University of Cape Town

Department of Psychiatry and Mental Health

Division of Neuropsychiatry

Insight into the modulation and mechanisms of experimentally induced secondary hyperalgesia

Luyanduthando Mqadi

MQDLUY001

Submitted to the University of Cape Town

In fulfilment of the requirements for the degree

MSC Med Neuroscience degree

Primary supervisor: Associate Victoria J Madden (PhD)

Co-supervisors: Associate Professor Romy Parker (PhD)

Professor John Joska (Prof)

Ms Gill Bedwell (PhD candidate)

13 February 2023

PLAGIARISM DECLARATION

I, Luyanduthando Mqadi, hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature: Signed by candidate

Date: 13/02/2023

ACKNOWLEDGEMENTS

I would like to express my thanks and appreciation for the following individuals.

Supervisors

I would like to express my boundless gratitude to my supervisors; Tory, Romy, Gill and John. Your constant support, advice and communication have helped me overcome many challenges, making it possible for me to achieve several milestones throughout my Master's training. Thank you for your patience and for investing in my personal and professional growth. I would have not made it this far without your immense knowledge, plentiful experience, mentorship and support.

Ukomelela research team, Pain Team (Department of Anaesthesia and Perioperative Medicine) and Neuropsychiatry journal club (Department of Psychiatry and Mental Health)

Ukomelela research team, thank you so much for making this research possible. Ncumisa Msolo and Andiswa Dikana, this research would have not been possible without your constant patience, support, and hard work during data collection. In addition, I want to thank Gwendoline Arendse for helping me with data analysis. To the members of the pain team and neuropsychiatry journal club, thank you for your support and contribution to this research.

Participants

Thank you to all the participants who contributed their time to this research project. I could not have undertaken this research project without your participation.

Funders

I want to thank the Vice-Chancellor's Scholarships, National Research Foundation, Ada & Bertie Levenstein Bursary and Postgraduate Funding Office at the University of Cape Town for providing me with financial support during my studies. Thank you for believing in me and investing in my education.

Family and friends

I want to thank my family, parents, and friends for their encouragement and support.

To my father, Derrick Mqadi, thank you for the emotional, moral and financial support you have given me. Thank you for always believing in me and encouraging me to shoot for the stars – I love and appreciate you, dad.

To my friends, Nthabiseng Phokoane, Nongcebo Mahlalela, and Sipho Hokwana thank you for supporting me throughout this journey. Further, I want to thank Lindokuhle Nkambule for helping me with data analysis.

ACRONYMS

The list of acronyms in this dissertation aims to provide a comprehensive reference tool that facilitates the understanding and accessibility of the content. The following list comprises acronyms that fulfil the criteria of being commonly used in literature, frequently used throughout this dissertation, or representing terms that require clarification:

SH	Secondary hyperalgesia
CS	Central sensitisation
HFS	High frequency electrical stimulation
CNS	Central nervous system
NMDA	N-methyl-D-aspartate
TRPV1	Transient receptor potential vanilloid-1
AMPA	Alpha-Amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid
HIV	Human Immunodeficiency Virus
PWH	People with HIV
VAS	Visual Analogue Scale
BPI	Brief Pain Inventory
HSCL-25	Hopkins Symptom Checklist-25
PSD	Psychosocial distress: represents symptoms of depression and anxiety.
GRADE	Grades of Recommendation, Assessment, Development and Evaluation

ABSTRACT

Background

Persistent pain is a frequent complaint associated with compromised mental health in many people with HIV (PWH). Central sensitisation (CS) is a hallmark of most persistent pain conditions. Pain and psychosocial distress (PSD) appear to have a bidirectional relationship. Therefore, it is plausible that PSD contributes to pain by increasing central sensitisation (CS), thus maintaining pain in PWH; however, this hypothesis has not been tested. Experimentally induced secondary hyperalgesia (SH) has been used in the laboratory study of other pain conditions (e.g., neuropathic pain) to increase understanding of CS mechanisms. Therefore, the model of SH is useful for investigating pain related

to CS, such as pain associated with HIV.
NMDA antagonists are often used to prevent and treat pain associated with CS based on the evidence from animal and human pain studies. Evidence suggests that NMDA antagonists likely target mechanisms underlying clinical pain, specifically wind-up, allodynia and hyperalgesia. However, the effect of NMDA antagonists on SH is under-recognised and poorly understood. Understanding the effect of NMDA antagonists on SH would help to inform the appropriate matching of treatment to individuals. We conducted two studies. First, a systematic review according to PRISMA guidelines on the existing evidence that targeting the NMDA system alters experimentally induced SH, in healthy human participants. Second, we conducted an experimental study to investigate whether distress predicted central sensitisation in people living with HIV (PWH).

Methods

Systematic review: The influence of NMDA targeting pharmacological manipulations on experimentally induced SH in healthy human adults without clinical pain

We identified studies that recruited healthy, pain-free human participants and experimentally induced SH (magnitude of SH, surface area of SH) and used a pharmacological method known to target the NMDA system to manipulate SH. Studies were identified by searching various electronic databases (conducted on the 24th of June 2019, updated on the 29th of September 2022). We also checked reference lists and contacted experts in the field, including authors who have recently published narrative reviews on experimental induction and manipulation of SH, to identify studies missed through electronic searching. We included studies that were published and in-press or accepted records with titles, abstracts, and full-text versions available in English. Authors were asked to provide missing data where necessary. Two or more reviewers assessed the risk of bias, extracted data, and judged the quality of evidence (GRADE) of the included studies. Data analysis was performed by narrative summary, and if more than two studies were available for a given

manipulation, data were included in the meta-analysis. Data were pooled in subgroups by study design and the type of manipulation (e.g., ketamine) for each outcome (magnitude of secondary hyperalgesia, area of secondary hyperalgesia, or both). We generated funnel plots to examine publication bias.

