, m.m.

Jeg vil rette en egen stor takk til hovedveilederen på oppgaven, Kristian Bernhard Nilsen, som har hjulpet meg med å holde driv gjennom oppgaven, og på uforklarlig vis har klart å ha god tid til jevne veiledninger. Kritiske tilbakemeldinger som går rett til kjernen av analysemetoder, argumentasjon eller bare enkeltsetninger, er høyt verdsatt, og jeg har vært bortskjemt som har fått slike også i helger.

Familien min fortjener en ordentlig takk. Siri har strukket seg langt for at jeg kunne gjennomføre eksperimenter på ettermiddager og kvelder i høst og vinter da det trengtes; Hun har latt meg sove ut de dagene det var viktig at jeg fikk jobbet, og har funnet på små utflukter med Eirill i vårhelgene for å gi meg tid og rom. Eirill kan ikke skryte av å ha bidratt like direkte, da ukesvis med planlagt skriving plutselig ble byttet ut med hyggelig pappa-datter-tid – Men ingen har vært en så standhaftig pådriver av moral og smittsomt godt humør som henne (medianhumør: strålende), og jeg drister meg til å påstå at det er enklere for de fleste å jobbe analytisk og kreativt når en er i godt humør!

Norwegian abstract

Bakgrunn: Kvantitativ termotesting er en psykofysisk funksjonstest av de tynne nervefibrene og tilhørende sentralnervøse baner, som er avhengig av referansemateriale for å vurdere normalitet. Normalgrenser for intra-subjekt sammenlikninger av termoterskler er mangelvare, og deres assosiasjon med alder, høyde og kjønn er ikke godt nok kartlagt. Målet med studien var å undersøke relative normalgrenser for termoterskler i underekstremitetene og thenar eminens, spesifikt kontralateralt homologe– eller distal-proksimale sammenlikninger, samt. å vurdere hvorvidt disse er asossiert med alder, høyde eller kjønn.

Metode: Førtiåtte frivillige i alderen 20-79 år deltok i studien. Terskler for kuldedeteksjon (CDT), varmedeteksjon (WDT), varmesmerte (HPT) og kuldesmerte (CPT) ble målt bilateralt på thenar eminens; anteriørt, midt på låret; distalt og medialt på leggen; og på fotryggen. Normalgrensene ble definert som gjennomsnitt ± 2 SD.

Resultater: CPT ble ekskludert fra analysene grunnet stor gulveffekt. Normalgrenser for sideforskjeller varierte mellom 2.0–7.4°C for CDT, 2.9–6.8°C for WDT og 3.2–4.6°C HPT. For distal-proksimale sammenlikninger varierte normalgrensene mellom 4.9–8.7°C for CDT, 6.0–14.0°C for WDT and 4.2–9.0°C for HPT, avhengig av hvilke områder som ble sammenliknet; Normalgrensene økte i tråd med distalitet. Alder var assossiert med sideforskjeller for CDT i thenar eminens ($p < 0.001$) og distale, mediale legger ($p < 0.002$), samt. med 11 av 18 av distal-proksimale sammenlikninger ($p < 0.01$). Den distal-proksimale gradienten var ikke lineær, med leggene mindre sensitive enn fotryggen for WDT og HPT ($p < 0.001$), og like sensitive for CDT ($p = 0.170$).

Konklusjon: Normalgrensene for kontralateralt homologe– og distal-proksimale sammenlikninger var brede, og forsiktighet bør utvises ved klinisk bruk. Alder, men ikke kjønn eller høyde, var assosiert med kontralaterale sammenlikninger av CDT i thenar eminens og distale, mediale legg, og med brorparten av distal-proksimale sammenlikninger, og bør korrigeres for i klinikk og eventuelt framtidig referansemateriale. Den distal-proksimale gradienten av økende sensitivitet var ikke lineær, noe som understreker behovet for tilstrekkelig spesifikke referanseverdier for termoterskler generelt.

Table of contents

0.0 Tables and Figures

APPENDIX

0.1 Common abbreviations

CDT – Cold detection thresholds

CPT - Cold pain threshold

DNFS - The German Research Network on Neuropathic Pain

HPT – Heat pain threshold

IENFD – Intra-epidermal nerve fiber density

QST – Quantitative sensory testing

QTT – Quantitative thermal testing

WDT – Warm detection threshold

1.0 Introduction

Neuropathic pain is responsible for a considerable part of the burden of pain. The prevalence of neuropathic pain is reported to be 7-10% (1), with findings ranging from 0.9% (2) to 17.9% (3), likely due to differences in definitions of neuropathic pain, assessment methods and patient selection (4). The incidence of neuropathic pain is growing, presumably in large part due to an ageing population, increased incidence of diabetes mellitus and improved cancer survival rates (5). Neuropathic pain is reportedly present in 8% of people after a cerebral infarction, in 15% of sufferers of diabetes neuropathies and post-herpetic neuralgia, in 28% of patients with multiple sclerosis, in 40–66% after spinal cord injuries, in 50–90% of amputees as phantom limb pain, in 37–55% of chronic back pain sufferers and in 28–40% of the knee arthritis and musculoskeletal-related pain population (6, 7). Furthermore, between one fifth and one half of the population reports chronic pain (8-11), and 15-25% of this is thought to be neuropathic (12).

Neuropathic pain is commonly defined as pain that is caused by a lesion or disease of the somatosensory nervous system (13). Consequently, neuropathic pain may stem from damage to myelinated (large) or unmyelinated (small) nerves anywhere from the peripheral nerve endings to the cortical neurons, e.g. a lesion of the peripheral nerve, nerve entrapment due to spinal disc herniation, or pain following a cerebral infarction (7). However, some relatively common causes of neuropathic pain principally targets the periphery, and may also largely or exclusively involve small-fiber nerves, either in general or at certain disease stages, such as diabetes mellitus, alcoholic neuropathy, chemotherapy-induced peripheral neuropathy or age-related peripheral perfusion impairment (14-17). In addition, small-fiber involvement may help to explain the wide range of symptoms seen in complex regional pain syndrome (CRPS), that is, the vasomotor, sudomotor and/or trophic changes that presents in addition to hypersensitivity and pain (18-21).

There is no gold standard for diagnosing small-fiber neuropathies. Both semi-objective assessment methods, i.e. clinical examination, quantitative sensory testing (QST) and reviewing the patient's history, and more objective ones, such as laser-evoked potentials and skin- or nerve biopsies, are commonly used in the diagnostic process (22, 23). As part of the QST test-battery, quantitative thermal testing (QTT) is suited to assess the small-fiber nerves and their corresponding central pathways (24). These thermal testing modalities are non-invasive and usually cause minimal pain, but require special equipment, a trained clinician, a motivated and cooperative patient, as well as valid reference values for the patient in question (24, 25).

The purported role of covariates on QTT values in the literature is somewhat conflicting. However, it would seem that adjusting for age (14, 26-29), sex (29-33) and possibly height (34-37) is necessary when creating reference material. As a consequence, a long list of reference materials now exist, for instance for children and adolescents (31, 38), Hispanic Latino and African American populations (39, 40), wide age-spans (28, 34, 41, 42), and for a large number of anatomical sites (36, 41, 43). In addition, it may be

important to match for equipment and protocol, such as thermode size, the rate of temperature change (ramp-rate), or the measuring method that is applied (method of limits *vs.* method of levels) (29, 30, 41, 43-46). As a result, existing reference material is largely heterogeneous and precludes the pooling of data. The German Research Network on Neuropathic Pain (DFNS) has attempted to overcome this by developing a standardized QST protocol and an associated training program, and endeavors to build a database of reference values on its foundations (27, 28, 42, 47). Even though this eliminates many technical and procedural variables, much work still remains.

A relative approach, where the patient is compared to his- or herself, either across anatomical- or contralateral homologous sites, could be complementary to absolute reference values. Within-subject comparisons could theoretically allow for assessing thermal hyperesthesia and hypoalgesia, and if they are subject to a lower inter-individual variability or show insignificant associations to common covariates for thermal thresholds, they may contribute to higher diagnostic sensitivity, and increase the external validity of relative reference material. Previous investigations have failed to find significant side-differences for QTT (28, 36, 43, 48) and it has been well established that there are differences between anatomical sites, with a likely distal-proximal gradient of increasing sensitivity (29, 30, 43, 49). Yet, attempts to quantify normal limits for such comparisons are scarce (48), and the role of common covariates such as age, sex and height, in relative comparisons is largely unknown.

Thus, the aim of this thesis is to investigate the normal limits for distal-proximal and contralateral homologous comparisons with QTT, and whether such limits should be adjusted for age, sex or height.

1.1 Aims

The aims of the thesis can be divided as follows:

I

To describe side-differences for thermal thresholds, and their limits of normality, in the extremities of healthy adults.

II

a) To investigate whether a distal-proximal gradient for thermal thresholds is present in the extremities of healthy adults; and

b) to describe such a distal-proximal gradient and the limits of normality for distal-proximal comparisons of thermal thresholds in the extremities of healthy adults.

III

To investigate whether age, sex or height is associated with relative comparisons of thermal thresholds in the extremities of healthy adults, and describe such an association.

2.0 Theoretical background

2.1 Definitions and terminology of pain

Pain is defined by the International Association for the Study of Pain (IASP) as an "unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage" (13). As such, pain can be understood in subjective terms, at the intersection between sensory input and emotion. Mosley and Butler (50) describes pain as *output*, emphasizing that the pain one experiences is the end-product of not only nociceptive signaling, but also factors such as kinesthesia, body image, mood, expectations, socio-economic status, thoughts and memories. As a consequence of these definitions, pain may be experienced even in the absence of tissue damage or nociception, and nociception can also be present without resulting in pain.

Chronic pain can be defined as pain that persists past the healing phase following an injury (51), for example radiculopathies following an intervertebral disc herniation. The healing phase can be difficult to determine, and arbitrary time-limits are often ascribed, such as 6 months for chronic back pain, or 3 months for post-herpetic neuralgia (51).

The IASP defines hyperalgesia as "increased pain from a stimulus that normally provokes pain", hypoesthesia as "decreased sensitivity to stimulation, excluding the special senses" and pain threshold as "the minimum intensity of a stimulus that is perceived as painful" (13).

2.1.1 Neuropathic pain

According to IASP (13), nociceptive pain arises from nociceptors, with or without actual or threatened tissue damage, and nociplastic pain relates to altered nociception with no clear evidence of a disease or lesion of the somatosensory system; Oppositely, neuropathic pain is defined as "pain caused by a lesion or disease of the somatosensory nervous system". The pain can be persistent with spontaneous variations in intensity, be spontaneously paroxysmal, or dependent on stimuli, e.g. allodynia (7). Two distinguishing factors between neuropathic and nociceptive pain is that the former doesn't require transduction of an external input, and that the prognosis seems to be worse (12); Otherwise, the ascending pathways, the descending modulatory pathways and the regions of the brain that are involved in pain-processing, are in large part the same for neuropathic and nociceptive pain (12, 50, 52).

As a consequence of neuropathic pain often being related to ageing and chronic diseases, the pain itself is often described as chronic. Chronic pain can follow different patterns, i.e. be relatively continuous, or have a phasic character, with fluctuating- and spontaneous pain. Notably, chronic neuropathic pain is always maladaptive, i.e. it confers no evolutionary benefit and serves no function (12).

The IASP notes in its definition that neuropathic pain is a clinical description which requires a demonstrable lesion or disease that satisfies established neurological diagnostic criteria. Importantly, neuropathic pain, as other pain conditions, cannot be objectively diagnosed in the same manner as a lesion or disease. However, the IASP's special interest group for neuropathic pain, NeuPSIG, published guidelines

for the assessment of neuropathic pain (4), updated in 2016 by Finnerup et al., that describe a grading system for determining the level of certainty for claiming that the patient's pain is neuropathic (Figure 1).

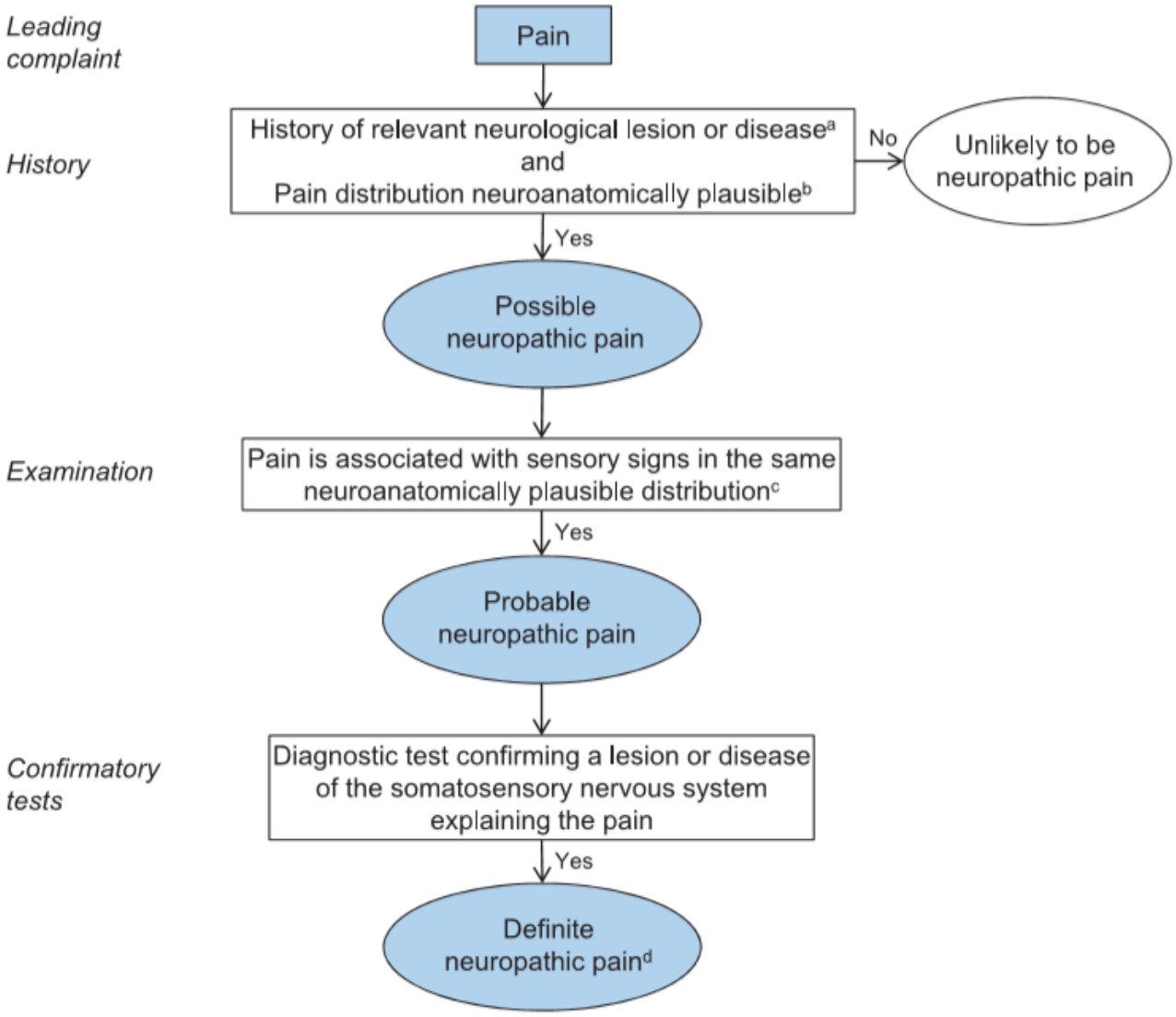

Figure 1: Flow chart of a grading system for neuropathic pain, reproduced from Finnerup et al. (55).

Following this grading system to a diagnosis of neuropathic pain, the patient must first have a relevant history that may plausibly explain the pain and its locality (e.g. nerve entrapment), this pain must then be associated with positive or negative neurological symptoms in the same area (e.g. tingling or loss of sensation), and lastly, a diagnostic test must confirm the lesion or disease (e.g. imaging). Of special noteworthiness, is the fact that NeuPSIG's guidelines do not advocate the use of QST for confirming a lesion or disease of the somatosensory system, but rather views the QST test-battery as more of an extension of the clinical examination. However, others, such as Walk et al. (53), Cruccu and Sommer (23) and Terkelsen et al. (54) proposes that it is appropriate to use QST more directly in the assessment of neuropathic pain. Importantly, QTT is widely used as a confirmatory test for diagnosing small-fiber neuropathies in clinical practice.

Many of the same mechanisms are implicated in the development- and chronification of neuropathic pain, representing a continuum of maladaptive changes. It should be noted that these mechanisms are rarely specific to neuropathic pain (52). A full review of the mechanisms involved is beyond the scope of this thesis, but a brief summary of an important overarching theme follows.

Although the underlying causes are not fully understood, neuroplasticity in particular seems to play an important role. In the periphery, nerve damage may alter the electrical properties of the sensory nerves, leading to hyperexcitability, spontaneous discharges and ephaptic transmissions, i.e. the electrical fields of a neuron altering its neighbor's excitability due to their proximity (12). In the dorsal horn, this increased input may cause activity-dependent sensitization, lowering the firing threshold of the secondary afferents (52). Additionally, peripheral nerve damage that leads to a loss of presynaptic input, causes degeneration of C-fiber terminals in the dorsal horn, allowing large myelinated fibers to sprout axons in their place, creating connections with non-nociceptive neurons and effectively creating new, nociceptive pathways (52).

Contributory mechanisms continue further up the nociceptive pathway. Not only can inhibitory interneurons be impaired, but there is an increase in descending excitatory- and a reduction in descending inhibitory signaling, further amplifying the signal that reaches the brain (5, 12, 52). In the brain, changes can occur in regions that are thought to influence descending transmission and pain-processing, for instance in the thalamus, insula, amygdala, anterior cingulate cortex, prefrontal cortex, periaqueductal gray, rostral ventromedial medulla and basal ganglia (12, 51, 52, 55, 56).

To add to this complexity, the final pain response is always multi-dimensional, consisting of both neurophysiological input and the contextual, psychological and sociocultural factors that follows being human (12, 50).

2.2 The bases of thermal sensitivity in humans

Thermal information, i.e. innocuous and noxious thermal input from the external environment, is conveyed to the central nervous system by cutaneous exteroreceptors, called thermoreceptors. These thermoreceptors are free nerve endings of thinly myelinated- (Aδ) or unmyelinated (C) fibers. The free nerve endings of C-fibers lie approximately 0.6mm within the dermis, while those of Aδ-fibers are found in 3-10 times the amount at a depth of about 0.15mm (14). In general, Aδ-fibers are responsible for cold detection, C-fibers for warm detection, while both contribute in varying degrees to heat- and cold pain (57, 58). Additionally, due to the difference in conduction velocity offered by a thin myelin sheath, the Aδ-fibers are responsible for the so called "first pain", which is usually immediate, sharp and well-localized, while the C-fibers carries the more diffuse and aching "second pain" (57). The thermoreceptors respond to temperature changes with a phasic and tonic component, allowing for a quick and precise feedback of change in temperature, followed by adaptation if exposed for a sufficient amount of time (57).