Experimental study: Does distress predict central sensitisation in PWH?

We aimed to investigate the relationships between PSD, SH (a known human surrogate model of CS), and persistent pain in PWH. We recruited consenting adults with well controlled HIV, reporting either persistent pain or no pain (assessed using a modified Brief Pain Inventory). Participants provided self-reports of PSD severity (on the Hopkins 25-item scale). We used high-frequency electrical stimulation to induce SH on one pain-free forearm and assessed the surface area (primary outcome) and magnitude (secondary outcome) of SH using a von Frey filament and 'pinprick' rods, respectively. It was hypothesised that the surface area and magnitude of experimentally induced SH would be positively associated with PSD severity (hypothesis 1) and that the persistent pain group would have a greater surface area and magnitude of experimentally induced SH than the group without pain (hypothesis 2).

Results

Systematic review: The influence of NMDA targeting pharmacological manipulations on experimentally induced SH in healthy human adults without clinical pain

Twenty-nine records were included in this review. Some records reported multiple manipulation methods, so each yielded multiple datasets. Therefore, the 29 records yielded 52 datasets. The effect of NMDA antagonists on experimentally induced SH was assessed by change in magnitude of SH in three records (4 of 52 studies) and assessed by surface area of SH in 20 records (37 of 52 studies). Six records (11 of 52 studies) assessed by change in surface area of SH. Six records (11 of 52 studies) assessed both the change in magnitude and surface area of SH in response to NMDA antagonist administration.

Narrative summary

Twelve studies manipulated the magnitude of SH using ketamine alone (n = 10) or combined with alfentanil (n = 2). Seven reported no effect, and three found a decrease in the magnitude of SH. Combining ketamine and alfentanil had no effect on the magnitude of SH.

Thirty-five studies manipulated the surface area of SH using ketamine alone (n = 31) or combined with alfentanil (n =2), morphine (n = 1) or remifentanil (n = 1). Of those using ketamine alone, fifteen reported no effect, and sixteen found a decrease in the surface area of SH. Combining ketamine with alfentanil or morphine had no effect on the surface area of SH; however, the combination of ketamine and remifentanil decreased the surface area of SH. Ten studies manipulated the surface

area of SH using dextromethorphan alone (n = 8) or combined with morphine (n = 2). Four of those using dextromethorphan alone reported no effect, and the rest reported a decrease in the surface area of SH. Those who combined dextromethorphan with morphine reported no effect (n=1) or a decrease (n=1) in the surface area of SH. Studies that used CH3381 (n = 1) or neramexane (n = 1) reported a decrease in the surface area of SH. One study that used magnesium sulphate found no effect of magnesium on the surface area of SH.

Meta-analysis

Forty studies (seven assessing the magnitude of SH) were included in the meta-analysis. Ketamine had no effect on the magnitude of SH. Similarly, ketamine, CHF3381, and dextromethorphan had no effect on the surface area of SH.

Experimental study: Does distress predict central sensitisation in PWH?

There was a positive relationship between PSD severity and the surface area of SH; however, the surface area was not predicted by pain status. There was no relationship between PSD severity and the magnitude of SH. Our plots suggested that the 'pain' group had a stronger positive relationship between PSD and the magnitude of SH than the 'no pain' group. However, we couldn't confirm this because we couldn't find a suitable statistical model. The proportion of participants reporting pain developed a significantly greater magnitude of SH than those without pain. In contrast, the proportion of participants without pain developed significantly greater surface area of SH than those with pain.

Conclusion

Systematic review: The influence of NMDA targeting pharmacological manipulations on experimentally induced SH in healthy human adults without clinical pain

The results of this study indicate that NMDA antagonists have no effect on the surface area or magnitude of experimentally induced SH. These findings carry important implications for clinical practice, highlighting the limited efficacy of NMDA antagonists in modulating SH. Future studies should explore alternative pharmacological and non-pharmacological interventions to assess their potential in modulating SH and optimising patient outcomes.

Experimental study: Does distress predict central sensitisation in PWH?

Psychosocial distress predicted the surface area but not the magnitude of SH in individuals living with HIV, independent of pain status. The positive relationship between the surface area of SH and PSD may support targeting of PSD to reduce pain. Specifically, interventions aimed at reducing PSD may hold therapeutic potential in addressing SH. Future studies should explore the effects of distress-reducing therapy on clinical SH.

Table of contents

TABLE OF TABLES

TABLE OF FIGURES

CHAPTER 1: INTRODUCTION

Pain is a global problem

Persistent pain is a global issue that transcends geographic and socioeconomic boundaries. Globally, persistent pain is reported by approximately 27.5% of adults, with prevalence rates varying across countries ranging from 9.9% to 50.3% (Zimmer et al., 2022). In South Africa, the prevalence of persistent pain falls within the global range: approximately 18% of adults experience persistent pain, indicating that nearly 1 in 5 individuals is affected (Kamerman et al., 2020). Inadequate pain management compromises the physical and mental well-being of individuals (Henwood & Ellis, 2004). Multiple factors and complex pathways are involved in the experience of pain, including sensory, neurological, and psychological components (Woolf et al., 1998). It is plausible that understanding the underlying mechanisms of pain allows for targeted interventions that act specifically on those mechanisms. Therefore, exploring and uncovering the underlying mechanisms of pain has the potential to help alleviate pain on both global and national scales.

Classifications of pain

Pain is a complex phenomenon influenced by various physiological and psychological mechanisms. Pain is described as, an unpleasant sensory and emotional experience associated with, or resembling that associated with, actual or potential tissue damage. (Raja et al., 2020). Pain can be classified according to its underlying mechanism: nociceptive, neuropathic or nociplastic. Nociceptive pain occurs after the activation of nociceptors as a response to actual or threatened damage to non-neural tissue in a normally functioning somatosensory nervous system (IASP, 2021h). In contrast, neuropathic pain occurs due to a lesion or disease in the central nervous system in an abnormally functioning somatosensory nervous system (IASP, 2021i). Nociplastic pain occurs when the activation of nociceptors is altered in the absence of actual or threatened tissue damage or evidence of a lesion or disease of the somatosensory system (IASP, 2021f). These classifications of pain identify variations in the underlying mechanisms contributing to each type. This thesis investigates mechanisms associated with certain pain features, specifically those underpinned by ongoing central sensitisation (CS).

Central sensitisation contributes to persistent pain

Persistent pain is associated with central sensitisation (CS) (Merlin, 2015; Petersen et al., 2015). Central sensitisation is the increased sensitivity and response of neurons in the central nervous system (CNS) to a nociceptive stimulus, occurring due to altered sensory processing in the CNS (IASP, 2021g). Long-term potentiation plays an important role in maintaining a sensitised CNS by increasing synaptic transmission in the CNS (Liu & Zhou, 2015; Nijs et al., 2011; Woolf, 2014). Long-term potentiation maintains a sensitised CNS by reducing pain thresholds and nociception. In a sensitised

CNS, the quantity of nociceptive input reaching the brain increases, and neural activity 'spreads' to adjacent neural pathways (i.e., uninjured site), which results in a broad distribution of neural activity. This broadened distribution of neural activity supports clinical manifestations, such as pain 'spreading' beyond the original site of injury (Arendt-Nielsen et al., 2018), as observed in individuals with secondary hyperalgesia (SH).

Secondary hyperalgesia is a measure of central sensitisation

Secondary hyperalgesia is a pain feature that provides valuable insight into CS. SH is heightened skin sensitivity to a usually painful stimulus in the area surrounding the injured/stimulated skin (IASP, 2021e; Madden et al., 2019; van den Broeke & Mouraux, 2014b). When CS occurs, there is an increased response of neurons in the CNS to a nociceptive stimulus. SH serves as a measure of CS because it is evidence of the 'spread' of pain sensitivity to non-injured skin areas. Clinically, SH typically extends beyond the visibly injured skin area, such as increased pain sensitivity surrounding a surgical incision site (Pogatzki-Zahn, Segelcke & Schug, 2017). Notably, SH can also occur in the absence of visible tissue injury, as observed in individuals living with painful neuropathies (Campbell & Meyer, 2006). While the precise mechanism of SH remains unclear, it is thought to involve an increase in synaptic efficiency within the synapses of the dorsal horn (Treede & Magerl, 2000), and may be triggered by repetitive afferent signalling. Various experimental pain models have been used to assess SH and investigate the contribution of CS to pain (Quesada et al., 2021).

Induction of SH using experimental pain models

Experimental pain models are valuable tools for investigating mechanisms of SH. These models use a variety of safe techniques to induce SH experimentally. Examples include low and high frequency electrical stimulation (Van Den Broeke et al., 2014a), capsaicin (Andersen et al., 1996; Ditre et al., 2018), burn (Brennum, Kaiser & Dahl, 2001; Rasmussen et al., 2015) or freeze injuries (Chassaing et al., 2006; Martin et al., 2019). Table 1 summarises the various pain models that have been developed to induce SH. Experimental SH is similar to its clinical counterpart; however, there is a notable difference between clinical and experimental SH. While clinical SH can persist for an extended period, experimental SH is short lived, typically lasting only a few hours (Pfau et al., 2011; Vo & Drummond, 2013). Experimentally inducing SH provides a time-efficient approach to studying SH within a shorter timeframe. This time-efficient approach facilitates efficient investigations into the mechanisms and characteristics of SH. Given the complex nature of pain, treatments targeting mechanisms are often preferred over symptom-targeting treatments (Baumgärtner et al., 2002; Jensen & Baron, 2003; Woolf et al., 1998). Consequently, researchers employ pain models to manipulate SH, providing insights into the underlying mechanisms and aiding the development and evaluation of treatments for pain

Table 1: Summary of pain models used to induce SH.