With the exception of the face, both the Aδ- and C-fibers ascend via the dorsal root ganglion to terminate

at second order neurons in the laminae I, II and V of the dorsal horn. From there, the second order axons predominately travel across to the contralateral side of the spinal cord, and ascend in the anterolateral spinothalamic tract (14, 57). Upon reaching the thalamus, third order neurons project to the cerebral cortex. While ascending, the topological arrangement of the nerves is preserved, meaning that a representation of the body's surface, including receptor density, is "mapped onto the surface of the brain", illustrated by the sensory *homunculus* of Penfield and Boldrey (59).

2.3 Small-fiber neuropathy

2.3.1 Definition and clinical manifestations

The IASP defines neuropathy as "a disturbance of function or pathological change in a nerve; in one nerve, mononeuropathy; in several nerves, mononeuropathy multiplex; if diffuse and bilateral, polyneuropathy" (13). Although most neuropathies involve both large and small nerve fibers, the term small-fiber neuropathy relates to neuropathies that predominately or selectively affects the Aδ- and C-fibers (54).

The range of possible symptoms is wide, as nerve damage may cause both positive and negative symptoms, and because of the small-fiber nerves' role in the autonomic nervous system. For instance, small-fiber neuropathy may lead to neuropathic pain, but may also cause e.g. hypo- or hyperesthesia, allodynia, numbness, burning sensations, abnormal sweating, gastric issues, skin discoloration, restless leg syndrome, erectile dysfunction, heart palpitations and orthostatic hypotension (60-62).

2.3.2 Causes and pathogenesis

Small-fiber neuropathies are the result of damage to the small peripheral nerves, and typically presents as a symmetrical, length-dependent polyneuropathy that can give either sensory or autonomic symptoms, or a combination of both (54, 60). The pathology is poorly understood, and as such, most patients are categorized as having idiopathic small-fiber neuropathy (61). However, a variety of diseases are linked to the development, with the most common identifiable cause of small-fiber neuropathies being diabetes mellitus (15). A likely mechanism for diabetes related neuropathy is that hyperglycemia-induced vascular inflammation and an overproduction of reactive oxygen species leads to atherosclerosis and endothelial dysfunction, which in turn impairs microvascularity and may cause hypoxic nerve damage (63). This may then result in impaired vascular autoregulation, leading to further loss of blood supply to the nerves and causing a negative, degenerative spiral (64).

A long list of possible mechanisms has been proposed for other causes of small-fiber neuropathies. In example, alcoholic neuropathy can be caused by malnutrition, oxidative stress or the direct effect of acetaldehyde, a toxic byproduct of ethanol metabolism (16), while the development of small-fiber neuropathy secondary to Sjögren's syndrome or systemic lupus erythematosus may be due to vasculitis or an auto-immune response (61, 65, 66). Small-fiber neuropathy may also be attributed to gene mutations in hereditary conditions (e.g. hereditary sensory autonomic neuropathy I and II) (67) or follow antiviral treatment of human immunodeficiency virus, causing inhibition of gamma DNA polymerase, reduced mitochondrial DNA content, and thus mitochondrial dysfunction (68).

An increasingly common cause, that underlines the variety of pathological processes and multi-factorial pathogenesis behind the development of small-fiber neuropathies, is chemotherapy-induced peripheral neuropathy; Proposed mechanisms include oxidative stress, apoptosis, altered calcium homeostasis, axonal degeneration, dysfunctional membrane remodeling, neuro-inflammation, changes in neuronal excitability, unfavorable immune responses and pharmacogenomic risk factors (17).

2.3.3 Diagnostics

There is no gold standard for diagnosing small-fiber neuropathies in clinical practice. Many tests have been developed, but a combination of patient history, clinical examination, skin biopsies and QTT is commonly used, in addition to (normal) nerve-conduction tests (22).

Clinical examination include inspection (e.g. discoloration, dry- or sweatiness, dystrophy), bedside neurological examination , i.e. testing the patient's response to heat, cold and pinprick-pain, and the use of standardized screening instruments (54). In addition, qualitative information regarding abnormal sensations to thermal stimuli (e.g. paradoxical heat sensations) or aftersensations may be recorded.

QTT takes the clinical examination one step further, in trying to quantify the altered sensory function of the small nerve fibers. By comparing the patient's thermal thresholds to a set of reference values, the function of the Aδ- and C-fibers (and their central pathways) can be assessed. QTT is non-invasive, but is time consuming and requires specialized equipment and the patient's focus and attention.

Though more invasive than the others, biopsies are reported to have relatively high diagnostic accuracy, and may provide additional information regarding inflammation, intra-epidermal nerve fiber density (IENFD), sweat gland innervation and axonal swelling, and possibly help establish a temporal degeneration pattern (22, 54). However, even as biopsies can reliably demonstrate structural changes in IENFD, and claim abnormality by comparison with reference material, the correlation with pain and function remains unclear. The fibers present can be sensitized or have altered function, and even complete denervation of the epidermis cannot be causally linked to pain (22, 54).

2.6 Quantitative thermal testing

QST can be defined as "a psychophysical method used to quantify somatosensory function" (24). It is comprised of a battery of tests that attempt to quantify a person's own perceptions, and thus give the clinician important information about the personal sensory experience (cf. IASP's definition of pain). The QST protocol developed by DNFS includes testing of the cold- and warm detection thresholds, number of paradoxical heat sensations during the thermal sensory limen procedure, cold- and heat pain thresholds, mechanical detection thresholds, mechanical pain threshold and mechanical pain sensitivity, dynamic mechanical allodynia, temporal pain summation, vibration detection threshold and pressure pain threshold (47).

The basic premise is to apply physical stimuli to an area of the body, that activates relatively specific receptors or anatomical parts of the sensory nervous system (44). As such, the goal of QST is to indicate whether there is a sensory alteration, e.g. hypoesthesia or hyperalgesia, as well as gather information about the patient's sensory experience, for instance paradoxical heat sensations in response to cold stimuli, or heat pain preceding warmth detection. QTT comprises the thermal testing modalities of QST, which henceforth refers to cold detection threshold (CDT), warm detection threshold (WDT), heat pain threshold (HPT) and cold pain threshold (CPT).

The QTT stimuli is applied with a thermode based on Peltier elements, allowing for the surface temperature of the thermode to be accurately changed. As the clinician is able to control the minimum and maximum thermal values, the ramp-rate, return-rate, inter-stimulus-intervals and program the testing order, duration and number of trials, the testing protocol may be standardized to a high degree.

Both the method of limits and the method of levels can be applied. In the method of limits, the temperature is gradually changed until the patient presses a trigger to signal that the stimuli has reached the desired threshold, while the method of levels applies predetermined high and low stimuli in succession to narrow in on the true threshold for the test in question (69). In general, the method of limits is less time-consuming and is somewhat reliant on reaction-time, while the method of levels is more comprehensive, but gives more accurate results, namely lower thermal thresholds (29, 69).

Several important limitations are noted by Krumova and Geber (25). First of all, the method requires active participation and is time-consuming, which means that e.g. concentration, tiredness, cognition or malingering affects the test. Secondly, the spatial resolution is limited, and so a clinical examination is needed beforehand to define the area(s) to be tested. Thirdly, it cannot differentiate between central or peripheral nerve damage, as the whole neuraxis is tested, i.e. it is of no localizing value. In addition, the readout must be compared to a reference material that is valid for the patient in question (24, 29).

2.6.1 Validity

QTT affords the clinician the ability to detect deficiencies or alterations in nervous signaling, including the central pathways, by revealing both positive and negative neurological phenomena (44). It may be particularly useful when asking about the type and extent of one or more phenomena in a particular area of interest, for example after patient history and a clinical examination reveals probable neuropathic pain or other symptoms relatable to small-fiber injury (44). As such, in a test cluster, QTT may increase both sensitivity, i.e. the ability to detect (and quantify) positive and negative sensory phenomena that the bedside examination may have missed, and specificity – identifying cases with no apparent changes in nervous function. Indeed, QTT is regarded as a well-validated method for the assessment of function in small fibers and their corresponding central pathways (24, 25, 69).

The diagnostic sensitivity and specificity of QTT for small-fiber neuropathies ranges from 67-100% (70-73) and 51-97% (70, 73-75), respectively. This varying accuracy in the literature may have several explanations. The validity of any psychophysical test is threatened by the test subject's attention, fatigue, cognitive deficits, malingering and so forth (29, 44). Furthermore, it is technically possible that central plasticity may preserve sensory normalcy in spite of reduced IENFD, and as a consequence of testing the whole neuraxis, normal thermal thresholds alone cannot exclude small-fiber impairment (25). One must also consider that, in the absence of a perfect, or agreed upon gold standard, comparing an imperfect diagnostic test to another imperfect test usually leads to the underestimation of a test's accuracy (76).

Some studies have shown a correlation between QTT and morphological measures, such as IENFD and mean dendritic length (75, 77, 78), while others have not (73, 79). However, it is commonly proposed that the methods that give morphological, psychophysical, perceptual or electrophysiological information should be viewed as complementary to each other, and be used in clusters to improve the diagnostic accuracy (24, 25, 44, 74, 75, 80).

Further impacting validity, is the choice of testing protocol, i.e. the method of limits versus the method of levels. Although the two methods are in relatively good agreement, it is well-established that the latter produces lower thresholds overall, generally attributed to a systematic reaction-time bias in method of limits, which is exacerbated by ramp-rates of ≥1°C/s (29, 30, 38, 81). The diagnostic accuracy of the method of levels is somewhat higher, and higher yet when the two methods are combined (70, 82). However, a combination is impractical for clinical purposes due to being too time-consuming.

The external validity of the reference material used is of great importance to the diagnostic accuracy of QTT. Matching for covariates such as age, sex, height and cultural or genetic descent may be necessary (14, 29, 36, 39, 40). Likewise, standardization of the protocol is also required. For instance, thermal thresholds are site-specific to an uncertain degree, and differences may arise as a result of different ramp-rates, safety cut-off values, duration of testing, inter-stimulus-intervals, verbal instructions, skin temperature, thermode size or [removal of] body hair (14, 29, 30, 41, 43-46). Ultimately, it is generally recommended that the

examiner assesses the test situation as a whole, and use their judgment regarding the validity of the test (29, 30, 44, 46).

2.6.2 Reliability

Studies that have investigated the repeatability coefficients of QTT have reported results between 0.54-10.6°C (30). Such a wide range makes it difficult to assess changes over time, and this is compounded by not knowing the contributions of experimental errors or biases or normal variations in healthy- or patient populations (29).

It has been shown that HPT and CPT have larger inter-individual variability than CDT and WDT, with CPT showing the worst results, even commonly being absent in healthy individuals (30). One hypothesis for why CPT sees such large variability between subjects, is that distinct afferent pathways are likely involved in cold detection and cold pain; Even with inhibition of the Aδ-nerves, e.g. by compression, it is still possible for the patient to feel cold-induced pain, without registering the cold (83, 84).

The thermode size may also be of importance for the reliability of QTT. Most thermodes used are relatively small (2.25-12.5cm^2), which may theoretically limit spatial summation and inadvertently target temperature insensitive fields, most commonly for warmth (30). Receptor density may vary locally or between body areas, and poor innervation could impair perception accuracy and spatial summation (36, 41, 49, 85). Indeed, a larger thermode may also be prudent for overweight or older individuals, as the skin stretches and reduces receptor density (49, 86, 87). It is therefore widely recommended to either measure adjacent skin areas, or use a sufficiently large thermode for the area of interest (24, 30, 36, 88)

The reproducibility of QTT is impacted by large intra- and inter-individual variability. The reproducibility varies from poor to excellent, and is not surprisingly affected by everything from the equipment used to the populations tested and algorithms applied (29). In general, QTT is reasonably reproducible (24, 29), with better results in the hands than feet (89), in single-center comparisons (90, 91), and over shorter periods, i.e. days or weeks (29, 43).

The reproducibility and repeatability of QTT may be affected by the dynamic nature of pain (44), but is considered to be good, especially for the method of levels (24, 30, 89). Gelber and Pfeifer (88) reported good test-retest reliability within- or between days, between technicians, and from center to center, although the coefficient of variation ranged from 60-145% for CDT in fingers and toes. Similarly, Geber et al. (92) showed good test-retest (within-day and day-to-day) and inter-observer reliability in patients with and without pain, for all thermal modalities ($r \geq 0.8$). Similar studies have shown good test-retest results over trials, sessions and days, and across an array of body sites (85, 93-96). For the most part, CDT and WDT display equally good test-retest reliability, while HPT and CPT are usually in lesser agreement or show weaker correlations, presumably due to their tendency towards larger variability (30, 35, 95). As a result, ICC values for the QTT modalities have been reported to range from 0.32 to 0.97 (30). Importantly, Kemler

et al. (97) reported that the repeatability was similar for an absolute- and relative approach, although the external validity of this finding may be limited as the patient population suffered from CRPS.

2.6.3 Relation to age, sex, height and skin temperature

2.6.3.1 Age

The association between age and thermal thresholds has been the target of many studies. In adults, age seems to increase both thermal thresholds and their variability – particularly detection thresholds – although the findings are not clear-cut (14, 27, 28, 34, 39, 42, 43, 98-101). Previous research paints a picture of age affecting thermal thresholds and variability more in the distal regions (14, 28, 36, 49, 102, 103), a finding that may have implications for relative comparisons across body sites.

Several explanations for the association between age and thermal thresholds have been put forth. As many studies apply the method of limits, a part of the association may cautiously be attributed to an age-related increase in reaction-time, however it seems that this is insufficient in explaining the changes observed (100). Normal age-related morphological changes are likely culprits, as age-related deterioration in the periphery include e.g. distal axonopathy, reduced IENFD, and increased firing thresholds in peripheral nerves (14, 104). Additionally, age is associated with some neuronal loss, somewhat preferentially targeting myelinated nerves such as the Aδ-fibers (104), and age-related reductions in peripheral microvascularity may cause hypoxic nerve damage that impairs function (14).

2.6.3.2 Sex

Sex and thermal thresholds may be associated, but the relationship is not well-defined, and may vary according to body site and thermal modality. While some studies have a found that female sex lowers thermal thresholds in both children and adults (27, 31, 33, 37, 43, 81, 98, 99, 101), others have found a negligible or absent effect of sex (26, 27, 35, 40-42, 49, 88, 105).

Although the link between sex and thermal thresholds has not yet been fully elucidated, a few explanations for why females could be more sensitive have been proposed. A difference between the sexes may exist in the central processing of thermal stimuli, as the peripheral nerve fibers functions similarly for thermal detection and discriminatory tasks (33). For instance, greater temporal summation is seen in females in response to sustained thermal stimuli (32). It is also possible that gonadal hormones contribute to lowering the pain thresholds in females (106); This effect seems to interact with age, with sex differences declining towards menopause, and eventually becoming negligible (87). The effect of sex could also be confounded by differences in height, as some studies have reported no effect when controlling for height (34, 37).

2.6.3.3 Height

A few studies have investigated the relationship between height and thermal thresholds. An association has been reported for CDT in the feet (35) and toes (86), for CDT and WDT in the hands (34), and for HPT in the distal lateral leg, thenar eminence and radial part of the lower arm (36), with height increasing the thermal thresholds. However, these studies simultaneously failed to find an association for a long list of thermal modalities and body parts. Other studies are more categorical, as e.g. Meh and Denišlič (43) found no effect of height on thermal thresholds.

The association with height is somewhat tenuous and unsystematic, and Bartlett and Stewart (34) argues that it is of little clinical significance. Wasner and Brock (95) follows this by suggesting that the length of the extremities themselves should be of little impact: For instance, a 15 cm difference and a conduction velocity of 0.5-2 m/s in C-fibers, with a ramp-rate of 1°C/s, would not even amount to half a degree difference when utilizing the method of limits, and would not affect the result if the method of levels was used. However, IENFD is related to body size (49), and it has been repeatedly reported that height impacts nerve conduction, possibly through a degree of distal axonal tapering (107-110). Consequently, more research is needed before height can be excluded as a covariate for thermal thresholds.

2.6.3.4 Skin temperature

Several studies have investigated the effect of skin temperature on thermal thresholds, and the findings are inconclusive. While some have reported an association between skin temperature, CDT and CPT (37, 38, 41, 86, 111), the effect sizes are small and generally limited to the hands and feet. Others have failed to find such an association, largely independent of site or thermal modality (34, 41, 43, 88). As a result, review articles differ in their recommendations, with Guergova and Dufour (14) and Siao and Cros (29) recommending to control for skin temperature, while Bakkers and Faber (30) proposes that the effect is insignificant.

3.0 Materials and methods

3.1 Study design

An experimental, cross-sectional study was designed to compare thermal thresholds with regards to distal-proximal gradients and side-differences. Four sites were measured bilaterally: the thenar eminence, the anterior thigh (10 cm superior to the patellar base in mid-line), the distal medial leg (directly superior and posterior to the medial malleolus) and the foot dorsum (dorsal aspect of metatarsals II-III) (Figure 2).

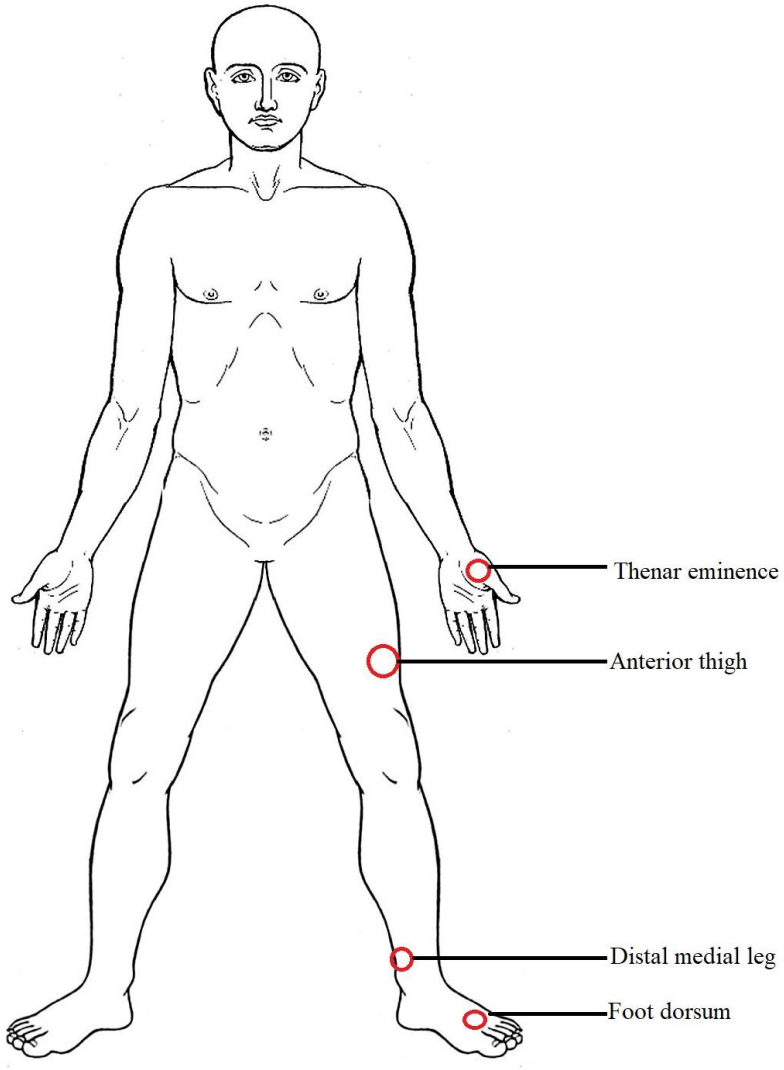

Figure 2: Bilateral sites of measurements for thermal thresholds (red circles)

The testing order of sites was randomized in advance (.NET pseudo-randomization). Skin temperature was measured at each body site before QTT measurements initiated, and a re-usable heat pack was applied when necessary, to achieve skin temperatures of ≥32°C for the thenar eminence and ≥30°C for sites in the lower extremities.