Pain model(s)	Mechanism of action
Electrical stimulation Low frequency electrical stimulation (LFS) High frequency electrical stimulation (HFS)	→ Direct (axonal) activation of free nerve endings in only the most superficial layers of the skin generates only nociceptive input. The very → frequent afferent signalling results in temporary strengthening of synapses, and 'spill-over' of neural activity to adjacent neural pathways (Meyer et al., 2005).
Capsaicin Intradermal injection Topical application Capsaicin combined with heat **Burn injuries**	→ Capsaicin is a chemical compound found in chilli peppers, which induces intense burning sensations (Fattori et al., 2016), a common feature in pain underpinned by CS. → Capsaicin and/or thermal stimuli activates transient receptor potential vanilloid-1 (TRPV1) ion channel receptors (Willis, 2009). → In the periphery, TRPV1 receptors are prevalently expressed in endings of small diameter nociceptive primary afferent neurons (Meza et al., 2022). These receptors can be activated by heat or chemical compounds (e.g., capsaicin) (Caterina et al., 1997). In animal studies, the activation of peripheral TPRV1 receptors is associated with nociception and pain sensation (assessed through pain behaviour) (Caterina et al., 2000), including thermal hyperalgesia (Davis et al., 2000). → Centrally, TRPV1 receptors are prevalently expressed on endings of nociceptive neurons in the spinal cord that make synaptic contact with second order neurons (Caterina et al., 1997; Hwang et al., 2004), including a small portion present in C and A-delta fibres, which may contain substance P. → Activation of TRPV1 receptors may trigger the release of substance P resulting in the generation of action potentials that are sent to the brain where they may be perceived as pain (Jara-Oseguera, Simon & Rosenbaum, 2008). → Activation of TRPV1 receptors results in the strengthening of synapses and the 'spreading' of neural activity to adjacent neural pathways, which results in the development of SH. The pain induced by capsaicin is typically described as a burning sensation (Iftinca, Defaye & Altier, 2021). Burning sensation is frequently reported in neuropathic pain conditions (Campbell & Meyer, 2006), which establishes a connection between the experimental capsaicin-induced pain model and the clinical descriptions provided by patients. The characterisation of the pain induced by capsaicin as a burning sensation provides a relatable reference to clinical pain and reflects the importance of this model in understanding pain mechanisms.
Freeze injury	→ Cold stimuli activate cold receptors from the TRP family. → TRPM8 receptors are activated at 25°C and TRPA1 receptors are activated at 17°C. → TRPM8 and TRPA1 receptors are expressed in the dorsal root ganglion (Ji et al., 2007). → Cold sensations activate A-delta and C-fibres. → Results in cold hypersensitivity, a common feature in neuropathic pain (Ji et al., 2007).

Involvement of the NMDA system in the mechanism of SH

The NMDA system is generally assumed to play a significant role in the development and maintenance of SH. Activation of NMDA receptors in the CNS contributes to the amplification of pain, the spread of hyperalgesia beyond the site of injury, and facilitates synaptic plasticity to enhance the excitability of nociceptors, thereby contributing to the sensitisation of the CNS (Angst et al., 2003;

Petrenko et al., 2003; Zhou, Chen & Pan, 2011). NMDA antagonists have been used to block NMDA receptors or modulate their activity in an attempt to alleviate pain (Carpenter & Dickenson, 1999). For example, the NMDA antagonist ketamine has demonstrated inhibitory effects on SH in patients with complex regional pain syndrome (Finch, Knudsen & Drummond, 2009). However, its impact appears to vary depending on the underlying pain condition. Ketamine was found to have no effect on SH in individuals with neuropathic pain (Lynch et al., 2005). Furthermore, while ketamine decreased capsaicin-induced SH in some studies with individuals without pain (Andersen et al., 1996; Gottrup et al., 2000b), others reported no significant influence of ketamine on capsaicin-induced SH (Gottrup et al., 2000a; Sethna et al., 1998). These divergent findings suggest that the influence of the NMDA system on SH may be contingent upon various factors, such as the type of pain condition being studied. Therefore, more research is required to elucidate the precise mechanisms underlying SH. In Chapter 2, we conducted a systematic review and meta-analysis of the existing evidence that targeting the NMDA system alters experimentally induced SH, in healthy human participants.

Hypothesised involvement of psychosocial distress in the mechanism of SH

Mental health and persistent pain share a bidirectional relationship that has a profound impact on individuals' overall well-being. Individuals living with persistent pain commonly experience psychosocial distress (PSD), including symptoms of depression and anxiety. The constant physical discomfort, limitations in daily activities, and challenges in maintaining relationships and fulfilling responsibilities associated with living with pain contribute to increased distress (Castillo et al., 2013; Hadlandsmyth et al., 2017; Sturgeon et al., 2015). For example, the experience of pain can trigger negative emotions such as frustration, anger, sadness, and fear, which can contribute to increased levels of PSD (Burns et al., 2015; Okifuji, Turk & Curran, 1999; van Middendorp et al., 2010). Similarly, PSD can exert a significant influence on pain perception and experience. For example, written emotional disclosure has been found to decrease the surface area of capsaicin-induced SH in women with a history of trauma (You et al., 2014). Stressful life events that translated to distress increased the surface area of capsaicin-induced SH (You, Creech & Meagher, 2016). Similarly, negative expectations increased SH following high frequency electrical stimulation (van den Broeke et al., 2014a). This evidence suggests that PSD-related factors can modulate pain. The precise mechanisms by which PSD influences pain are not yet fully understood. However, PSD may contribute to SH, thereby supporting the maintenance and persistence of pain. We acknowledge the limited understanding of the specific pathways involved in the impact of PSD on pain. However, SH serves as an example of how PSD can extend pain sensitivity beyond the injured skin areas, suggesting that psychosocial factors can contribute to the maintenance and persistence of pain. Therefore, in Chapter 3, we investigated the potential role of PSD as an upstream driver of SH in PWH.

CHAPTER 2: The influence of NMDA-targeting pharmacological manipulations on experimentally induced SH in healthy human adults without clinical pain

Introduction

The NMDA system is a well-described target of several medications used to treat pain. The efficacy of NMDA-targeted treatments has been demonstrated in neuropathic pain (Carlsson et al., 2004; Felsby et al., 1996; Galer et al., 2000; Pud et al., 1998). However, there is limited evidence clarifying the mechanism of action of these treatments. Secondary hyperalgesia is a common clinical feature observed in neuropathic pain. Central sensitisation is a process assumed to underlie SH, which plays a role in the persistence and maintenance of pain. We can investigate the effects and mechanisms of NMDA antagonists on CS using experimental pain models (Klein et al., 2005).

The NMDA system is generally assumed to play a critical role in the development and maintenance of CS (Hansen et al., 2018). NMDA receptors are one of three types of ionotropic glutamate receptors, the other two being AMPA and kainate receptors (Mori & Mishina, 1995). Activation of NMDA receptors in the CNS facilitates synaptic plasticity and increases the responsiveness of nociceptors (Latremoliere & Woolf, 2009). During persistent pain states, the repetitive activation of NMDA receptors increases the influx of calcium ions. The influx of calcium ions activates the calcium/calmodulin-dependent protein kinase II pathway, which strengthens synaptic connections between nociceptors (Ji & Woolf, 2001; Latremoliere & Woolf, 2009). This process ultimately results in an increased sensitivity to noxious input (Latremoliere & Woolf, 2009). In addition, activation of NMDA receptors also contributes to releasing excitatory neurotransmitters, such as glutamate. Glutamate activates AMPA channels and contributes to the transmission and amplification of nociceptive signals in the CNS (Fundytus, 2001; Zhuo, 2017). The NMDA system's contribution to CS highlights its critical role in developing and maintaining persistent pain states.

Clarifying the effect of NMDA antagonists on the underlying processes of CS can provide valuable mechanistic insight. Above, we explained how CS is a key process contributing to SH and the persistence of pain. NMDA antagonists have been identified as a potential treatment for modulating CS and alleviating pain (Carpenter & Dickenson, 1999). Understanding how NMDA antagonists influence the processes underlying CS can shed light on their effectiveness in targeting particular mechanisms of CS. Here, we draw on studies that used NMDA antagonists with an experimentally controlled model of SH to clarify the effect of NMDA antagonists on the processes that underlie SH (which are also assumed to underlie CS). Understanding the effects of these antagonists will bridge the gap between preclinical research and clinical applications, to fine-tune our understanding of the potential benefits and limitations of NMDA receptor-targeted treatments in pain management. Therefore, this review will critically examine the existing literature on the influence of NMDA

targeting pharmacological manipulations on experimentally induced SH in healthy human adults, exploring the current knowledge, methodological considerations, and implications for future research and clinical applications.

Objectives

The objectives of this systematic review were to:

1. Identify studies that recruited healthy, pain-free human participants and assessed the magnitude and/or surface area of experimentally induced SH;
2. Describe the methods used to induce SH, including the type of induction, equipment required, pain reported by participants before, after and/or during the induction, and the rate and nature of associated adverse events;
3. Describe the NMDA-targeting pharmacological methods (and sham controls) used to manipulate experimentally induced SH, including the agent, dose, mode of administration, and the rate and nature of associated adverse events;
4. Determine the efficacy of each manipulation method based on the effect of the manipulation on the magnitude (*primary outcome*) and/or surface area (*secondary outcome*) of experimentally induced SH.

Methods

This review draws records from a sub-sample of studies deemed eligible for a more extensive review that systematically identifies, collates, and appraises pharmacological and non-pharmacological methods used to manipulate experimentally induced SH in healthy humans, the protocol for which has been published (Madden et al., 2019). Here, we are interested in the subgroup of studies that tested pharmacological methods theorised to target the NMDA system.

Eligibility criteria

This review included studies that used *pharmacological methods* that target the *NMDA system* to manipulate experimentally induced SH to mechanical punctate stimuli in healthy humans. The complete set of eligibility criteria for the current review are (i) published and in-press or accepted records with title; abstract; and full-text versions available in English, (ii) studies that used an experimental procedure to induce and manipulate SH in healthy participants (i.e. clinical SH excluded), (iii) studies that used a pharmacological method of manipulation thought to target the NMDA system to manipulate SH, iv) studies that assessed SH within 120 minutes after the induction of SH, and assessed SH before, after or during the manipulation, and used participant self-report as the outcome measure.

Types of study participants

Regardless of age limitations, data from healthy, pain-free human participants were deemed eligible. It was planned that, where applicable, data from adults (age > 18 years) would be analysed separately from data from children (< 18 years old).