A pre-test was conducted on the subject's non-dominant volar forearm, consisting of two CDT- and two WDT measurements, to familiarize the subject with the procedure and trigger mechanism. Subsequently, CDT, WDT and HPT was measured in succession at each body site, followed by an independent

measurement of CPT in the distal medial legs and feet dorsa, in the same randomized order (Figure 3).

The subjects were informed of the testing procedure by means of a standardized instruction sheet. A single, male experimenter carried out all experiments. The placement of instruments, room temperature (20-23°C), experimenter's clothing and lighting was standardized. Participants were blinded to the study's hypotheses and instrument readouts.

Figure 3: Timeline of each experiment. Skin temperature was measured immediately after subjects undressing. A pre-test was performed to familiarize subjects with the protocol and stimulations. Cold detection threshold (CDT), warm detection threshold (WDT) and heat pain threshold (HPT) was measured in succession for each of the eight test-sites. Lastly, cold pain threshold (CPT) was measured in the distal medial legs and feet dorsa, in the same order.

3.2 Sample

3.2.1 Sample Size

Sample size for comparisons between body sites was calculated in accordance with the equations provided by Rosner (112). The α value (two-tailed) was set to 0.05, while the β value was set to 0.2 (i.e 80% power). Values for effect sizes and standard deviations were based on data provided by Rolke and Baron (28), Kemler and Schouten (48) and Hafner and Lee (26), as well as clinical experience.

For side-differences, a minimum difference of 1°C with standard deviations of 2°C was used, resulting in a minimum of 31 subjects needed. For distal-proximal comparisons of body sites, these values were 2°C and 3°C, respectively, producing a requirement of 18 subjects.

3.2.2 Recruitment and exclusion criteria

Healthy men and women, ages 20-79 were recruited through advertisements at Oslo University Hospital, local universities, gyms, centers for the elderly, and on social media.

Exclusion criteria were as follows: cancer (current or previously), diabetes, radiculopathy, chronic pain (average NRS ≥ 1 for ≥ 3 months, last two years), pregnant or breastfeeding, limited capacity for consent, personal acquaintance of experimenter, or any disease of nerves, muscles, or of the brain that could influence normal nervous function, including psychiatric illnesses.

Subjects were requested not to work nightshifts within 48 hours of the experiment, to not consume alcohol in the last 12 hours before the experiment, or consume pain-killers the same day as the experiment.

Recruitment efforts focused on obtaining equal numbers male and female subjects, with all age groups similarly represented. In addition, the data would ideally be stratified by decades and later be used in the creation of an absolute reference material, and as such, the recruitment process aimed to recruit as many subjects as possible.

3.3 Experimental protocol

Thermal stimulus was applied with a 30 x 30 mm Peltier thermode (Medoc, Ramat Yishai, Israel). The thermode was manually held in place by the experimenter. Baseline temperature was set to 32°C, with safety cut-off values of 0°C and 52°C.

Scripted, verbal instructions were used. Participants were informed of the procedure in its entirety before testing began, and reminded of the current modality before each test. They were instructed to converse as little as possible, but were allowed to ask questions.

Subjects lay supine on a treatment table, with the back rest at approximately 120-135° incline. Pillows were used for head-support and placed under the subjects' knees, and a duvet helped regulate skin temperature. A wired computer mouse served as the trigger, and was held by the subjects in their dominant hand, except when testing the dominant thenar eminence. For each site, excessive body-hair was gently removed with scissors, so as not to irritate the skin.

The skin temperature was measured at each site with an 826-T2 hand-held infrared thermometer (Testo SE & Co., Pennsylvania, USA), held perpendicular to the skin's surface at a standardized distance of 1 cm. A re-usable heat pack was applied where skin temperature was <30°C for the lower extremities, and <32°C for the thenar eminence.

The method of limits was employed: with the thermode held in place, the temperature increased gradually from baseline until the subject pressed the trigger, at which point the temperature was recorded and the thermode's surface temperature returned to baseline. For CDT and WDT, subjects were asked to press the trigger at the first sensation of cool or warmth. Similarly, for HPT and CPT, the cue was to press the trigger at the first sensation of pain, typically when the thermal stimulus begins to induce a stinging, burning or aching sensation. A response was considered invalid and repeated once if it deviated substantially from

contemporaneous measurements, or if the subject admitted to an accidental response. Each thermal detection threshold was measured five times in quick succession, while the thermal pain thresholds were measured three times.

Based on previous experience, subjects were advised that pain thresholds are not a measure of pain tolerance, i.e. the purpose of the investigation was to establish the thresholds at which a stimulus changes from warm or cool to painfully hot or cold, and not what thermal intensity the subject can endure.

The ramp-rate was set to 1°C/s, and the thermode returned to baseline after each measurement at the rate of 1°C/s and 5°C/s for detection- and pain thresholds, respectively. Inter-stimulus-intervals were randomly assigned to 4-6 seconds for all consecutive tests, while changes between thermal modalities were initiated manually by the experimenter.

CDT, WDT and HPT were measured in succession for each site, followed by CPT in the distal medial legs and feet dorsa in the same order. Absolute temperature thresholds were recorded. The testing session lasted between 60 and 90 minutes in total.

3.4 Data computation and analysis

Statistical analyses were performed using IBM SPSS Statistics v. 25 (Armonk, NY: IBM Corp.). P-values were regarded as significant at ≤ 0.05 and correlation values were interpreted in accordance with Mukaka (113): negligible correlation ±0.0–0.3, low correlation ±0.3–0.5, moderate correlation ±0.5–0.7, high correlation ±0.7–0.9 and very high correlation ±0.9–1.0. The absolute temperature thresholds recorded were calculated to express absolute change from baseline (Δ°C).

Data distribution was assessed in preliminary analyses by use of descriptive tables, histograms, boxplots and Q-Q plots, and uncertainty regarding a distribution was solved through discussion with supervisors. The arithmetic mean of five (CDT, WDT) or three (HPT, CPT) measurements was used in the analysis.

Data from each subject was excluded from the calculation of sample thresholds and for performing linear regression if the following was present:

i) If the delta value of a thermal threshold was >3 times (detection thresholds) or <1/3 times (pain thresholds) the arithmetic mean of the remaining data points; or

ii) If the delta value of a thermal threshold exceeded ± 3 SD from the arithmetic mean of the remaining data points.

Invalid measurements due to a test's floor- or ceiling effect were not replaced or included in the final analysis.

Side-differences were determined by use of multiple paired t-tests. The distal-proximal gradients were examined by repeated measures ANOVA with post-hoc analysis for pairwise comparisons. Bonferroni's adjustment for multiple testing was applied to post-hoc comparisons and to the series of paired t-tests. Sample normal limits for side-differences and distal-proximal gradients were calculated as mean ± 2 SD.

Pearson- or Spearman correlation was calculated to determine the association between side-differences or

distal-proximal gradients, and age, sex and height. In the event of a significant association with sex or height, an interaction between the two would be investigated through multiple linear regression.

Equations for determining sample normal limits (Δ°C) for CDT and WDT as a function of age in the distal medial leg and foot dorsum, as well as for distal-proximal comparisons between the foot dorsum and anterior thigh, were calculated by use of linear regression. Appropriate variance-stabilizing transformations were applied when necessary. The normal limits for regression-derived data was defined as the upper 95% prediction interval.

3.5 Ethical considerations

This study was integrated in a large, existing project, aimed at better understanding the mechanisms behind the development of chronic pain. Important ethical considerations include the study's merit (e.g. value and validity of the research), respect of the individual participant (e.g. the recruitment procedure, sample size, duration of the experiment, protection of privacy and informed consent), as well as risk-benefit assessment with regards to exposing participants to noxious thermal stimuli. The study will be published with open access at the earliest convenience, so as to maximize the participant's contribution to the field.

The parent study was approved by Regional Committees for Medical and Health Research Ethics (REC), project no. 2010/2927.

All participants provided written consent, and the study was conducted in accordance with the Declaration of Helsinki. Subjects received a gift certificate of NOK 250 for participation.

4.0 Results

All results are presented in Appendix A, but a brief synopsis follows.

Forty-eight subjects were included in the analysis. Participants were 46 (SD ± 15.6) years old, 52% of whom female. Two subjects were excluded from the analysis of CDT in the distal medial leg, 34 from CPT in the foot dorsum, and 37 from CPT in the distal medial leg, due to reaching the test's floor value of 0°C. The large floor-effect precluded calculation of sample CPT thresholds or relative CPT comparisons.

Some outlying data was identified and removed for CDT in the foot dorsum (n = 2), distal medial leg (n = 5) and thenar eminence (n = 2); for WDT in anterior thigh (n = 2) and thenar eminence (n = 1); and for HPT in anterior thigh (n = 1) and thenar eminence (n = 1).

A significant right-left side-difference of -1.4°C (-2.1°C – -0.6°C) was found between the feet dorsa for WDT ($p = 0.001$), while other comparisons were non-significant. Sample normal limits for side-differences ranged from 2.0–7.4°C for CDT, 2.9–6.8°C for WDT and 3.2–4.6°C for HPT.

Repeated measures ANOVA with Greenhouse-Geisser correction detected differences in means for distal-proximal comparisons for CDT, WDT and HPT ($p < 0.001$). Post-hoc analysis revealed a general distal-proximal gradient of increasing thermal sensitivity, with the exception of the distal medial leg being similarly– (CDT) or less sensitive (WDT, HPT) than the foot dorsum, as well as no difference found between the foot dorsum and anterior thigh or thenar eminence for HPT. Sample normal limits for distal-proximal comparisons ranged from 4.9–8.7°C for CDT, 6.0–14.0°C for WDT and 4.2–9.0°C for HPT.

No correlation was found between side-differences or distal-proximal comparisons and height or sex. A low correlation was found between age and CDT in the thenar eminences ($p = 0.001$) and distal medial legs ($p = 0.002$). Correlations ranging from moderate to high were found for 11 of 18 distal-proximal comparisons ($p < 0.01$), largely representing comparisons involving the feet dorsa or distal medial legs.

5.0 Discussion

5.1 Results and clinical implications

Within-subject comparisons are in a sense a holy grail of clinical neurological assessments; For instance, if healthy subjects show no side-differences and little variability, one could reliably assess clinically relevant side-differences when examining a patient, organically controlling for a range of potential covariates. For the basic neurological examination, comparing with the contralateral site is indeed considered a fundamental principle and constitutes a time-honored practice (46, 114). Our results support previous findings that side-differences for QTT are generally non-existent, while at the same time, large inter-individual variability exists (28, 31, 36, 42, 43, 45, 47, 48, 88). This entails that as high inter-individual variability leads to wide normal limits, QTT's diagnostic sensitivity decreases. Comparing sides would prove difficult when a wide range of values can be considered normal, and as such, the ability of QTT to assess thermal hypoesthesia or hyperalgesia diminishes. However, as the diagnostic sensitivity of QTT should decrease with wide normal limits, the specificity would increase, potentially allowing for a confirmatory role in a cluster of tests assessing small-fiber function. One must also consider that contralateral comparisons may be the only viable option when assessing hyperesthesia and hypoalgesia, as absolute reference material cannot be utilized for this purpose (48). Thus, even though contralateral comparisons with QTT are unlikely to be as useful as one could hope, they may still have a complementary role to play.

Interestingly, one study shows contralateral comparisons to be more sensitive than absolute reference material: Rolke and Baron (28) found that right-left comparisons were 49% and 65% more sensitive for heat- and cold hypoalgesia, respectively, in a pooled dataset of the face, hand and feet. In a similar vein, Maier et al. (115) used contralateral comparisons as a supplementary method to increase total sensitivity when subjects' values fell within normal limits of absolute reference values, leading to the identification of an additional 4.6–8.4% pathological cases, depending on the thermal modality in question. It is important to note that neither of these studies stratified their results by age. The variability of thermal thresholds seems to increase with age (14, 34, 39, 102, 103), particularly in the distal regions, and it is likely that this would also be true for relative comparisons. Indeed, the fact that relative comparisons may increase diagnostic sensitivity in some age-pooled data, speaks to its potential and ought to warrant further investigation. Although the present study did not examine diagnostic sensitivity directly, the normal limits reported are likely to be widened by age-pooling and the small sample size. It is thus possible that future studies may still find that contralateral comparisons of some thermal modalities are of clinical value in younger adults, as the test's sensitivity would theoretically increase with a narrowing of the inter-individual variability.

In line with Yarnitsky and Sprecher (81), we found that the inter-individual variability is dependent on body-site, underlining the importance of sufficiently anatomically specific reference values. In fact, the normal limits for contralateral comparisons can vary by a factor of more than three, exemplified by our sample normal limits for CDT of 2.0°C for the thenar eminences and 7.4°C for the distal medial calves.

Following this, regarding a set difference between contralateral homologous sites of e.g. ≥±1°C or ≥±2°C as pathological, like Leffler and Hansson (114), may not only be erroneous because of the narrow limit proposed, but also because different limits should systematically be applied depending on the body site examined.

We show a similarly large variation in inter-individual variability for inter-region comparisons, and an association with age. As a result, it seems clear that a set of age-stratified normal limits are also required for each distal-proximal comparison of clinical interest. Fortunately, our data implies that sex and height need not be accounted for in such a future relative reference material, a finding that would of consequence to the external validity. However, neither previous research (28, 36, 42, 43, 45, 47, 48, 87, 88), nor our present work, permits drawing any conclusions regarding other plausible confounders, such as the use of the method of limits versus the method of levels, environmental factors or technical variables such as thermode size or ramp-rate. Clearly, not much is known about possible covariates and confounders for relative comparisons, but our findings do contribute meagerly to suggest that sufficiently large samples may be more easily achieved in future studies by omitting to adjust for sex and height.

A secondary finding of note was the large floor-effect of CPT in the feet dorsa and distal medial legs. With 70-77% of subjects reaching the test-floor of 0°C, planned calculations became underpowered, and possibly even clinically meaningless. As most subjects provided invalid measurements, and the remainder ranged through the entire thermal spectrum tested, we confirm previous findings that CPT may not be as clinically useful as the other thermal modalities (24, 30). Indeed, Neziri and Scaramozzino (87) excluded CPT results due to most participants (n = 300, age range 20-77 years) reaching 0°C, while Rolke and Baron (28) showed a normal range of 0–31°C in healthy women under the age of 40. Thus, just as CPT is of limited value when utilizing absolute reference material, attempting relative comparisons of CPT may prove to be futile.

5.2 Relevance for physical therapy

Physical therapists are found in all parts of the healthcare sector, and many countries, including Norway, Australia, Brazil and parts of the United States, patients have direct access to physical therapists, without the need of a referral (116). As such, physical therapists have a role in diagnosing and treating patients in a one-on-one setting in primary care, and may also be part of interdisciplinary teams, e.g. performing clinical examinations in secondary care, including thermal testing. This means that the physical therapist is directly involved and invested in uncovering the mechanisms that underlie the patient's problems, in order to provide best-practice individualized care.

Basing the treatment of pain on the underlying mechanisms is widely accepted to be superior to disease- or cause-based treatment, although it can be difficult to achieve (12). For physical therapists, a mechanism-based approach makes sense, as many of the common treatments, such as massage, transcutaneous electrical nerve stimulation, soft-tissue- or joint manipulation, or the choice of exercise and total exercise load may work differently depending on the pain mechanisms involved (117). For instance, manual nerve flossing

exercises have been shown to ameliorate nerve compression in cadaver studies (118, 119), and both aerobic and resistance training of moderate intensity may promote nerve healing and analgesia, and lead to an increase in IENFD (117, 120). At the same time, muscle relaxation exercises, dry-needling or joint manipulation are likely to be of limited use in the treatment of neuropathies. One of the most important actions of treatment is the education of the patient about their condition, and when the mechanism of pain is known, this education may be tailored to a higher degree (117). As such, it should naturally follow that physical therapists take interest in elucidating pain mechanisms in general, and in applying this knowledge in the clinical setting; An understanding of the pain mechanisms influencing the patient in question, could both improve the physical therapist's examination and choice of clinical tests, and help to inform the most appropriate path of treatment.

This study also highlights an important area for both the clinical physical therapist and the field of physical therapy research in general. Just as one would want to compare contralateral homologous sites or body regions in a basic neurological examination or with QTT, the physical therapist commonly compares sides when assessing e.g. muscle strength or joint laxity, or when palpating soft tissues or performing provocation tests. However, without a clear image of the normal variation for a given test, both between- and within subjects, this practice may be of limited diagnostic value (Figure 4). This source of error would be further compounded by performing the same test at a later date to determine progress following a rehabilitative effort, where two wrongs does not make a right. The wide range of normalcy shown for QTT should raise some doubt as to what other diagnostic tools may need to be revisited, to make sure that the tools themselves, and indeed the way in which they are used, is valid and reliable.

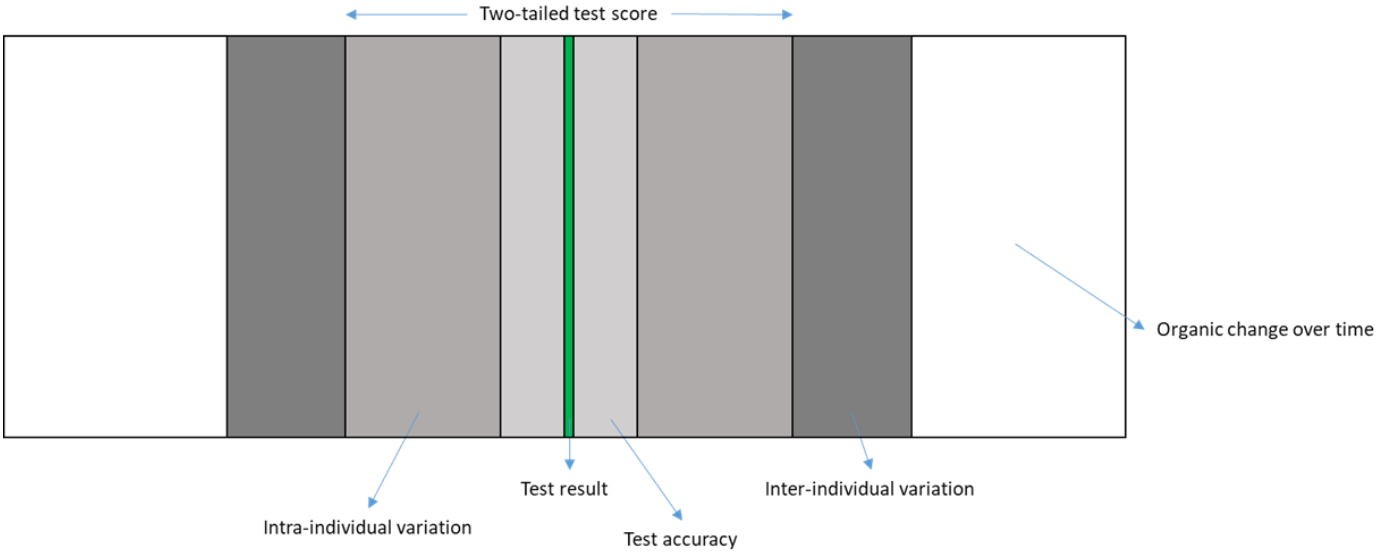

Figure 4: Sources of error for a theoretical test, with arbitrary divisions. The test results lie within the limits of the test's accuracy, and is subject to intra-individual variation, e.g. a patient having a good or bad day. When comparing with reference material or clinical experience with other individuals, the inter-individual variation expands the limits of what may still be considered normal. Further adding to the uncertainty of the test result, is that the phenomena being quantified may organically change from a test to another, e.g. regression towards the mean, or rapid phasic cycling of pain. For test-retest scenarios or relative comparisons, the end result would be a product of the relevant parts of two such figures, i.e. the sources of error are multiplied. Thus, the physical therapist must be aware that normality for a given test exists on a spectrum, which may sometimes be much wider than initially thought.