Types of interventions

We included studies that experimentally induced SH in healthy participants. Studies were eligible if they induced and/or assessed SH after or during the manipulation. We included studies that used a pharmacological method to manipulate the NMDA system, with effects on SH. For studies to be eligible, participants had to have received a chemical substance through injection, ingestion or topical application to influence physiological functioning. A typical example of an eligible study is one that used oral or intravenous Ketamine to reduce SH after an experimental burn injury (McGuinness et al., 2011). Studies had to have used a valid control condition, for example, (i) a separate group of participants who received a sham manipulation (e.g. inactive drug) or (ii) the same group of participants who received the manipulation (e.g. active drug) and the control on different days, or (iii) a control site (e.g. opposite arm with induction but without the experimental manipulation, if the effect of the manipulation was anticipated to be local, rather than systemic).

Types of outcome measures

Included studies had to have used a valid and reliable outcome measure to determine pain or sensation by participant self-report (Fillingim et al., 2016). In chapter 1, we defined pain as "an unpleasant sensory and emotional experience associated with, or resembling that associated with, actual or potential tissue damage" (Raja et al., 2020). We restricted pain rating to self-reports given by participants to align with the definition of pain as a personal experience that can be influenced by biological, psychological, and social factors (IASP, 2021a).

Primary and secondary outcomes

The primary outcome was the magnitude of SH, assessed using a mechanical punctate stimulus applied to the area surrounding the induction site and the relevant control, and conceptualised as a change/difference between ratings of the intensity of the stimulus at different time points or sites. This reflects the definition of hyperalgesia as '*increased* pain' (IASP, 2021b). We were not interested in SH to other test stimuli, such as contact heat. Studies were grouped separately based on their use of within-subject (i.e., between time points or between sites) or between-group comparisons.

The secondary outcomes of this review are (i) the surface area of SH, measured using a reproducible method (e.g., radial lines approach) and (ii) the risks associated with the induction and manipulation, defined as adverse events (e.g., damage).

Information sources

Electronic searches

We searched the following electronic databases on the 24th of June 2019 and updated our search on the 29th of September 2022: Biosis (via Web of Science), Cochrane library, PubMed (includes MEDLINE), PsycArticles, PsychInfo, ScienceDirect, Scopus, Web of Science (Core and Biosis). All terms in the search strategy were searched for in the title, keywords, or abstract. Where possible, the search strategy was limited to humans. The search strategy for all database was:

((“human*” OR “women” or “woman” OR “man” OR “men” OR “participant*” OR “volunteer* OR individual*”)

OR

“normal skin” OR “healthy skin”)

AND

(“secondary hyperalgesia” OR “punctate hyperalgesia” OR “pinprick pain” OR “pinprick hyperalgesia” OR “mechanical hyperalgesia” OR “mechanical pain” OR “heat hyperalgesia” OR “neurogenic hyperalgesia”))

Other resources

We identified additional eligible studies that were not discovered by electronic searching. To do this, we checked the citations and reference lists for searches of retrieved papers. In addition, we contacted experts in the field and authors who have recently published narrative reviews on experimental induction and manipulation of SH and asked for their assistance in identifying any missed studies.

Data collection and analysis

Data management

The initially proposed platform (Systematic Review Facility; http://syrf.org.uk/) was cumbersome for managing the review process, so we deviated from the protocol to use Covidence (https://covidence.org/).

Selection process

Two of four reviewers (GJB, LM, PM, PCC) independently determined the eligibility of each study identified by the search in two phases. In the first phase, the reviewers independently screened the titles and abstracts of identified studies and eliminated studies that did not satisfy the inclusion criteria.

The reviewers obtained full-text versions of the remaining studies in the second phase. They independently screened the full-text version of these studies using the customised eligibility form (see Appendix A). When information was missing from the full texts, we planned to contact study authors a minimum of three times over at least six weeks (see Dealing with missing data). Discussion or independent adjudication by a fifth reviewer (VJM) resolved disagreements about study eligibility. We did not anonymise the studies before assessment.

Risk of bias assessment

We used a custom-modified version of the Cochrane Collaboration's tool (Higgins et al., 2011) to assess the risk of bias (RoB) of the included studies (Appendix B). The RoB tool was refined during piloting to suit the aim and objectives of this review. Three reviewers (GJB, LM, PM) independently reviewed each study for six types of bias: selection, performance, detection, manipulation veracity, attrition, measurement, and reporting bias. Notably, the RoB assessment was not intended to reflect overall study quality; instead, the assessment focused on the risk that the results that were *of interest to this review* might be biased. Each domain included pre-specified questions of importance to direct reviewers' attention to design-relevant features. Reviewers were asked to review the pre-specified domains of importance alongside any other relevant features and then allocate a summary judgement of 'high risk' (i.e., seriously weakens confidence in the results), 'low risk' (unlikely to seriously change or reduce trust in the results) or 'unclear risk' (i.e. insufficient information provided to make judgement) of each type of bias. Judgements from the three independent reviewers were compared, and disagreements were resolved through discussion or independent adjudication by a fourth reviewer (VJM). If studies scored a 'high risk' for the performance, detection, measurement and selection bias domains, the study was scored as having an overall 'high risk' of bias because these domains affect the results that are of interest to this review.

Selection bias

The criteria for assessing selection bias addressed the methods used to (i) recruit participants, (ii) determine the pain status of participants, (iii) determine whether participants had similar baseline demographics (where applicable), and (iv) generate random allocation for participants into manipulation groups. For example, in (iv), we checked the method used to randomly allocate participants into groups (with or without manipulation). If an appropriate method (e.g., a computer-operated random number generator) was used to achieve random allocation, we scored the article as having a low RoB. If studies failed to use an appropriate randomisation method, we scored the article as having a high RoB. If insufficient information was given (e.g., reported randomisation without method used for random allocation), we scored the article as having an unclear RoB.

Performance and detection bias

The criteria for assessing performance and detection bias addressed the methods used to blind participants and outcome assessors to the research question, paradigm, and/or group allocation, including providing a blinding check. For example, if (i) an appropriate method (e.g., manipulation given in identical syringes) was used to blind participants and outcome assessors, and (ii) a blinding check was provided, we scored the article as having a low performance/detection bias. If (i) an inappropriate method (manipulation given in different syringes) was used, or (ii) an appropriate method was used, and a blinding check confirmed that blinding was broken, the article scored a high performance/detection bias despite providing a blinding check. If (i) an appropriate method was used but no blinding check was provided, or (ii) insufficient information was provided, the article was scored as having an unclear RoB.

Manipulation veracity

The criteria for assessing manipulation veracity included verifying the 'manipulation check' tool/method used to confirm or negate the manipulation efficacy. An example of this for a pharmacological manipulation would be checking the blood levels of the drug after the manipulation is given. The idea of manipulation checks is not novel, and evidence from the literature suggests that manipulation checks affect experimental conclusions (Hauser, Ellsworth & Gonzalez, 2018). Here, studies that verified the effectiveness of the manipulation were scored as having a low RoB. However, studies that failed to verify manipulation effectiveness were rated as having a high RoB. Similarly, if a manipulation check was completed, and the results confirmed the ineffectiveness of the manipulation, these studies were rated as having a 'high' RoB. If insufficient information was reported, the study was rated as having an unclear RoB.

Attrition bias

The criteria for assessing attrition bias included reporting of attritions (i.e., exclusions, withdrawals) and handling of incomplete outcome data. Studies that reported attritions and appropriately dealt with them in their analyses were scored as having a low RoB. Those that failed to report attritions were rated as having a high RoB. If insufficient information was reported, the study was rated as having an unclear RoB.

Measurement bias

To assess measurement bias, reviewers were prompted to verify the outcome measures used to assess SH (e.g., self-report scales) and consistency of equipment (e.g., pinprick probes) and assessor across groups, sites, or time points (where relevant). Studies that used the same outcome assessment approach consistently across the course of the study were scored as having a low RoB, and those that

failed to report measurements were rated as having a high RoB. If insufficient information was reported, the study was rated as having an unclear RoB.

Reporting bias

To assess reporting bias, reviewers were prompted to evaluate reporting of (i) pre-specified study outcomes related to this review and (ii) if conflicts of interest and/or funding sources were declared. Studies that reported study outcomes, conflicts of interest and/or funding sources were scored as having a low RoB. In addition, studies that were funded by pharmaceutical companies that produce NMDA targeting drugs were scored as having a low risk of bias, provided they disclosed their funder(s). Those that failed to report study outcomes, conflicts of interest and funding sources were rated as having a high RoB. If insufficient information was reported, the study was rated as having an unclear RoB.

Data extraction

We extracted information for the control and experimental group/site from the included studies. We extracted information on the (i) study designs, (ii) participants' characteristics (e.g. eligibility criteria, age, sex), (iii) induction and manipulation methods, (iv) rate and nature of adverse events associated with the induction (e.g. blisters, erythema) and manipulation methods (e.g. dizziness, sedation), (v) secondary hyperalgesia outcome measurement tools, (vi) time points when SH was measured, (vii) pain reported by participants before and/or after the induction (excluding pain ratings to induction stimulation), and (viii) data related to the primary and secondary outcomes.

Two reviewers (GJB, LM) independently extracted data from each included study using a data extraction tool (https://redcap.uct.ac.za/surveys/?s=EYAYT9CY3A) hosted on Redcap (https://redcap.uct.ac.za/). Data extraction was duplicated by the two reviewers using excel. Before formal data extraction, we piloted and improved the data extraction tool using five studies. Reviewers compared data extracted before formal analysis.

Dealing with missing data

When data were missing, we planned to contact study authors a minimum of three times over at least six weeks. If the authors did not respond or provide the data during the 6 weeks, then we would classify the data as unavailable for this review.