5.3 Strengths and limitations

Several decisions were made regarding the data acquisition, that may have impacted the results and their internal– and external validity. While the method of limits is not as accurate as the method of levels (24, 29, 38, 81), it was chosen to mirror the clinical protocol used at the hospital where the research was conducted, which is based on that of DNFS. Since the results were not compared to absolute reference material, somewhat inflated thermal thresholds would likely not impact our tests or conclusions. In the same manner, the skin was heated when needed to approximately 30°C and 32°C for the lower extremities and thenar eminence, respectively, in accordance with the protocol. While this precludes examining a potential effect of skin temperature on relative comparisons, it's role was controlled for experimentally. Furthermore, the sites of measurements were picked carefully. The sites represent areas of high prevalence for neuropathies involving small fibers, they are used clinically at the hospital where the study was conducted, and they form a distal-proximal pattern. While we would ideally have an additional point of measurement in the upper extremities, this was decided against as the protocol became too time-consuming. The sites were also chosen to better compare with other research groups, in an attempt to increase data compatibility across both centers and borders. Lastly, our strict adherence to the standardized experimental protocol adapted from DNFS helped ensure that important experimental variables such as ramp-and-return-rate, number of trials or instructions were controlled for during data collection. We believe that the sum of these decisions strengthens the validity and integrity of the study.

A crucial limitation was our failure to stratify by age. While originally planning to stratify by decades, the study ended up not being sufficiently powered to achieve this goal. The determination of sample-size for the creation of a reference material constitutes a trade-off between the amount theoretically needed and what is practically possible to achieve, and while aiming for more than 100 subjects, we ended up with 48 for the final analysis. Too small a sample could generate wider normal limits, which may be especially detrimental for the elderly; Not only were the older adults most poorly represented in our study, they are reported to show particularly high inter-individual variability (14, 34, 39, 102, 103). Unfortunately, they also constitute the age group most likely to present with neuropathies (121), and thus in need of accurate normal limits for QTT. As age is associated with many of the relative comparisons in the present study, our age-pooled sample presents normal limits that are likely to be of too low accuracy and precision to be used for diagnostic purposes.

The problem of sample size also affects the linear regression performed for the supplementary table of thermal thresholds. Whereas thermal thresholds generally increased proportionally with age in our data, so did the variability, leading to heteroscedasticity. Because of the high variability seen in older adults, our few data points form an incomplete picture, and produce significant doubt with regards to slope steepness. While square root transformation of the data reduced the variance and allowed the data to meet the assumptions for linear regression, uncertainty regarding the accuracy of the slope remains, and thus also the 95% prediction intervals representing the normal limits. Relatively small decreases in slope steepness would have a large

impact on the prediction of normal limits for the oldest adults, which, for example, reaches theoretical safety cut-off levels (52°C) when passing 70 years of age for WDT in the foot dorsum in our model. While the absolute normal limits predicted by this model can be seen in light of other reference material, to our knowledge, the normal limits for comparisons between the foot dorsum and anterior thigh is novel. As a result, the normal limits predicted for inter-region comparisons should be interpreted with great care, particularly for the elder half.

Although not strictly necessary to answer the main research questions, several variables of clinical value may have been wise to record during the experimental phase. For instance, we did not record "misclicks", i.e. erroneous triggering during testing. During data collection, it became evident that many subjects, seemingly mostly over the age of 50, pressed the trigger several times during periods where this was not registered by the machine. This mainly happened while testing for WDT, especially in the distal medial legs, but also for the feet dorsa; While a subject could press the trigger in response to what they felt was a sensation of warmth 8-9 times during five tests, these misclicks would not be registered while the thermode's surface temperature was returning to baseline or during the inter-stimulus-interval. In the same vein, recording the presence of paradoxical thermal sensations, or when a subject did not detect a change in temperature before sensing thermal pain, could contribute to our understanding of normal responses to QTT. This data would allow for a more accurate picture to be painted of how difficult it can be for healthy subjects to accurately detect thermal stimuli in the distal lower extremities, and even whether this is dependent on e.g. age or sex. Hence, not recording misclicks is an important shortcoming that should be considered in future research.

How to deal with outliers is another problematic area when such a broad testing is employed. With a psychophysical study design consisting of 116 recorded measurements per subject, some outliers are to be expected. Although tests with obvious or admitted misclicks were re-done immediately, they were never repeated more than once, and some intra-individual variation in thermal thresholds was expectedly present. When investigating normal thresholds in healthy subjects, it is important to stay as true to the actual results as possible, lest the data become sterile with low external validity for the clinical setting. Still, our approach did involve some conservative treatment of outliers. The impact of single-test outliers was blunted by calculating the mean of 3-5 repeated tests (depending on modality). Furthermore, prior to performing linear regression and calculating sample thermal thresholds, only relatively extreme outliers of the *delta values* were excluded by rules of common local practice for determining reference values for thermal thresholds, resulting in few removals in total, and the preservation of an arithmetic mean also in these subjects. Very few mean outliers were removed, and invalid measurements (floor- or ceiling effects) were not replaced with imputed ones, as they were not missing, but only considered extreme by our limited test-range. With this careful effort, we believe that the data represents much of the true variation seen between healthy subjects.

Finally, we made a principal decision to correct for multiple comparisons in our analysis. The statistical theory underpinning whether to correct for multiple comparisons, and if so, which correction method to use,

is an area widely discussed in statistics (122), and far beyond the scope of this thesis. However, by accepting an alpha level of 0.05, we simultaneously accept a familywise error-rate that becomes quite high when performing up to 18 pairwise comparisons (i.e. $1-(0.95^{18})$). A simple question one can ask, is whether it is important to the research question at hand that all possibly statistically significant findings are reported for future investigations, or whether a more conservative approach of limiting false positives could or should be applied. Following this logic, it might be considered important to report all significant outcomes when testing a potentially life-saving medicine in an array of Petri dishes, and exclude false findings later in the research process. In the present study, however, excluding false positives served the purpose of identifying the leads most likely worth following up on, while the possible failure to find certain associations were not considered to be too impactful. Still, as statisticians disagree (123, 124), we decided that the most prudent path was to apply the simple, well-known and conservative Bonferroni adjustment, while also reporting accurate p-values for all comparisons, so as to give the reader the opportunity to interpret the raw results.

6.0 Conclusion and future research

We showed fairly wide normal limits for all within-subject comparisons, particularly in the distal lower extremities. This may limit the diagnostic sensitivity of relative comparisons of thermal thresholds, but could simultaneously increase its specificity; As such, within-subject comparisons of thermal thresholds may be better suited as a supplement to the use of absolute reference material, or as part of a test-battery, while caution is advised for utilization as an isolated clinical test.

Side-differences for CDT in the thenar eminences and distal medial legs, as well as most inter-region comparisons, were correlated with age, but not with sex or height. Consequently, it may be necessary to adjust for age, but not sex or height in relative reference material.

Our results revealed a non-linear distal-proximal gradient of increasing thermal sensitivity from the foot dorsum to the thenar eminence. Coupled with the finding of varying inter-individual variability across anatomical sites, we propose that it is crucial that the clinician utilizes highly site-specific normal limits.

Future research should consider stratifying results by age to further elucidate the association between age and relative comparisons, and strive to increase the resolution of anatomical sites. In general, research involving QTT may benefit from recording all triggers by participants during the QTT procedure, particularly when testing WDT in the distal lower extremities.

7.0 References

1. Van Hecke O, Austin SK, Khan RA, Smith B, Torrance N. Neuropathic pain in the general population: a systematic review of epidemiological studies. PAIN®. 2014;155(4):654-62.
2. Gore M, Dukes E, Rowbotham DJ, Tai KS, Leslie D. Clinical characteristics and pain management among patients with painful peripheral neuropathic disorders in general practice settings. Eur J Pain. 2007;11(6):652-64.
3. Toth C, Lander J, Wiebe S. The prevalence and impact of chronic pain with neuropathic pain symptoms in the general population. Pain Med. 2009;10(5):918-29.
4. Haanpää M, Attal N, Backonja M, Baron R, Bennett M, Bouhassira D, et al. NeuPSIG guidelines on neuropathic pain assessment. PAIN®. 2011;152(1):14-27.
5. Colloca L, Ludman T, Bouhassira D, Baron R, Dickenson AH, Yarnitsky D, et al. Neuropathic pain. Nature reviews Disease primers. 2017;3(1):1-19.
6. Andrew R, Derry S, Taylor RS, Straube S, Phillips CJ. The costs and consequences of adequately managed chronic non-cancer pain and chronic neuropathic pain. Pain Practice. 2014;14(1):79-94.
7. Jørum E. Utredning av nevropatisk smerte. Tidsskr Nor Lægeforen 2005. 2005;125: 2652-4.
8. NIPH. Norwegian Institute of Public Health: Public Health Report - Chronic Pain. 2019.
9. Dahlhamer J, Lucas J, Zelaya C, Nahin R, Mackey S, DeBar L, et al. Prevalence of chronic pain and high-impact chronic pain among adults—United States, 2016. Morbidity and Mortality Weekly Report. 2018;67(36):1001.
10. Fayaz A, Croft P, Langford RM, Donaldson LJ, Jones GT. Prevalence of chronic pain in the UK: a systematic review and meta-analysis of population studies. BMJ Open. 2016;6(6):e010364.
11. Sá KN, Moreira L, Baptista AF, Yeng LT, Teixeira MJ, Galhardoni R, et al. Prevalence of chronic pain in developing countries: systematic review and meta-analysis. PAIN Reports. 2019;4(6):e779.
12. Cohen SP, Mao J. Neuropathic pain: mechanisms and their clinical implications. Bmj. 2014;348:f7656.
13. IASP. International Association for the Study of Pain 2018 [Available from: https://www.iasp-pain.org/terminology?navItemNumber=576.
14. Guergova S, Dufour A. Thermal sensitivity in the elderly: A review. Ageing Research Reviews. 2011;10(1):80-92.
15. Tavee J, Zhou L. Small fiber neuropathy: a burning problem. Cleve Clin J Med. 2009;76(5):297-305.
16. Chopra K, Tiwari V. Alcoholic neuropathy: possible mechanisms and future treatment possibilities. Br J Clin Pharmacol. 2012;73(3):348-62.
17. Starobova H, Vetter I. Pathophysiology of Chemotherapy-Induced Peripheral Neuropathy. Front Mol Neurosci. 2017;10:174-.
18. Harden, Bruehl S, Stanton-Hicks M, Wilson PR. Proposed new diagnostic criteria for complex regional pain syndrome. Pain Med. 2007;8(4):326-31.
19. Oaklander AL, Fields HL. Is reflex sympathetic dystrophy/complex regional pain syndrome type I a small-fiber neuropathy? Ann Neurol. 2009;65(6):629-38.
20. Oaklander AL, Rissmiller JG, Gelman LB, Zheng L, Chang Y, Gott R. Evidence of focal small-fiber axonal degeneration in complex regional pain syndrome-I (reflex sympathetic dystrophy). Pain. 2006;120(3):235-43.
21. Oaklander AL, Romans K, Horasek S, Stocks A, Hauer P, Meyer RA. Unilateral postherpetic neuralgia is associated with bilateral sensory neuron damage. Annals of neurology. 1998;44(5):789-95.
22. Devigili G, Tugnoli V, Penza P, Camozzi F, Lombardi R, Melli G, et al. The diagnostic criteria for small fibre neuropathy: from symptoms to neuropathology. Brain. 2008;131(7):1912-25.
23. Cruccu G, Sommer C, Anand P, Attal N, Baron R, Garcia-Larrea L, et al. EFNS guidelines on neuropathic pain assessment: revised 2009. Eur J Neurol. 2010;17(8):1010-8.
24. Backonja MM, Attal N, Baron R, Bouhassira D, Drangholt M, Dyck PJ, et al. Value of quantitative sensory testing in neurological and pain disorders: NeuPSIG consensus. Pain. 2013;154(9):1807-19.
25. Krumova EK, Geber C, Westermann A, Maier C. Neuropathic pain: is quantitative sensory testing helpful? Current diabetes reports. 2012;12(4):393-402.
26. Hafner J, Lee G, Joester J, Lynch M, Barnes EH, Wrigley PJ, et al. Thermal quantitative sensory testing: a study of 101 control subjects. Journal of Clinical Neuroscience. 2015;22(3):588-91.
27. Magerl W, Krumova EK, Baron R, Tolle T, Treede RD, Maier C. Reference data for quantitative sensory testing (QST): refined stratification for age and a novel method for statistical comparison of group data. Pain. 2010;151(3):598-605.

28. Rolke R, Baron R, Maier C, Tolle TR, Treede RD, Beyer A, et al. Quantitative sensory testing in the German Research Network on Neuropathic Pain (DFNS): standardized protocol and reference values. Pain. 2006;123(3):231-43.
29. Siao P, Cros DP. Quantitative sensory testing. Physical Medicine and Rehabilitation Clinics. 2003;14(2):261-86.
30. Bakkers M, Faber CG, Peters MJ, Reulen JP, Franssen H, Fischer TZ, et al. Temperature threshold testing: a systematic review. J Peripher Nerv Syst. 2013;18(1):7-18.
31. Blankenburg M, Boekens H, Hechler T, Maier C, Krumova E, Scherens A, et al. Reference values for quantitative sensory testing in children and adolescents: developmental and gender differences of somatosensory perception. Pain. 2010;149(1):76-88.
32. Fillingim RB, Maixner W, Kincaid S, Silva S. Sex differences in temporal summation but not sensory-discriminative processing of thermal pain. Pain. 1998;75(1):121-7.
33. Wang Y, Mo X, Zhang J, Fan Y, Wang K, Peter S. Quantitative sensory testing (QST) in the orofacial region of healthy Chinese: influence of site, gender and age. Acta Odontologica Scandinavica. 2018;76(1):58-63.
34. Bartlett G, Stewart JD, Tamblyn R, Abrahamowicz M. Normal distributions of thermal and vibration sensory thresholds. Muscle & Nerve: Official Journal of the American Association of Electrodiagnostic Medicine. 1998;21(3):367-74.
35. Kelly KG, Cook T, Backonja MM. Pain ratings at the thresholds are necessary for interpretation of quantitative sensory testing. Muscle & Nerve: Official Journal of the American Association of Electrodiagnostic Medicine. 2005;32(2):179-84.
36. Lilliesköld VH, Nordh E. Method-of-limits; Cold and warm perception thresholds at proximal and distal body regions. Clinical neurophysiology practice. 2018;3:134-40.
37. Torgén M, Swerup C. Individual factors and physical work load in relation to sensory thresholds in a middle-aged general population sample. European journal of applied physiology. 2002;86(5):418-27.
38. Meier PM, Berde CB, DiCanzio J, Zurakowski D, Sethna NF. Quantitative assessment of cutaneous thermal and vibration sensation and thermal pain detection thresholds in healthy children and adolescents. Muscle & Nerve: Official Journal of the American Association of Electrodiagnostic Medicine. 2001;24(10):1339-45.
39. Gonzalez-Duarte A, Lem-Carrillo M, Guerrero-Torres L. Normative values of quantitative sensory testing in Hispanic Latino population. Brain Behav. 2016;6(7):e00466.
40. Powell-Roach KL, Yao Y, Rutherford JN, Schlaeger JM, Patil CL, Suarez ML, et al. Thermal and mechanical quantitative sensory testing values among healthy African American adults. J Pain Res. 2019;12:2511.
41. Hilz M, Stemper B, Axelrod F, Kolodny E, Neundörfer B. Quantitative thermal perception testing in adults. Journal of clinical neurophysiology. 1999;16(5):462.
42. Pfau DB, Krumova EK, Treede R-D, Baron R, Toelle T, Birklein F, et al. Quantitative sensory testing in the German Research Network on Neuropathic Pain (DFNS): reference data for the trunk and application in patients with chronic postherpetic neuralgia. PAIN®. 2014;155(5):1002-15.
43. Meh D, Denišlič M. Quantitative assessment of thermal and pain sensitivity. Journal of the neurological sciences. 1994;127(2):164-9.
44. Backonja M-M, Walk D, Edwards RR, Sehgal N, Moeller-Bertram T, Wasan A, et al. Quantitative sensory testing in measurement of neuropathic pain phenomena and other sensory abnormalities. The Clinical journal of pain. 2009;25(7):641-7.
45. Claus D, Hilz M, Hummer I, Neundörfer B. Methods of measurement of thermal thresholds. Acta neurologica scandinavica. 1987;76(4):288-96.
46. Hansson P, Backonja M, Bouhassira D. Usefulness and limitations of quantitative sensory testing: clinical and research application in neuropathic pain states. Pain. 2007;129(3):256-9.
47. Rolke R, Magerl W, Campbell KA, Schalber C, Caspari S, Birklein F, et al. Quantitative sensory testing: a comprehensive protocol for clinical trials. European journal of pain. 2006;10(1):77-.
48. Kemler MA, Schouten HJ, Gracely RH. Diagnosing Sensory Abnormalities with Either Normal Values or Values from Contralateral SkinComparison of Two Approaches in Complex Regional Pain Syndrome I. Anesthesiology: The Journal of the American Society of Anesthesiologists. 2000;93(3):718-27.
49. Stevens JC, Choo KK. Temperature sensitivity of the body surface over the life span. Somatosensory & motor research. 1998;15(1):13-28.
50. Mosley GL, Butler DS. Explain pain supercharged: NOI; 2017.
51. Apkarian AV, Baliki MN, Geha PY. Towards a theory of chronic pain. Progress in neurobiology. 2009;87(2):81-97.

52. Woolf CJ. Central sensitization: implications for the diagnosis and treatment of pain. Pain. 2011;152(3):S2-S15.
53. Walk D, Sehgal N, Moeller-Bertram T, Edwards RR, Wasan A, Wallace M, et al. Quantitative sensory testing and mapping: a review of nonautomated quantitative methods for examination of the patient with neuropathic pain. The Clinical journal of pain. 2009;25(7):632-40.
54. 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(11):934-44.
55. Cauda F, Palermo S, Costa T, Torta R, Duca S, Vercelli U, et al. Gray matter alterations in chronic pain: a network-oriented meta-analytic approach. NeuroImage: Clinical. 2014;4:676-86.
56. Feizerfan A, Sheh G. Transition from acute to chronic pain. Continuing Education in Anaesthesia, Critical Care & Pain. 2015;15(2):98-102.
57. Feher JJ. Quantitative human physiology: an introduction: Academic press; 2017.
58. Julius D, Basbaum AI. Molecular mechanisms of nociception. Nature. 2001;413(6852):203-10.
59. Penfield W, Boldrey E. Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. Brain. 1937;60(4):389-443.
60. Levine TD. Small Fiber Neuropathy: Disease Classification Beyond Pain and Burning. J Cent Nerv Syst Dis. 2018;10:1179573518771703-.
61. Lacomis D. Small-fiber neuropathy. Muscle & Nerve: Official Journal of the American Association of Electrodiagnostic Medicine. 2002;26(2):173-88.
62. Themistocleous AC, Ramirez JD, Serra J, Bennett DLH. The clinical approach to small fibre neuropathy and painful channelopathy. Practical Neurology. 2014;14(6):368.
63. Paneni F, Beckman JA, Creager MA, Cosentino F. Diabetes and vascular disease: pathophysiology, clinical consequences, and medical therapy: part I. Eur Heart J. 2013;34(31):2436-43.
64. Yagihashi S, Mizukami H, Sugimoto K. Mechanism of diabetic neuropathy: Where are we now and where to go? J Diabetes Investig. 2011;2(1):18-32.
65. Perzyńska-Mazan J, Maślińska M, Gasik R. Neurological manifestations of primary Sjögren's syndrome. Reumatologia. 2018;56(2):99-105.
66. Xianbin W, Mingyu W, Dong X, Huiying L, Yan X, Fengchun Z, et al. Peripheral neuropathies due to systemic lupus erythematosus in China. Medicine (Baltimore). 2015;94(11):e625-e.
67. Eggermann K, Gess B, Häusler M, Weis J, Hahn A, Kurth I. Hereditary Neuropathies. Dtsch Arztebl Int. 2018;115(6):91-7.
68. Evans SR, Ellis RJ, Chen H, Yeh T-m, Lee AJ, Schifitto G, et al. Peripheral neuropathy in HIV: prevalence and risk factors. AIDS. 2011;25(7):919-28.
69. Chong PS, Cros DP. Technology literature review: quantitative sensory testing. Muscle Nerve. 2004;29(5):734-47.
70. Devigili G, Rinaldo S, Lombardi R, Cazzato D, Marchi M, Salvi E, et al. Diagnostic criteria for small fibre neuropathy in clinical practice and research. Brain. 2019;142(12):3728-36.
71. Jamal GA, Hansen S, Weir AI, Ballantyne JP. The neurophysiologic investigation of small fiber neuropathies. Muscle & Nerve: Official Journal of the American Association of Electrodiagnostic Medicine. 1987;10(6):537-45.
72. Tobin K, Giuliani MJ, Lacomis D. Comparison of different modalities for detection of small fiber neuropathy. Clinical neurophysiology. 1999;110(11):1909-12.
73. Holland NR, Stocks A, Hauer P, Cornblath DR, Griffin JW, McArthur JC. Intraepidermal nerve fiber density in patients with painful sensory neuropathy. Neurology. 1997;48(3):708-11.
74. Lefaucheur J-P, Wahab A, Planté-Bordeneuve V, Sene D, Ménard-Lefaucheur I, Rouie D, et al. Diagnosis of small fiber neuropathy: a comparative study of five neurophysiological tests. Neurophysiologie Clinique/Clinical Neurophysiology. 2015;45(6):445-55.
75. Scherens A, Maier C, Haussleiter IS, Schwenkreis P, Vlckova-Moravcova E, Baron R, et al. Painful or painless lower limb dysesthesias are highly predictive of peripheral neuropathy: comparison of different diagnostic modalities. European Journal of Pain. 2009;13(7):711-8.
76. Hassanzadeh J, Rezaianzadeh A. Assessing the validity of diagnostic tests. Iran J Med Sci. 2012;37(1):2.
77. Lauria G, Cornblath DR, Johansson O, McArthur JC, Mellgren SI, Nolano M, et al. EFNS guidelines on the use of skin biopsy in the diagnosis of peripheral neuropathy. European Journal of Neurology. 2005;12(10):747-58.
78. Pittenger GL, Ray M, Burcus NI, McNulty P, Basta B, Vinik AI. Intraepidermal Nerve Fibers Are Indicators of Small-Fiber Neuropathy in Both Diabetic and Nondiabetic Patients. Diabetes Care. 2004;27(8):1974.

79. Periquet M, Novak V, Collins M, Nagaraja H, Erdem S, Nash S, et al. Painful sensory neuropathy: prospective evaluation using skin biopsy. Neurology. 1999;53(8):1641-.
80. Løseth S, Lindal S, Stålberg E, Mellgren S. Intraepidermal nerve fibre density, quantitative sensory testing and nerve conduction studies in a patient material with symptoms and signs of sensory polyneuropathy. European journal of neurology. 2006;13(2):105-11.
81. Yarnitsky D, Sprecher E. Thermal testing: normative data and repeatability for various test algorithms. Journal of the neurological sciences. 1994;125(1):39-45.
82. Bakkers M, Faber CG, Reulen JPH, Hoeijmakers JGJ, Vanhoutte EK, Merkies ISJ. Optimizing temperature threshold testing in small-fiber neuropathy. Muscle & Nerve. 2015;51(6):870-6.
83. Wasner G, Schattschneider J, Binder A, Baron R. Topical menthol—a human model for cold pain by activation and sensitization of C nociceptors. Brain. 2004;127(5):1159-71.
84. Yarnitsky D, Ochoa JL. Release of cold-induced burning pain by block of cold-specific afferent input. Brain. 1990;113(4):893-902.
85. Agostinho CMS, Scherens A, Richter H, Schaub C, Rolke R, Treede RD, et al. Habituation and short-term repeatability of thermal testing in healthy human subjects and patients with chronic non-neuropathic pain. European Journal of Pain. 2009;13(8):779-85.
86. Gerr F, Letz R. Covariates of human peripheral nerve function: II. Vibrotactile and thermal thresholds. Neurotoxicology and teratology. 1994;16(1):105-12.
87. Neziri AY, Scaramozzino P, Andersen OK, Dickenson AH, Arendt-Nielsen L, Curatolo M. Reference values of mechanical and thermal pain tests in a pain-free population. Eur J Pain. 2011;15(4):376-83.
88. Gelber DA, Pfeifer MA, Broadstone VL, Munster EW, Peterson M, Arezzo JC, et al. Components of variance for vibratory and thermal threshold testing in normal and diabetic subjects. Journal of Diabetes and its Complications. 1995;9(3):170-6.
89. Moravcová E, Bednarik J, Svobodník A, Dušek L. Reproducibility of thermal threshold assessment in small-fibre neuropathy patients. Scr Med(Brno). 2005;78(3):177-84.
90. Armstrong F, Bradbury J, Ellis S, Owens D, Rosen I, Sonksen P, et al. A study of peripheral diabetic neuropathy. The application of age-related reference values. Diabetic medicine. 1991;8(S2):S94-S9.
91. Valensi P, Attali JR, Gagant S, Research FGf, Neuropathy SoD. Reproducibility of parameters for assessment of diabetic neuropathy. Diabetic medicine. 1993;10(10):933-9.
92. Geber C, Klein T, Azad S, Birklein F, Gierthmühlen J, Huge V, et al. Test–retest and interobserver reliability of quantitative sensory testing according to the protocol of the German Research Network on Neuropathic Pain (DFNS): a multi-centre study. PAIN®. 2011;152(3):548-56.
93. Valk GD, Grootenhuis PA, van Eijk JT, Bouter LM, Bertelsmann FW. Methods for assessing diabetic polyneuropathy: validity and reproducibility of the measurement of sensory symptom severity and nerve function tests. Diabetes Res Clin Pract. 2000;47(2):87-95.
94. Wang R, Cui L, Zhou W, Wang C, Zhang J, Wang K, et al. Reliability study of thermal quantitative sensory testing in healthy Chinese. Somatosens Mot Res. 2014;31(4):198-203.
95. Wasner GL, Brock JA. Determinants of thermal pain thresholds in normal subjects. Clinical Neurophysiology. 2008;119(10):2389-95.
96. Zwart J-A, Sand T. Repeatability of dermatomal warm and cold sensory thresholds in patients with sciatica. European Spine Journal. 2002;11(5):441-6.
97. Kemler MA, Reulen JP, van Kleef M, Barendse GA, van den Wildenberg FA, Spaans F. Thermal thresholds in complex regional pain syndrome type I: sensitivity and repeatability of the methods of limits and levels. Clinical neurophysiology. 2000;111(9):1561-8.
98. Doeland HJ, Nauta JJ, van Zandbergen JB, van der Eerden HA, van Diemen NG, Bertelsmann FW, et al. The relationship of cold and warmth cutaneous sensation to age and gender. Muscle & Nerve: Official Journal of the American Association of Electrodiagnostic Medicine. 1989;12(9):712-5.
99. Harju E-L. Cold and warmth perception mapped for age, gender, and body area. Somatosensory & motor research. 2002;19(1):61-75.
100. Huang H-W, Wang W-C, Lin C-CK. Influence of age on thermal thresholds, thermal pain thresholds, and reaction time. Journal of Clinical Neuroscience. 2010;17(6):722-6.
101. Lautenbacher S, Kunz M, Strate P, Nielsen J, Arendt-Nielsen L. Age effects on pain thresholds, temporal summation and spatial summation of heat and pressure pain. Pain. 2005;115(3):410-8.
102. Dyck P, Karnes J, O'brien P, Zimmerman I. Detection thresholds of cutaneous sensation in humans. Peripheral neuropathy. 1993;1:706-28.

103. Fagius J, Wahren LK. Variability of sensory threshold determination in clinical use. Journal of the neurological sciences. 1981;51(1):11-27.
104. Chakour M, Gibson S, Bradbeer M, Helme R. The effect of age on Aδ-and C-fibre thermal pain perception. Pain. 1996;64(1):143-52.
105. Yarnitsky D, Sprecher E, Zaslansky R, Hemli JA. Heat pain thresholds: normative data and repeatability. Pain. 1995;60(3):329-32.
106. Martin VT. Ovarian hormones and pain response: a review of clinical and basic science studies. Gender medicine. 2009;6:168-92.
107. Soudmand R, Ward LC, Swift TR. Effect of height on nerve conduction velocity. Neurology. 1982;32(4):407-.
108. Sulaxane YD, Bhavasar RP. Impact of height on sensory nerve conduction. National Journal of Physiology, Pharmacy and Pharmacology. 2017;7(8):851-4.
109. Campbell Jr WW, Ward LC, Swift TR. Nerve conduction velocity varies inversely with height. Muscle & Nerve: Official Journal of the American Association of Electrodiagnostic Medicine. 1981;4(6):520-3.
110. Robinson L, Rubner D, Wahl P, Fujimoto W, Stolov W. Influences of height and gender on normal nerve conduction studies. Archives of physical medicine and rehabilitation. 1993;74(11):1134-8.
111. Hagander LG, Midani HA, Kuskowski MA, Parry GJ. Quantitative sensory testing: effect of site and skin temperature on thermal thresholds. Clin Neurophysiol. 2000;111(1):17-22.
112. Rosner B. Fundamentals of biostatistics: Nelson Education; 2015.
113. Mukaka MM. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. Malawi medical journal : the journal of Medical Association of Malawi. 2012;24(3):69-71.
114. Leffler AS, Hansson P. Painful traumatic peripheral partial nerve injury-sensory dysfunction profiles comparing outcomes of bedside examination and quantitative sensory testing. European Journal of Pain. 2008;12(4):397-402.
115. Maier C, Baron R, Tölle T, Binder A, Birbaumer N, Birklein F, et al. Quantitative sensory testing in the German Research Network on Neuropathic Pain (DFNS): somatosensory abnormalities in 1236 patients with different neuropathic pain syndromes. Pain. 2010;150(3):439-50.
116. Bury TJ, Stokes EK. A global view of direct access and patient self-referral to physical therapy: implications for the profession. Phys Ther. 2013;93(4):449-59.
117. Chimenti RL, Frey-Law LA, Sluka KA. A Mechanism-Based Approach to Physical Therapist Management of Pain. Phys Ther. 2018;98(5):302-14.
118. Brown CL, Gilbert KK, Brismee JM, Sizer PS, Roger James C, Smith MP. The effects of neurodynamic mobilization on fluid dispersion within the tibial nerve at the ankle: an unembalmed cadaveric study. J Man Manip Ther. 2011;19(1):26-34.
119. Gilbert KK, Smith MP, Sobczak S, James CR, Sizer PS, Brismee JM. Effects of lower limb neurodynamic mobilization on intraneural fluid dispersion of the fourth lumbar nerve root: an unembalmed cadaveric investigation. J Man Manip Ther. 2015;23(5):239-45.
120. Kluding PM, Pasnoor M, Singh R, Jernigan S, Farmer K, Rucker J, et al. The effect of exercise on neuropathic symptoms, nerve function, and cutaneous innervation in people with diabetic peripheral neuropathy. Journal of diabetes and its complications. 2012;26(5):424-9.
121. Hanewinckel R, van Oijen M, Ikram MA, van Doorn PA. The epidemiology and risk factors of chronic polyneuropathy. Eur J Epidemiol. 2016;31(1):5-20.
122. Bender R, Lange S. Adjusting for multiple testing—when and how? Journal of clinical epidemiology. 2001;54(4):343-9.
123. Rothman KJ. No adjustments are needed for multiple comparisons. Epidemiology. 1990:43-6.
124. Saville DJ. Multiple comparison procedures: the practical solution. The American Statistician. 1990;44(2):174-80.

8.0 Appendix

8.1 Appendix A: Research paper – original draft for publication

Title of the paper

Can within-subject comparisons of thermal thresholds be used for diagnostic purposes?

Authors

Ø. Dunker

Journal

European journal of pain

https://onlinelibrary.wiley.com/journal/15322149

Title of the paper

Can within-subject comparisons of thermal thresholds be used for diagnostic purposes?

Running head

Within-subject comparisons of thermal thresholds

Authors and institutions

Ø. Dunker[1,2], M.U. Lie[1,4], K.B. Nilsen[1,3]

1Research and Communication Unit for Musculoskeletal Health (FORMI), Oslo University Hospital, Oslo, Norway

2Oslo Metropolitan University

3Department of Neurology and Clinical Neurophysiology, Oslo University Hospital, Oslo, Norway

4Faculty of Medicine, University of Oslo, Oslo, Norway

Author responsible for correspondence

Øystein Dunker, Research and Communication Unit for Musculoskeletal Health (FORMI), Oslo University Hospital, Pb. 4956 Nydalen, 0424 Oslo, Norway. Email: oeystein.dunker@gmail.com

Category

Original article

Funding

Funding from Oslo University Hospital and Oslo Metropolitan University

Conflicts of interest

None

Significance

- The limits of normality for side-differences and distal-proximal comparisons are wide

- Our findings underline the need for site-specific reference values of adequate resolution

- The difference between distal and proximal thermal thresholds increase drastically with age, and should be accounted for in clinical practice

Abstract

Background: Quantitative thermal testing (QTT) is a psychophysical assessment method of small nerve fibers that relies on reference material to assess function. Normal limits for within-subject comparisons of thermal thresholds are lacking, and their association with age, height and sex is largely unknown. The aim of this study was to investigate the normal limits for distal-proximal– and contralateral homologous comparisons of thermal thresholds with QTT, and their association with age, sex or height.

Methods: Forty-eight healthy volunteers ages 20-79 participated in the experiment. Cold detection thresholds (CDT), warm detection thresholds (WDT), heat pain thresholds (HPT), and cold pain thresholds (CPT) were measured bilaterally at the thenar eminence, anterior thigh, distal medial leg and foot dorsum. Sample normal limits were calculated as mean ± 2 SD.

Results: CPT was excluded from all analysis due to a large floor-effect. Sample normal limits for side-differences ranged from 2.0–7.4°C for CDT, 2.9–6.8°C for WDT and 3.2–4.6°C for HPT, depending on anatomical site. For distal-proximal comparisons, sample normal limits ranged from 4.9–8.7°C for CDT, 6.0–14.0°C for WDT and 4.2–9.0°C for HPT. Age was associated with side-differences for CDT in the thenar eminences ($p < 0.001$) and distal medial legs ($p < 0.002$), and with 11 of 18 distal-proximal comparisons ($p < 0.01$).

Conclusions: The normal limits for distal-proximal- and contralateral homologous thermal thresholds were wide. Age, but not sex or height, was associated with contralateral comparisons of CDT in the thenar eminences and distal medial legs, and with most distal-proximal comparisons.

1.0 INTRODUCTION

Pain due to lesions of small nerve-fibers is an inherent part of many pain syndromes, and a definite diagnosis may depend on the adequate testing of small-fiber function. Quantitative thermal testing (QTT) is an assessment method of sensory function in small, thinly myelinated (A-delta) and unmyelinated (C) nerve fibers, including their central pathways (Krumova et al., 2012). The ability to detect neuropathies and small-fiber lesions may be improved by using QTT in concordance with other tests, such as measurements of intra-epidermal nerve fiber density or nerve conduction studies (Backonja et al., 2013; Krumova et al., 2012; Lefaucheur et al., 2015; Løseth et al., 2006; Scherens et al., 2009; Siao and Cros 2003).

Valid reference values are an important prerequisite for QTT. Although the literature is somewhat equivocal with regards to covariates, adjusting for age (Guergova and Dufour 2011; Hafner et al., 2015; Magerl et al., 2010; Rolke et al., 2006a; Siao and Cros 2003), sex (Bakkers et al., 2013; Blankenburg et al., 2010; Fillingim et al., 1998; Siao and Cros 2003; Wang et al., 2018) and possibly height (Bartlett et al., 1998; Kelly et al., 2005; Lilliesköld and Nordh 2018; Torgén and Swerup 2002) is advised. Consequently, a wide range of distinct reference materials have been reported, e.g. for children and adolescents (Blankenburg et al., 2010; Meier et al., 2001), Hispanic Latino and African American populations (Gonzalez-Duarte et al., 2016; Powell-Roach et al., 2019), wide age-spans (Bartlett et al., 1998; Hilz et al., 1999; Pfau et al., 2014; Rolke et al., 2006a), and a number of anatomical sites (Hilz et al., 1999; Lilliesköld and Nordh 2018; Meh and Denišlič 1994). Additionally, efforts to standardize experimental variables and methods of analysis have been made (Magerl et al., 2010; Pfau et al., 2014; Rolke et al., 2006a; Rolke et al., 2006b). Yet for their differences, the reference materials mostly share the commonality of wide normal limits.

Within-subject comparisons across anatomical- or contralateral sites are also possible with QTT. Such relative comparisons could be beneficial if they showed lower variability, or if they were unassociated with common covariates for thermal thresholds. Previous investigations have failed to find significant side-differences for QTT (Kemler et al., 2000; Lilliesköld and Nordh 2018; Meh and Denišlič 1994; Rolke et al., 2006a) and it has been well established that there are differences between anatomical sites, with a likely distal-proximal gradient of increasing sensitivity (Bakkers et al., 2013; Meh and Denišlič 1994; Siao and Cros 2003; Stevens and Choo 1998). However, attempts to quantify normal limits for such comparisons are scarce (Kemler et al., 2000), and the association with age, sex and height is largely untested. This crucial research gap should be explored further, as part of the clinical value of within-subject comparisons of thermal thresholds rests on the knowledge of normal variability in healthy individuals.

Thus, the aim of this study is to investigate the normal limits for distal-proximal– and contralateral homologous comparisons of thermal thresholds with QTT, and their association with age, sex or height.

2 METHODS

2.1 Study design

An experimental, cross-sectional study was designed to compare thermal thresholds with regards to side-differences and distal-proximal gradients. Four sites were measured bilaterally: the thenar eminence; the anterior thigh, 10 cm superior to the patellar base in mid-line; the distal medial leg, directly superior and posterior to the medial malleolus, and; the foot dorsum, at the dorsal aspect of metatarsals II-III.

The protocol included cold detection threshold (CDT), warm detection threshold (WDT), heat pain threshold (HPT) and cold pain threshold (CPT). Testing order of the specific sites was randomized in advance, and a pre-test of two cold- and warm detection threshold stimuli was performed to familiarize the subjects with the procedure.

A single, male experimenter carried out all experiments. The QTT protocol, placement of instruments, room temperature, experimenter's clothing, the instructions and lighting was standardized. Participants were blinded to the study's hypotheses and instrument readouts.

The study was approved by Regional Committees for Medical and Health Research Ethics (REC), project no. 2010/2927. All participants provided written consent, and the study was conducted in accordance with the Declaration of Helsinki. Subjects received a gift certificate of NOK 250 for participation.

2.2 Subjects

The required sample size for paired comparisons were calculated in accordance with Rosner (2015). For side-differences, a minimal clinical significance of 1°C and standard deviation of 2°C was used, while these values were 2°C and 3°C for distal-proximal comparisons, respectively. A minimum of 31 subjects were needed to detect a side-difference of 1°C, with type I and II error rates of 0.05 and 0.2, respectively, and 18 were needed for distal-proximal comparisons.

Healthy men and women, ages 20-79, were recruited through advertisements at Oslo University Hospital, local universities, gyms, centers for the elderly, and on social media. Recruitment efforts focused on achieving an equivalent representation of sex and ages. Exclusion criteria were: cancer (current or previously), diabetes, radiculopathy, chronic pain (average NRS ≥ 1 for ≥ 3 months, last two years), pregnant or breastfeeding, limited capacity for consent, personal acquaintance of experimenter, or any disease of nerves, muscles, or of the brain that could influence normal nervous function, including psychiatric illnesses. Subjects were requested not to work nightshifts within 48 hours of the experiment, to not consume alcohol in the last 12 hours before the experiment, or consume pain-killers the same day as the experiment.

2.3 Test protocol

Thermal stimulus was applied with a 30 x 30 mm Peltier thermode (Medoc, Ramat Yishai, Israel), by method of limits. The thermode was held in place by the experimenter. Baseline temperature was 32°C, with a range of 0-52°C. The ramp-rate was 1°C/s, and returned to baseline at rates of 1°C/s for detection thresholds and 5°C/s for pain thresholds. Inter-stimulus-intervals were 4-6 seconds for all tests.

Subjects lay supine on a treatment table, with the back rest at approximately 120-135° incline. Pillows were used for head-support and placed under the subjects' knees, and a duvet helped regulate skin temperature. Immediately prior to testing, the skin temperature was measured at each site with an 826-T2 hand-held infrared thermometer (Testo SE & Co., Pennsylvania, USA), held perpendicular to the skin's surface at a standardized distance of 1 cm. A re-usable heat pack was applied where skin temperature was <30°C for the lower extremities and <32°C for the thenar eminence. Excessive body-hair was removed with scissors.

CDT, WDT and HPT were measured in succession for each site, followed by CPT in the distal medial legs and feet dorsa, in the same order. Absolute temperature thresholds were recorded.

Scripted, verbal instructions were used. Participants were informed of the procedure in its entirety before testing began, and reminded of the current modality before each test. For CDT and WDT, subjects were asked to press a trigger at the first sensation of cool or warmth. Similarly, for HPT and CPT, the cue was to press the trigger at the first sensation of pain, typically when the thermal stimulus begins to induce a stinging, burning or aching sensation. Subjects were advised that thermal pain thresholds are not a measure of pain tolerance. A response was considered invalid and repeated once if it deviated substantially from contemporaneous measurements, or if the subject admitted to an accidental response.

2.4 Data computation and analysis

Statistical analyses were performed using IBM SPSS Statistics v. 25 (Armonk, NY: IBM Corp.) P-values were regarded as significant at ≤ 0.05, and Bonferroni correction for multiple testing was applied. Correlation values were interpreted in accordance with Mukaka (2012): negligible correlation ±0.0–0.3, low correlation ±0.3–0.5, moderate correlation ±0.5–0.7, high correlation ±0.7–0.9 and very high correlation ±0.9–1.0. Thermal thresholds were calculated to express absolute change from baseline (Δ°C).

Data distribution was assessed in preliminary analyses by use of histograms, boxplots and Q-Q plots. The arithmetic mean of five (CDT, WDT) or three (HPT, CPT) measurements was used in the analysis. Data from each subject was excluded from the calculation of sample thresholds and for performing linear regression if the delta value of a thermal threshold was > 3 times the arithmetic mean (detection thresholds)-, <1/3 the arithmetic mean (pain thresholds)-, or if exceeding ± 3 SD from the arithmetic mean of the remaining data points. Invalid measurements due to a test's floor- or ceiling effect were not replaced or included in the final analysis.

Side-differences were determined by use of multiple paired t-tests. The distal-proximal gradients were examined by repeated measures ANOVA with post-hoc analysis for pairwise comparisons. Sample normal limits were calculated as mean ± 2 SD.

Pearson- or Spearman correlation was used to determine the association between side-differences or distal-proximal gradients, and age, sex and height.

Equations for determining sample normal limits (Δ°C) for CDT and WDT as a function of age in the distal medial leg and foot dorsum, as well as for distal-proximal comparisons between the foot dorsum and anterior thigh, were calculated by use of linear regression. Square root transformations were applied to the data on CDT- and WDT in the foot dorsum. Normal limit was defined as the 95% prediction interval.

3 RESULTS

3.1 Study sample

Forty-eight subjects were included in the analysis. Participants were 46 (SD ± 15.6) years old, 52% of whom female. The inclusion process is displayed in Figure 1 and sample characteristics are presented in Table 1.

FIGURE 1

TABLE 1

Two subjects were excluded from the analysis of CDT in the distal medial leg, 34 from CPT in the foot dorsum, and 37 from CPT in the distal medial leg, due to reaching the test's floor value of 0°C. Consequently, Table 2 shows sample normal limits, but the large floor-effect for CPT precluded calculation of sample CPT thresholds or relative CPT comparisons.

Some outlying data was identified and removed for CDT in the foot dorsum (n = 2), distal medial leg (n = 5) and thenar eminence (n = 2); for WDT in anterior thigh (n = 2) and thenar eminence (n = 1); and for HPT in anterior thigh (n = 1) and thenar eminence (n = 1).

TABLE 2

3.2 Side-differences for quantitative thermal testing

Side-differences in thermal thresholds for right *vs.* left are presented in Table 3. A significant difference of -1.4°C (-2.1°C – -0.6°C) was found for WDT in the feet dorsa ($p = 0.001$), significant at $p < 0.05$ after Bonferroni correction for multiple testing. Sample normal limits for side-differences ranged from 2.0–7.4°C for CDT, 2.9–6.8°C for WDT and 3.2–4.6°C for HPT.

TABLE 3

3.3 Distal-proximal gradients for quantitative thermal testing

Repeated measures ANOVA with Greenhouse-Geisser correction detected differences in means when comparing CDT ($F(2.2, 21.3)$), $p < 0.001$), WDT ($F(2.0, 125)$), $p < 0.001$) and HPT ($F(1.7, 14.8)$, $p < 0.001$). Post-hoc analysis with Bonferroni correction for multiple testing displayed a general distal-proximal gradient for CDT, WDT and HPT (Table 4). The distal-proximal gradient's linearity was violated by the distal medial leg being less sensitive than the foot dorsum for WDT ($p < 0.001$) and HPT ($p < 0.001$), with no difference detected for CDT ($p = 0.170$) (Figure 2). In addition, no significant difference was found between the foot dorsum and anterior thigh ($p = 0.932$) or thenar eminence ($p = 0.016$) for HPT. Sample normal limits for distal-proximal comparisons ranged from 4.9–8.7°C for CDT, 6.0–14.0°C for WDT and 4.2–9.0°C for HPT.

TABLE 4

FIGURE 2

3.4 Relation to age, sex and height

A low correlation was found between age and side-differences for CDT between the thenar eminences ($r = 0.46$, $p = 0.001$) and distal medial legs ($r = 0.44$, $p = 0.002$). Age was moderately correlated with the CDT gradient between the foot dorsum and anterior thigh ($r = 0.52$, $p < 0.001$), the foot dorsum and thenar eminence ($r = 0.59$, $p < 0.001$) and between the distal medial leg and thenar eminence ($r = 0.52$, $p < 0.001$); with the WDT gradient between the foot dorsum and anterior thigh ($\rho = 0.65$), the foot dorsum and thenar eminence ($r = 0.65$, $p < 0.001$), the distal medial leg and anterior thigh ($r = 0.52$, $p < 0.001$) and between the distal medial leg and thenar eminence ($r = 0.50$, $p < 0.001$); and with the HPT gradient between the foot dorsum and anterior thigh ($r = 0.53$, $p < 0.001$), the distal medial leg and thenar eminence ($r = 0.59$, $p < 0.001$) and between the anterior thigh and thenar eminence ($r = 0.56$, $p < 0.001$). A high correlation between age and thermal threshold gradient was found for HPT between the foot dorsum and thenar eminence ($r = 0.72$, $p < 0.001$).

No significant correlations were found for height or sex.

3.5 Prediction of normal limits by linear regression

Based on our sample, predicted normal limits for CDT and WDT in the foot dorsum and distal medial leg, as well as for comparisons between the foot dorsum and the anterior thigh, are presented in TableS1 in the appendix.

4 DISCUSSION

4.1 Side-differences

We found a side-difference of -1.4°C (95% CI -2.1 – -0.6°C) for WDT in the feet dorsa (right *vs.* left), while all other comparisons were non-significant. The large inter-subject variability resulted in sample normal limits that ranged from 2.0–7.4°C for CDT, 2.9–6.8°C for WDT and 3.0–4.6°C for HPT, with the distal medial calves and feet dorsa predominantly represented in the higher end. Based on the present study, this means that e.g. a WDT difference between the distal medial calves of >6.1°C would be necessary to claim that the difference is abnormal, while the same is true for HPT at >3.2°C. This is in line with similar studies that have shown that the sensitivity gains from contralateral comparisons are negligible for thermal detection thresholds, but may be advantageous for HPT, compared to absolute reference data (Blankenburg et al., 2010; Pfau et al., 2014; Rolke et al., 2006a).

With most studies failing to find side-differences for QTT (Blankenburg et al., 2010; Claus et al., 1987; Gelber et al., 1995; Kemler et al., 2000; Lilliesköld and Nordh 2018; Meh and Denišlič 1994; Neziri et al., 2011; Pfau et al., 2014; Rolke et al., 2006a), it could seem reasonable to compare sides in cases of unilaterally altered somatosensory function. Indeed, assessing asymmetries is a basic principle of the neurological examination, and has been recommended for QTT (Backonja et al., 2013; Hansson et al., 2007; Kemler et al., 2000). Such a relative comparison in the clinic would have the strengths of identical equipment, protocol and setting, in addition to the advantage of being one's own neuropsychological control (Backonja et al., 2013). Not surprisingly then, some reports show that contralateral homologous comparisons can increase the diagnostic sensitivity of QTT, either on its own, or as a supplement when the patient's results are within normal ranges of absolute reference values (Blankenburg et al., 2010; Maier et al., 2010; Rolke et al., 2006a). However, although inter-subject variability may yet prove to be less pronounced for within-subject comparisons than in absolute reference material, they are still considerable. Indeed, our data shows that contralateral homologous comparisons are also subject to wide limits of normality for most regions and thermal modalities, which diminishes their clinical utility.

Although we cannot rule out that a side-difference may exist, we expect our singular finding to be due to an unknown systematic error, as there are few reasons as to why a side-difference should exist for WDT in the feet dorsa of healthy individuals, and other studies have not found such a difference. A possible explanation is the use of a relatively small thermode (30x30mm), and subsequent low spatial resolution on the uneven surface of the foot dorsum. Through a systematic difference in how the thermode was manually applied by the experimenter, small areas with a lower warmth receptor density could inadvertently be targeted, or spatial summation could be influenced by varying degrees of contact with the skin (Backonja et al., 2013; Bakkers et al., 2013; Gelber et al., 1995).

An important limitation of comparing sides is that the contralateral site must be normal, limiting its use in e.g. symmetrical distal polyneuropathies. Besides, both pain and functional alterations can spread with time,

i.e. enlarging the affected area or mirroring the pathology (Jancalek 2011), for instance in complex regional pain syndrome (Maleki et al., 2000; Rasmussen et al., 2018; Veldman and Goris 1996) or post-herpetic neuralgia (Oaklander et al., 1998). In such cases, another relative approach is possible: to compare thermal thresholds across anatomical sites.

4.2 Distal-proximal comparisons

A general distal-proximal gradient of increasing thermal sensitivity, with differences as high as 100-fold between face and feet, has been well established (Bakkers et al., 2013; Blankenburg et al., 2010; Claus et al., 1987; Kelly et al., 2005; Lilliesköld and Nordh 2018; Meh and Denišlič 1994; Stevens and Choo 1998; Yarnitsky and Sprecher 1994). Although a few comparisons revealed no differences been anatomical sites, i.e. CDT foot dorsum-distal medial leg, HPT foot dorsum-anterior thigh and HPT foot dorsum-thenar eminence, our findings confirm the presence of a rough gradient. However, the data aligns with a previous report (Zhang et al., 2017), in showing that the linearity of the distal-proximal gradient of increasing sensitivity is violated by the calves; In the present study, the distal medial calf exhibited equal sensitivity to the foot dorsum for CDT, and was less sensitive for WDT and HPT, while the large floor effect deterred any conclusion regarding CPT. This has clinical implications, as for instance Maier et al., (2010) reports tentatively using absolute reference values for the hand and feet in the upper and lower body, respectively. According to our findings, e.g. utilizing reference values for the feet dorsum when assessing small-fiber function in the relatively adjacent distal medial legs, could result in an increase in false positives for WDT hypoesthesia or false negatives for HPT hyperalgesia. Consequently, we surmise that adequate care should be taken to ensure sufficiently high resolution of anatomical sites when creating absolute reference material or determining normal limits for inter-region comparisons, and we advise that exclusively site-specific reference values are used clinically, until the necessary resolution is fully elucidated.

It is uncertain whether inter-subject variability of thermal thresholds increase with distality, as equivocal findings can be drawn from previous research (Bartlett et al., 1998; Dyck et al., 1993; Lilliesköld and Nordh 2018; Meh and Denišlič 1994; Moravcová et al., 2005; Rolke et al., 2006b; Zhang et al., 2017). However, we show that this may be true for inter-region comparisons. This could mean that comparisons across anatomical sites are less useful when they involve the feet dorsa or distal medial calves, which, regrettably, are high-prevalence areas for distal neuropathies.

Rolke et al., (2006a) reported that region, i.e. face *vs.* foot had a larger effect on thermal thresholds than age or sex, but found no increase in sensitivity by comparing regions instead of using absolute reference data. It cannot be ruled out that the findings of Rolke et al., (2006a) are due to quite distal comparisons – that adjacency of the anatomical sites compared may influence diagnostic sensitivity somehow; Yet our findings of wide normal limits for neighboring anatomical sites raises doubts that this is the case. Indeed, comparisons of detection thresholds between adjacent sites in the lower extremities show normal limits of 6.0–11.2°C in our sample. Still, it must be noted that the aged part of our age-pooled sample may inflate

these limits somewhat, and larger samples that allow for age stratifications are needed in future research to more accurately explore inter-subject variability across age-groups.

4.3 Relation between relative comparisons and age, sex and height

No association was found between sex or height and side-differences or distal-proximal gradient, suggesting that these covariates need not be accounted for when establishing normal limits for within-subject comparisons. The finding of no relation between sex and side-differences is in line with previous investigations (Kemler et al., 2000). If sex and height is in fact inconsequential to the external validity of relative reference values, achieving sufficiently large sample sizes in the future may be easier than when creating absolute reference material.

In some contrast to Kemler et al., (2000), we found that age was significantly correlated with side-differences for CDT in the thenar eminences and distal medial legs. Furthermore, age was associated with most of the distal-proximal comparisons, with larger differences being normal in the older patient. The effect of age on distal-proximal comparisons may be due to faster ageing of the peripheral nerves, for instance age-related perfusion impairment near the peripheral receptors, age-related reduction in inter-epidermal nerve fiber density, distal axonopathy, or age-related changes in the synthesis, transport and action of key neurotransmitters leading to higher firing thresholds (Chakour et al., 1996; Guergova and Dufour 2011; Stevens and Choo 1998). Accordingly, it seems necessary to adjust for age when performing relative comparisons of thermal thresholds, particularly when assessing the most distal sites.

4.4 Strengths and limitations

The external validity of our findings is strengthened by our close adherence to the DNFS protocol, and limited by the low sample-size and failure to stratify by age.

Each trial lasted about one hour without breaks; It is thus possible that attention and reaction-time was diminished through the last tests, and since the test-order was randomized, this could theoretically lead to a global increase in inter-subject variability and widening of the sample normal limits.

5 CONCLUSION

The normal limits for distal-proximal- and contralateral homologous thermal thresholds were wide, advising caution for exclusive use in a clinical setting. Age, but not sex or height, was associated with most distal-proximal comparisons, and contralateral comparisons of CDT in the thenar eminences and distal medial legs. In addition, our data confirms that the distal medial legs may be equally– or less sensitive than the feet dorsa, resulting in a non-linear distal-proximal gradient in the lower extremities, and highlighting the need for site-specific reference values for thermal thresholds in general.

ACKNOWLEDGMENTS

The authors are grateful to Marie Skovli Pettersen for her continuous assistance with recruitment and inclusion, and to Sara Maria Allen for her work in adapting the experimental protocol and assuring the quality of the data collection.

AUTHOR CONTRIBUTIONS

Ø. D. takes responsibility for the integrity of the work as a whole. All authors contributed to conception and design, while Ø.D. performed acquisition- and analysis of data, and drafted the article.

6 REFERENCES

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, Walalce MS, Ware M, Yarnitsky D, Ziegler D. Value of quantitative sensory testing in neurological and pain disorders: NeuPSIG consensus. Pain 2013;154: 1807-1819 10.1016/j.pain.2013.05.047.

Bakkers M, Faber CG, Peters MJ, Reulen JP, Franssen H, Fischer TZ, Merkies IS. Temperature threshold testing: a systematic review. J Peripher Nerv Syst 2013;18: 7-18 10.1111/jns5.12001.

Bartlett G, Stewart JD, Tamblyn R, Abrahamowicz M. Normal distributions of thermal and vibration sensory thresholds. Muscle & Nerve: Official Journal of the American Association of Electrodiagnostic Medicine 1998;21: 367-374 10.1002/(sici)1097-4598(199803)21:3<367::aid-mus11>3.0.co;2-x.

Blankenburg M, Boekens H, Hechler T, Maier C, Krumova E, Scherens A, Magerl W, Aksu F, Zernikow B. Reference values for quantitative sensory testing in children and adolescents: developmental and gender differences of somatosensory perception. Pain 2010;149: 76-88 10.1016/j.pain.2010.01.011.

Chakour M, Gibson S, Bradbeer M, Helme R. The effect of age on Aδ-and C-fibre thermal pain perception. Pain 1996;64: 143-152 10.1016/0304-3959(95)00102-6.

Claus D, Hilz M, Hummer I, Neundörfer B. Methods of measurement of thermal thresholds. Acta neurologica scandinavica 1987;76: 288-296 10.1111/j.1600-0404.1987.tb03583.x.

Dyck P, Karnes J, O'brien P, Zimmerman I. Detection thresholds of cutaneous sensation in humans. Peripheral neuropathy 1993;1: 706-728

Fillingim RB, Maixner W, Kincaid S, Silva S. Sex differences in temporal summation but not sensory-discriminative processing of thermal pain. Pain 1998;75: 121-127 10.1016/s0304-3959(97)00214-5.

Gelber DA, Pfeifer MA, Broadstone VL, Munster EW, Peterson M, Arezzo JC, Shamoon H, Zeidler A, Clements R, Greene DA. Components of variance for vibratory and thermal threshold testing in normal and diabetic subjects. Journal of Diabetes and its Complications 1995;9: 170-176 10.1016/1056-8727(94)00042-m.

Gonzalez-Duarte A, Lem-Carrillo M, Guerrero-Torres L. Normative values of quantitative sensory testing in Hispanic Latino population. Brain Behav 2016;6: e00466 10.1002/brb3.466.

Guergova S and Dufour A. Thermal sensitivity in the elderly: A review. Ageing Research Reviews 2011;10: 80-92 10.1016/j.arr.2010.04.009.

Hafner J, Lee G, Joester J, Lynch M, Barnes EH, Wrigley PJ, Ng K. Thermal quantitative sensory testing: a study of 101 control subjects. Journal of Clinical Neuroscience 2015;22: 588-591 10.1016/j.jocn.2014.09.017.

Hansson P, Backonja M, Bouhassira D. Usefulness and limitations of quantitative sensory testing: clinical and research application in neuropathic pain states. Pain 2007;129: 256-259 10.1016/j.pain.2007.03.030.

Hilz M, Stemper B, Axelrod F, Kolodny E, Neundörfer B. Quantitative thermal perception testing in adults. Journal of clinical neurophysiology 1999;16: 462 10.1097/00004691-199909000-00008.

Jancalek R. Signaling mechanisms in mirror image pain pathogenesis. Ann Neurosci 2011;18: 123-127 10.5214/ans.0972-7531.11183010.

Kelly KG, Cook T, Backonja MM. Pain ratings at the thresholds are necessary for interpretation of quantitative sensory testing. Muscle & Nerve: Official Journal of the American Association of Electrodiagnostic Medicine 2005;32: 179-184 10.1002/mus.20355.

Kemler MA, Schouten HJ, Gracely RH. Diagnosing Sensory Abnormalities with Either Normal Values or Values from Contralateral Skin Comparison of Two Approaches in Complex Regional Pain

Syndrome I. Anesthesiology: The Journal of the American Society of Anesthesiologists 2000;93: 718-727 10.1097/00000542-200009000-00021.
Krumova EK, Geber C, Westermann A, Maier C. Neuropathic pain: is quantitative sensory testing helpful? Current diabetes reports 2012;12: 393-402 10.1007/s11892-012-0282-7.
Lefaucheur J-P, Wahab A, Planté-Bordeneuve V, Sene D, Ménard-Lefaucheur I, Rouie D, Tebbal D, Salhi H, Créange A, Zouari H. Diagnosis of small fiber neuropathy: a comparative study of five neurophysiological tests. Neurophysiologie Clinique/Clinical Neurophysiology 2015;45: 445-455 10.1016/j.neucli.2015.09.012.
Lilliesköld VH and Nordh E. Method-of-limits; Cold and warm perception thresholds at proximal and distal body regions. Clinical neurophysiology practice 2018;3: 134-140 10.1016/j.cnp.2018.06.004.
Løseth S, Lindal S, Stålberg E, Mellgren S. Intraepidermal nerve fibre density, quantitative sensory testing and nerve conduction studies in a patient material with symptoms and signs of sensory polyneuropathy. European journal of neurology 2006;13: 105-111 10.1111/j.1468-1331.2006.01232.x.
Magerl W, Krumova EK, Baron R, Tolle T, Treede RD, Maier C. Reference data for quantitative sensory testing (QST): refined stratification for age and a novel method for statistical comparison of group data. Pain 2010;151: 598-605 10.1016/j.pain.2010.07.026.
Maier C, Baron R, Tölle T, Binder A, Birbaumer N, Birklein F, Gierthmühlen J, Flor H, Geber C, Huge V. Quantitative sensory testing in the German Research Network on Neuropathic Pain (DFNS): somatosensory abnormalities in 1236 patients with different neuropathic pain syndromes. Pain 2010;150: 439-450 10.1016/j.pain.2010.05.002.
Maleki J, LeBel AA, Bennett GJ, Schwartzman RJ. Patterns of spread in complex regional pain syndrome, type I (reflex sympathetic dystrophy). Pain 2000;88: 259-266 10.1016/s0304-3959(00)00332-8.
Meh D and Denišlič M. Quantitative assessment of thermal and pain sensitivity. Journal of the neurological sciences 1994;127: 164-169 10.1016/0022-510x(94)90069-8.
Meier PM, Berde CB, DiCanzio J, Zurakowski D, Sethna NF. Quantitative assessment of cutaneous thermal and vibration sensation and thermal pain detection thresholds in healthy children and adolescents. Muscle & Nerve: Official Journal of the American Association of Electrodiagnostic Medicine 2001;24: 1339-1345 10.1002/mus.1153.
Moravcová E, Bednarik J, Svobodník A, Dušek L. Reproducibility of thermal threshold assessment in small-fibre neuropathy patients. Scr Med(Brno) 2005;78: 177-184
Mukaka MM. Statistics corner: A guide to appropriate use of correlation coefficient in medical research. Malawi medical journal : the journal of Medical Association of Malawi 2012;24: 69-71 PMCID: PMC3576830.
Neziri AY, Scaramozzino P, Andersen OK, Dickenson AH, Arendt-Nielsen L, Curatolo M. Reference values of mechanical and thermal pain tests in a pain-free population. Eur J Pain 2011;15: 376-383 10.1016/j.ejpain.2010.08.011.
Oaklander AL, Romans K, Horasek S, Stocks A, Hauer P, Meyer RA. Unilateral postherpetic neuralgia is associated with bilateral sensory neuron damage. Annals of neurology 1998;44: 789-795 10.1002/ana.410440513.
Pfau DB, Krumova EK, Treede R-D, Baron R, Toelle T, Birklein F, Eich W, Geber C, Gerhardt A, Weiss T. Quantitative sensory testing in the German Research Network on Neuropathic Pain (DFNS): reference data for the trunk and application in patients with chronic postherpetic neuralgia. PAIN® 2014;155: 1002-1015 10.1016/j.pain.2014.02.004.
Powell-Roach KL, Yao Y, Rutherford JN, Schlaeger JM, Patil CL, Suarez ML, Shuey D, Angulo V, Carrasco J, Ezenwa MO. Thermal and mechanical quantitative sensory testing values among healthy African American adults. J Pain Res 2019;12: 2511 10.2147/JPR.S211855.

Rasmussen VF, Karlsson P, Drummond PD, Schaldemose EL, Terkelsen AJ, Jensen TS, Knudsen LF. Bilaterally Reduced Intraepidermal Nerve Fiber Density in Unilateral CRPS-I. Pain Med 2018;19: 2021-2030 10.1093/pm/pnx240.

Rolke R, Baron R, Maier C, Tolle TR, Treede RD, Beyer A, Binder A, Birbaumer N, Birklein F, Botefur IC, Braune S, Flor H, Huge V, Klug R, Landwehrmeyer GB, Magerl W, Maihofner C, Rolko C, Schaub C, Scherens A, Sprenger T, Valet M, Wasserka B. Quantitative sensory testing in the German Research Network on Neuropathic Pain (DFNS): standardized protocol and reference values. Pain 2006a;123: 231-243 10.1016/j.pain.2006.01.041.

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 2006b;10: 77-77 10.1016/j.ejpain.2005.02.003.

Rosner B. Fundamentals of biostatistics Nelson Education. 2015.

Scherens A, Maier C, Haussleiter IS, Schwenkreis P, Vlckova-Moravcova E, Baron R, Sommer C. Painful or painless lower limb dysesthesias are highly predictive of peripheral neuropathy: comparison of different diagnostic modalities. European Journal of Pain 2009;13: 711-718 10.1016/j.ejpain.2008.07.014.

Siao P and Cros DP. Quantitative sensory testing. Physical Medicine and Rehabilitation Clinics 2003;14: 261-286 10.1016/s1047-9651(02)00122-5.

Stevens JC and Choo KK. Temperature sensitivity of the body surface over the life span. Somatosensory & motor research 1998;15: 13-28 10.1080/08990229870925.

Torgén M and Swerup C. Individual factors and physical work load in relation to sensory thresholds in a middle-aged general population sample. European journal of applied physiology 2002;86: 418-427 10.1007/s00421-001-0567-z.

Veldman PH and Goris RJA. Multiple reflex sympathetic dystrophy. Which patients are at risk for developing a recurrence of reflex sympathetic dystrophy in the same or another limb. Pain 1996;64: 463-466 10.1016/0304-3959(95)00160-3.

Wang Y, Mo X, Zhang J, Fan Y, Wang K, Peter S. Quantitative sensory testing (QST) in the orofacial region of healthy Chinese: influence of site, gender and age. Acta Odontologica Scandinavica 2018;76: 58-63 10.1080/00016357.2017.1383511.

Yarnitsky D and Sprecher E. Thermal testing: normative data and repeatability for various test algorithms. Journal of the neurological sciences 1994;125: 39-45 10.1016/0022-510x(94)90239-9.

Zhang JW, Wang Q, Wang XF, Gao MM, Yun XP, Wu HF, Hong Y. Sensory thresholds at different sites of the foot: a valuable reference for neurologic examinations. Spinal Cord 2017;55: 396-398 10.1038/sc.2016.167.

Tables and figures

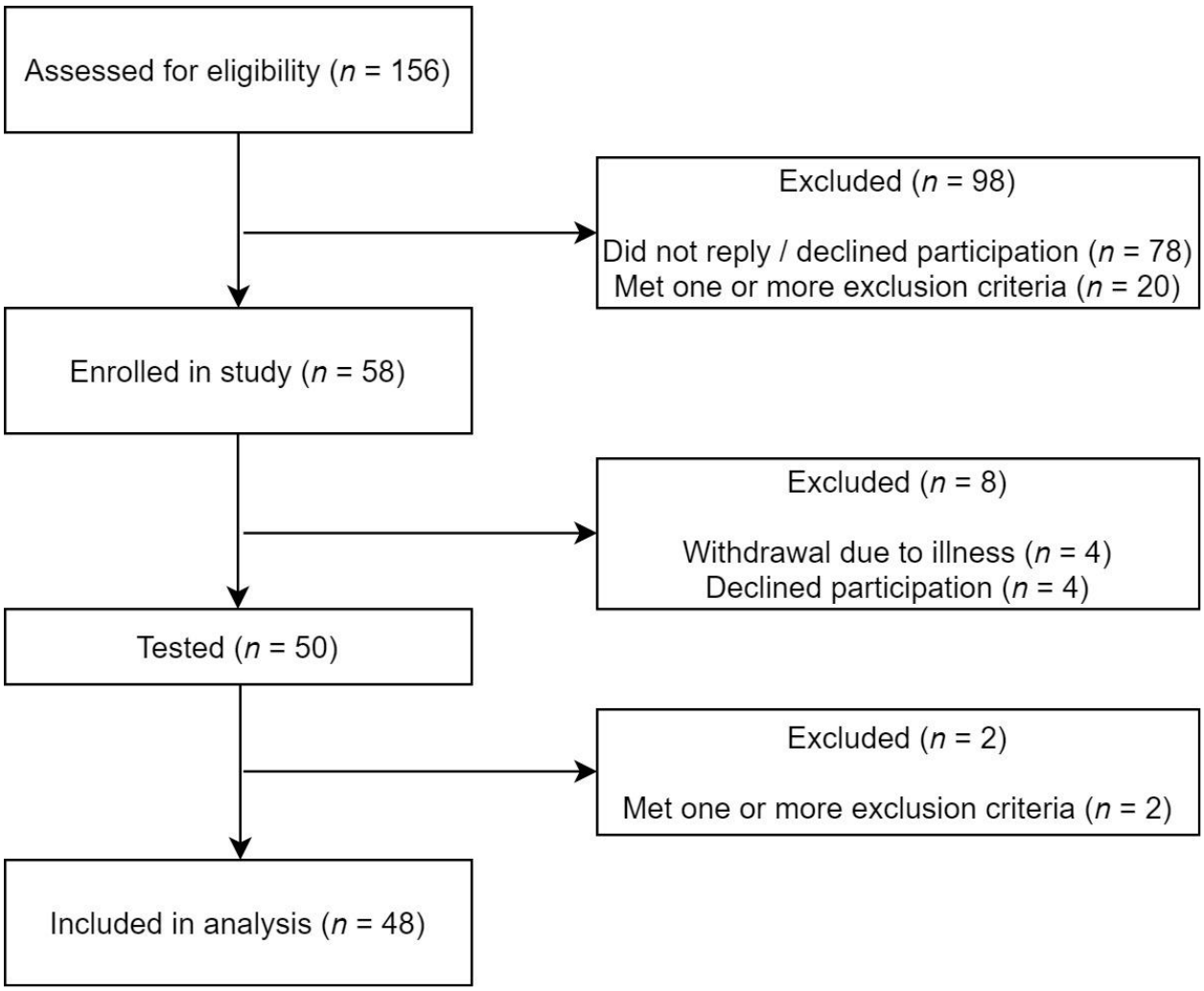

Figure 1: Flow chart of the study inclusion process

TABLE 1 Sample demographics (*n* = 48)

Variable	**n (%/SD)**
Sex, females	25 (52)
Age, years	46 (15.6)
Height, cm	174 (8.8)
BMI, kg/m², mean	25 (3.9)
Relationship status	
Married/partner	28 (58)
Single	20 (42)
Education	
Primary school, 10 years	0 (0)
High school, 1-2 years	0 (0)
High school, 3 years	0 (0)
Vocational high school	5 (10)
College/university <4 years	19 (40)
College/university ≥4 years	24 (50)
Current smoker, yes	3 (6)
Current snus-user, yes	5 (10)
Uses alcohol, yes	46 (96)
Abbreviations: SD = Standard Deviation	

TABLE 2 Sample thermal thresholds (Δ°C)

	Mean (SD)	Sample normal limit
CDT Foot dorsum[b]	4.1 (2.8)	9.7
WDT Foot dorsum	7.6 (3.8)	15.2
HPT Foot dorsum	13.5 (3.3)	6.9
CDT Distal medial leg[c]	4.0 (2.0)	8.0
WDT Distal medial leg	9.8 (3.4)	16.6
HPT Distal medial leg	14.8 (2.5)	9.8
CDT Anterior thigh	3.1 (1.7)	6.5
WDT Anterior thigh[b]	4.3 (1.5)	7.3
HPT Anterior thigh[a]	13.8 (2.7)	8.4
CDT Thenar eminence[b]	1.5 (0.8)	3.1
WDT Thenar eminence[a]	2.3 (1.0)	4.3
HPT Thenar eminence[a]	12.6 (3.5)	5.6

Abbreviations: SD = Standard Deviation
CDT = Cold Detection Threshold, WDT = Warm Detection Threshold, HPT = Heat Pain Threshold, Sample normal limit = Mean ± 2SD, CDT/WDT limit = Hypoesthesia, HPT limit = Hyperalgesia
[a]*n = 47*
[b]*n = 46*
[c]*n = 43*

TABLE 3 Side-differences for thermal thresholds (Δ°C)

Test site	Right side (mean, SD)	Left side (mean, SD)	Side-difference (SD)	Sample normal limit	P-value	95% CI
Cold detection threshold, °C						
Thenar eminence	0.9 (1.9)	0.8 (1.6)	-0.2 (0.9)	-2.0	0.254	-0.4 – 0.1
Anterior thigh	2.9 (1.7)	3.3 (1.9)	0.4 (1.4)	3.2	0.090	-0.1 – 0.8
Distal medial leg[b]	4.5 (3.5)	4.7 (3.4)	0.2 (3.6)	7.4	0.647	-0.8 – 1.3
Foot dorsum	4.0 (3.0)	4.2 (3.0)	0.2 (2.1)	4.4	0.543	-0.4 – 0.8
Warm detection threshold, °C						
Thenar eminence	2.7 (1.5)	2.1 (0.9)	0.5 (1.2)	2.9	0.004	0.2 – 0.8
Anterior thigh	4.4 (2.0)	4.4 (2.0)	0.0 (2.3)	4.6	0.978	-0.7 – 0.7
Distal medial leg	9.4 (3.8)	10.1 (3.5)	-0.7 (2.7)	-6.1	0.083	-1.5 – 0.1
Foot dorsum	6.9 (4.0)	8.2 (4.1)	-1.4 (2.7)	-6.8	0.001[a]	-2.1 – -0.6
Heat pain threshold, °C						
Thenar eminence	12.2 (4.1)	12.5 (3.7)	-0.3 (2.0)	-4.3	0.324	-0.9 – 0.3
Anterior thigh	13.6 (3.3)	13.5 (3.1)	0.0 (2.0)	4.0	0.891	-0.5 – 0.6
Distal medial leg	14.7 (2.5)	14.8 (2.7)	0.0 (1.6)	3.2	0.873	-0.5 – 0.4
Foot dorsum	13.1 (3.6)	13.9 (3.3)	-0.8 (1.9)	-4.6	0.005	-1.4 – -0.3

Abbreviations: SD = Standard Deviation, CI = Confidence Interval
Sample normal limit = Mean ± 2 SD
[a]*Significant at p < 0.05 after Bonferroni adjustment for multiple testing*
[b]*n = 46*

TABLE 4 Distal-proximal gradients for thermal thresholds (Δ°C)

Pairwise comparisons		Mean difference (SD)	Sample normal limit	P-value	95% CI
Cold detection threshold[c], °C					
Foot dorsum	Distal medial leg	-0.6 (2.7)	-6.0	0.170	-0.5 – 1.7
	Anterior thigh	-0.9 (2.0)	-4.9	0.002[a]	-1.7 – -0.1
	Thenar eminence	-2.1 (2.5)	-7.1	<0.001[b]	-3.2 – -1.1
Distal medial leg	Anterior thigh	-1.5 (2.8)	-7.1	0.001[b]	-2.7 – -0.3
	Thenar eminence	-2.7 (3.0)	-8.7	<0.001[b]	-3.9 – -1.5
Anterior thigh	Thenar eminence	-1.2 (1.6)	-4.4	<0.001[b]	-1.9 – -0.5
Warm detection threshold, °C					
Foot dorsum	Distal medial leg	-2.2 (2.5)	-7.2	<0.001[b]	-3.2 – -1.2
	Anterior thigh	3.2 (3.2)	9.6	<0.001[b]	1.9 – 4.4
	Thenar eminence	5.2 (3.4)	12.0	<0.001[b]	3.8 – 6.5
Distal medial leg	Anterior thigh	5.4 (2.9)	11.2	<0.001[b]	4.2 – 6.5
	Thenar eminence	7.4 (3.3)	14.0	<0.001[b]	6.1 – 8.7
Anterior thigh	Thenar eminence	2.0 (2.0)	6.0	<0.001[b]	1.4 – 2.6
Heat pain threshold, °C					
Foot dorsum	Distal medial leg	-1.2 (1.5)	-4.2	<0.001[b]	-1.8 – -0.6
	Anterior thigh	0.0 (2.1)	±4.2	0.932	-0.9 – 0.8
	Thenar eminence	1.2 (3.4)	8	0.016	-0.1 – 2.5
Distal medial leg	Anterior thigh	1.2 (2.0)	5.2	<0.001[b]	0.4 – 2.0
	Thenar eminence	2.4 (3.3)	9.0	<0.001[b]	1.1 – 3.7
Anterior thigh	Thenar eminence	1.2 (2.2)	5.6	<0.001[b]	0.3 – 2.1

Abbreviations: SD = Standard Deviation, CI = Confidence Interval, Sample normal limit = Mean ± 2 SD
[a,b]Significant at p ≤ 0.05 and 0.01, respectively, after Bonferroni adjustments for multiple comparisons.
[c]n = 46

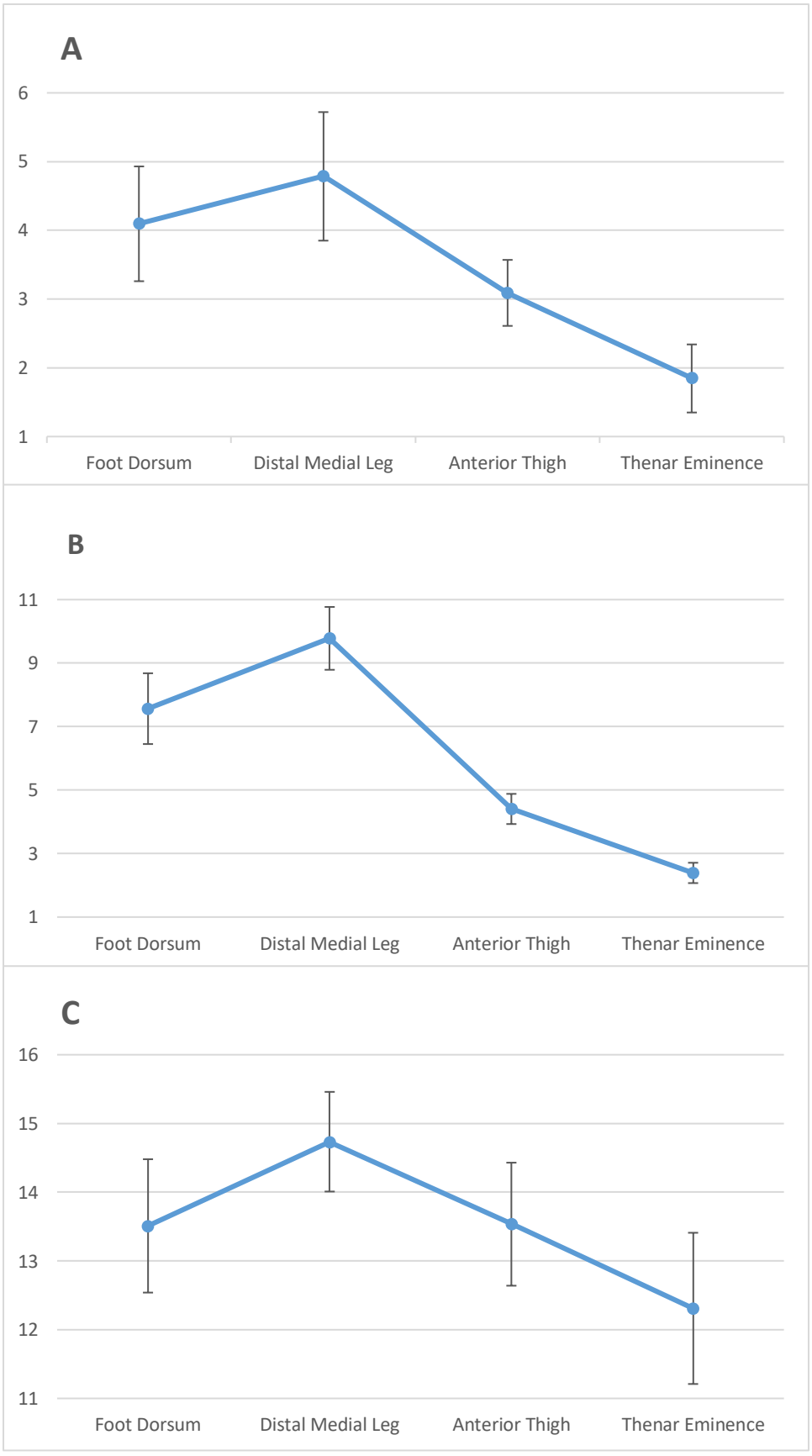

*Figure 2: Thermal thresholds (Δ°C) for (**A**) cold detection threshold, (**B**) warm detection threshold and (**C**) heat pain threshold for the foot dorsum, distal medial leg, anterior thigh and thenar eminence. Error bars represent 95% confidence intervals.*

APPENDIX A: Supporting material

SUPPORTING TABLE 1 Predicted normal limits for warm and cold detection thresholds in the foot dorsum and distal medial legs (Δ°C), and for the difference between foot dorsum and anterior thigh

					Normal limits	
Age (yr)	CDT Foot dorsum	WDT Foot dorsum	CDT Distal medial leg	WDT Distal medial leg	CDT Foot-Anterior thigh	WDT Foot-Anterior thigh
20	5.05	8.38	5.36	11.95	-2.91	4.58
21	5.18	8.56	5.45	12.08	-2.99	4.72
22	5.31	8.75	5.53	12.22	-3.07	4.86
23	5.44	8.94	5.62	12.35	-3.15	5
24	5.57	9.13	5.70	12.48	-3.23	5.14
25	5.70	9.33	5.78	12.61	-3.31	5.28
26	5.84	9.52	5.87	12.75	-3.39	5.42
27	5.97	9.72	5.95	12.88	-3.47	5.56
28	6.11	9.92	6.04	13.01	-3.55	5.7
29	6.25	10.13	6.12	13.15	-3.63	5.84
30	6.39	10.33	6.20	13.28	-3.71	5.98
31	6.53	10.54	6.29	13.41	-3.79	6.12
32	6.68	10.75	6.37	13.55	-3.87	6.26
33	6.82	10.96	6.46	13.68	-3.95	6.4
34	6.97	11.17	6.54	13.81	-4.03	6.54
35	7.12	11.38	6.62	13.94	-4.11	6.68
36	7.27	11.60	6.71	14.08	-4.19	6.82
37	7.42	11.82	6.79	14.21	-4.27	6.96
38	7.57	12.04	6.88	14.34	-4.35	7.1
39	7.73	12.26	6.96	14.48	-4.43	7.24
40	7.88	12.49	7.04	14.61	-4.51	7.38
41	8.04	12.72	7.13	14.74	-4.59	7.52
42	8.20	12.95	7.21	14.88	-4.67	7.66
43	8.36	13.18	7.30	15.01	-4.75	7.8
44	8.53	13.41	7.38	15.14	-4.83	7.94
45	8.69	13.65	7.46	15.27	-4.91	8.08
46	8.86	13.88	7.55	15.41	-4.99	8.22
47	9.02	14.12	7.63	15.54	-5.07	8.36
48	9.19	14.36	7.72	15.67	-5.15	8.5
49	9.36	14.61	7.80	15.81	-5.23	8.64
50	9.54	14.85	7.88	15.94	-5.31	8.78
51	9.71	15.10	7.97	16.07	-5.39	8.92
52	9.88	15.35	8.05	16.21	-5.47	9.06
53	10.06	15.60	8.14	16.34	-5.55	9.2
54	10.24	15.86	8.22	16.47	-5.63	9.34
55	10.42	16.11	8.30	16.60	-5.71	9.48
56	10.60	16.37	8.39	16.74	-5.79	9.62
57	10.78	16.63	8.47	16.87	-5.87	9.76
58	10.97	16.89	8.56	17.00	-5.95	9.9
59	11.16	17.16	8.64	17.14	-6.03	10.04
60	11.34	17.42	8.72	17.27	-6.11	10.18
61	11.53	17.69	8.81	17.40	-6.19	10.32
62	11.72	17.96	8.89	17.54	-6.27	10.46
63	11.92	18.23	8.98	17.67	-6.35	10.6
64	12.11	18.51	9.06	17.80	-6.43	10.74
65	12.31	18.78	9.14	17.93	-6.51	10.88
66	12.50	19.06	9.23	18.07	-6.59	11.02
67	12.70	19.34	9.31	18.20	-6.67	11.16
68	12.90	19.62	9.40	18.33	-6.75	11.3
69	13.10	19.91	9.48	18.47	-6.83	11.44
70	13.31	20.00[a]	9.56	18.60	-6.91	11.58
71	13.51	20.00[a]	9.65	18.73	-6.99	11.72
72	13.72	20.00[a]	9.73	18.87	-7.07	11.86
73	13.93	20.00[a]	9.82	19.00	-7.15	12
74	14.14	20.00[a]	9.90	19.13	-7.23	12.14
75	14.35	20.00[a]	9.98	19.26	-7.31	12.28
76	14.56	20.00[a]	10.07	19.40	-7.39	12.42
77	14.78	20.00[a]	10.15	19.53	-7.47	12.56
78	14.99	20.00[a]	10.24	19.66	-7.55	12.7
79	15.21	20.00[a]	10.32	19.80	-7.63	12.84

Abbreviations: CDT = Cold Detection Threshold; WDT = Warm Detection Threshold, Diff = Difference

Values are calculated from the 95% prediction intervals of the linear regression equations for mean Δ°C as a function of age (yrs)

[a] *Warm detection threshold cut-off at 52 °C*

8.2 Appendix B: Author guidelines for original articles in the European Journal of Pain

Author guidelines relevant to the paper in Appendix A are cut and pasted from https://onlinelibrary.wiley.com/page/journal/15322149/homepage/forauthors.html.

2. ARTICLE TYPES AND CONTENT

EJP invites the following types of submission:

Original Articles

Original Articles are the journal's primary mode of communication.

Original articles must include a structured abstract including at the end a statement "Significance", indicating the main aspects where this work adds significantly to existing knowledge in the field, and if appropriate to clinical practice. The significance statement should be short, attention-grabbing, non-redundant with the conclusions and rigorously in line with the contents of the full article ' (see Section 4).

4. PREPARATION OF MANUSCRIPTS

Manuscripts must be written in English.

Manuscript text must be saved in Word (.doc) or Rich Text Format (.rtf). Please do not submit text in PDF format (.pdf).

Due to space restrictions and a better readability papers generally should not exceed ten typeset pages (780 words/page, 32 references/page, including figures and tables). *EJP* can publish additional material as "supporting material" with a special link guiding from the manuscript to this material.

Suggestions for the cover inset are invited. The illustration may be from a manuscript accepted for publication in the European Journal of Pain.

Manuscript Structure and Word Count

1) Manuscript
   - Title page (see further details below)
   - Abstract (should not exceed 250 words, see further details below)
   - Text
     - o Introduction (no subheadings, should not exceed 500 words)
     - o Methods (or Literature Search Methods for Review Articles)
     - o Results
     - o Discussion and conclusions (should not exceed 1500 words)
   - Acknowledgements
   - Author contributions (see Section 6)
   - References (limited to 80 for original manuscripts)
   - Legends for illustrations and tables

2) Tables (to be uploaded as separate files)
3) Figures (to be uploaded as separate files)
4) Supporting material (additional material that will be published online-only, to be uploaded separately, see further details below)

Title Page

The title page should give:

1) The title of the article. Titles should be short and should not contain acronyms
2) A running head not exceeding 50 characters
3) The authors' names (initial(s) of first name(s) and last name of each author)
4) The names of the institutions at which the research was conducted, clearly linked to respective authors
5) The name, address, telephone and fax numbers, and e-mail address of the author responsible for correspondence
6) The category for which the manuscript is being submitted (original article, review, short communication)
7) A statement of all funding sources that supported the work
8) Any conflicts of interest disclosures (see Section 6).
9) A statement “Significance”, indicating the main aspects where this work adds significantly to existing knowledge in the field, and if appropriate to clinical practice. The signifiance statement should be short, attention-grabbing, non-redundant with the conclusions and rigorously in line with the contents of the full article. It should not exceed 80 words and will be added to the end of the abstract at the time of typesetting. It does not count to thte abstract's word limit (250 words).The statement "Significance" also applies to Review papers. It has to be given on the title page and will be added at the end of the abstract at the time of typesetting. It does not count to the abstract's word limit (250 words).

Abstract

The abstract should not exceed 250 words and should describe the background, the aims, the methods, the results and the conclusions reached. It should contain only standard abbreviations and no references. For *Original Manuscripts* the following subheadings are required:

- Background
- Methods
- Results
- Conclusions

Acknowledgements

The acknowledgements section should specify acknowledgement of technical help, but no sources of financial and material support. These should be given in the "Funding Sources" on the Title page.

Author Contributions

Authors are required to include a statement of responsibility at the end of their manuscript's text that specifies the contribution of every author (see Section 6). Please state that all authors discussed the results and commented on the manuscript.

References

In the text: references should be cited in parantheses at the appropriate point in the text by author(s) and year in chronological order, e.g. for one author: (Mustola, 1996), for two authors: (Mustola and Baer 1998), for more than two authors: (Mustola et al., 1999). If two or more references with the same first author and year are cited, use lower-case letters a, b, etc., after the year both in the text and in the reference list. The complete reference list should appear alphabetically by the name at the end of the paper.

Figures

All figures must be uploaded as separate files. Figure legends should be listed **on a separate page** in numerical order and should contain brief but comprehensible explanations.

Figures should be referred to in the text in numerical sequence as follows: Fig. 1, Figs 2–4. The place at which a figure is to be inserted in the printed text should be indicated clearly on a manuscript. Where a figure has more than one panel, each panel should be labeled in the top left-hand corner using lower case letters in parentheses i.e. ‘(a)’, ‘(b)’ etc., and a brief description of each panel given in the figure legend.

Print publication requires high quality, EPS (lineart) or TIFF/PDF (halftone/photographs) files are preferable (though GIF, JPEG, PICT or Bitmap files are acceptable for submission). MS PowerPoint and Word Graphics are **unsuitable** for printed pictures. Scans (TIFF only) should have a resolution of 300 dpi (halftone) or 600 to 1200 dpi (line drawings) in relation to the reproduction size (see below). EPS files should be saved with fonts embedded (and with a TIFF preview if possible). For scanned images, the scanning resolution (at final image size) should be as follows to ensure good reproduction: lineart: >600 dpi; half-tones: >300 dpi; figures containing both halftone and line images: >600 dpi.

Tables

Tables should be referred to in the text in numerical sequence as follows: Table 1, Table 2. Each table, with an appropriate brief legend, comprehensible without reference to the text, should be typed on a separate page. For footnotes, use superscripts 'a', 'b', 'c', etc., not asterisks or other symbols.

Supporting Information

EJP encourages the submission of underlying data sets, appendices, additional figures or tables, movie files, animations, etc. as online supporting information. Supporting information should be uploaded during manuscript submission (see Section 3). Supporting information should be important ancillary information that is relevant to the parent article but which is not included in the typeset PDF but can be accessed via a link from this PDF.

To submit any material to be published as supporting material please choose the item "supporting information" when uploading the files of tables or figures.

Please use the following terms:
- for tables: "tableS1", "tableS2" etc.
- for figures: "figureS1, "figureS2" etc. The figure legends should be included in the figure's file.
- for parts of the manuscript's text: "methodsS1", resultsS1", "discussionS1" or "AppendixS1" (please note that it is not possible to publish additional material for the introduction)
- please indicate and cite clearly in your manuscript the supporting information using the terms given above.

Units & Abbreviations

Measurements of length, height and volume should be reported in metric units (metre, kilogram, litre). Temperatures should be given in degrees Celsius and blood pressures in millimetres of mercury or kPa with the alternative units in parentheses. All other measurements including laboratory measurements should be reported in the metric system in terms of the International System of Units (SI).

Abbreviations should be limited and defined after the first use of the term.

5. IMPORTANT DECLARATIONS

Original Publication

Submission of a manuscript will be held to imply that it contains original unpublished work and is not being submitted for publication elsewhere at the same time. The author must supply a full statement to the Editor-in-Chief about all submissions and previous reports that might be regarded as redundant or duplicate publication of the same or very similar work.

Conflicts of Interest

Authors are responsible for disclosing all financial and personal relationships between themselves and others that might be perceived by others as biasing their work. To prevent ambiguity, authors must state explicitly whether potential conflicts do or do not exist.

Ethics

When reporting experiments on human subjects, indicate whether the procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional or regional) and with the Helsinki Declaration of 1975, as revised in 1983. Do not use patients' names, initials or hospital numbers, especially in illustrative material.

Authorship

All persons designated as authors should qualify for authorship and all those who qualify should be listed. Each author should have participated sufficiently in the work to take public responsibility for appropriate portions of the content. One or more authors should take responsibility for the integrity of the work as a whole, from inception to published article. Authorship credit should be based only on 1) substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data; 2) drafting the article or revising it critically for important intellectual content; 3) final approval of the version to be published. Conditions 1, 2 and 3 must all be met. Acquisition of funding, the collection of data or general supervision of the research group, by themselves, do not justify authorship. All others who contributed to the work who are not authors should be named in the Acknowledgements section.