Data analysis and synthesis

This is a non-Cochrane systematic review. Nonetheless, data analysis was guided by the Cochrane Handbook of Systematic Reviews for Interventions (Higgins et al., 2019) which explains the statistical considerations that should be undertaken when analysing data from crossover and within-subject study

designs. Data analysis was performed in two formats: (i) all data were summarised narratively, and (ii) data from a portion of the studies were pooled into a meta-analysis. After completing the narrative summary, data were pooled for a meta-analysis if more than two studies were available for a certain manipulation and if there were no variations between studies in the reported study design, including the methods used to induce and manipulate SH outcomes. We conducted a meta-analysis using R software. Final analysis script used for data analysis is available in Appendix C. The following packages were used for analysis: tidyverse (Wickham et al., 2019b), readxl (Wickham et al., 2019a), gridExtra (Auguie & Antonov, 2017), here (Müller, 2017), kableExtra (Zhu, 2019), ggstatsplot (Patil, 2021), dplyr (Wickham et al., 2022), readr (Wickham, Hester & Bryan, 2022), formatR (Xie, 2022), meta (Balduzzi, Rücker & Schwarzer, 2019), forestplot (Gordon & Lumley, 2022), estmeansd (McGrath et al., 2022), dmetar (Harrer et al., 2019), robvis (McGuinness, 2019), ftExtra (Yasumoto, 2022), tidyr (Wickham & Girlich, 2022), stringr (Wickham, 2022), tibble (Müller & Wickham, 2022), flextable (Gohel & Skintzos, 2022), magrittr (Bache & Wickham, 2022), patchwork (Pedersen, 2022), and magick (Ooms, 2021).

Measures of manipulation effect

We were interested in the size of the effect of the manipulation on SH outcomes. We planned to measure the effect of the manipulation by comparing data either between conditions (experiment versus sham) or between assessments (baseline versus follow-up). Subgroup analyses were planned for both outcomes of SH, with data sub-grouped by the type of manipulation and induction.

Rationale for sub-grouping

We wanted to know which pharmacological manipulation is most effective for manipulating experimentally induced SH. We planned to perform a sub-group analysis by manipulation because evidence suggests that NMDA receptors have different binding affinities (Bresink et al., 1995). In addition, NMDA antagonists have different pharmacodynamic properties. For example, memantine is a moderate affinity NMDA antagonist with rapid unblocking kinetics resembling that of magnesium. We anticipate that the heterogeneity of NMDA receptors and the different pharmacodynamic properties of NMDA antagonists may influence the binding of NMDA antagonists to its receptors, therefore supporting our approach to sub-group by manipulation.

Pooling of data

Data were pooled in subgroups by study design and the type of manipulation (e.g., ketamine) for each outcome. We summarised the results narratively. If more than two studies were available for a certain manipulation data were included in the meta-analysis. Data were pooled from studies of similar manipulations into pooled analyses. Considering the anticipated heterogeneity between studies, the meta-analysis used a random effects model to pool effect sizes and weigh the studies relatively

uniformly (Higgins et al., 2019). The restricted maximum likelihood method for continuous data was used to estimate the between-study heterogeneity (I^2 statistics) (Veroniki et al., 2016; Viechtbauer, 2005), and heterogeneity was considered substantial if the I^2 statistic was greater than 50% (Higgins et al., 2019).

Relative ranking of manipulations

We compared the manipulations by efficacy and risk.

Publication bias

Publication bias was assessed by visually checking the asymmetry of contour-enhanced funnel plots. For each funnel plot, we plotted each study's standardised mean difference (x-axis) against the standard error (y-axis). Contour-enhanced funnel plots show the potential influence of the statistical significance of each study on publication bias (Peters et al., 2008). Publication bias was formally tested using Egger's statistical test for asymmetry, which requires a minimum of 10 studies included in the meta-analysis (Egger et al., 1997). A p-value of less than 0.10 ($p < 0.10$) was considered the significance level for publication bias. The 'funnel meta*'* meta-analysis function from the meta package (Balduzzi S, Rücker G & Schwarzer G, 2019) in R was used to generate funnel plots.

Assessment of the certainty of the body of evidence

We used the GRADE (Grades of Recommendation, Assessment, Development and Evaluation) criteria to rank the quality of evidence for each manipulation method (Guyatt et al., 2008). The GRADE approach uses five domains to assess the quality of the body of evidence for each outcome: (i) risk of bias, (ii) consistency of effect, (iii) imprecision, and (iv) indirectness (Guyatt et al., 2008). Grading of manipulations was conducted using GRADEpro (https://www.gradepro.org/). Each domain was rated as having 'not serious', 'serious' and 'very serious' limitations. If a domain is rated as having a 'serious' or 'very serious' limitation, the certainty of evidence was downgraded by 1. The overall certainty the of evidence was rated on a three-point scale, including 'very low, 'low', 'moderate', and 'high' certainty. If the certainty of evidence is 'very low' it indicates that any estimate of effect is very uncertain. A 'low' certainty indicates that future research is likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate. A 'moderate' certainty indicates that future research would probably have an important impact on our confidence in the estimate of effect and may change the estimate. A 'high' certainty means that future research is very likely to change our confidence in the estimate of effect.

Ethical considerations

This review gathered evidence from studies to which healthy human participants had previously consented. None of the data included in this review is from participants who did not consent.

Results

Results of search

In the extensive review, the initial search (conducted on 24 June 2019) of the 8 databases retrieved 4809 records investigating pharmacological and non-pharmacological methods used to manipulate experimentally induced SH in healthy participants. After rerunning the search, additional eligible records were identified: 30 (October 2019) and 143 (August 2020). One record was identified through communication with an expert in the field. After removing duplicate studies, we retained 2425 records. We remained with 244 records after screening the titles and abstracts and retained 165 records after screening the full texts of the records. We retained 143 studies investigating pharmacological methods; the excluded non-pharmacological studies are not reported here. Finally, we identified and included 29 records with pharmacological manipulations that directly target the NMDA system. The search was updated again on the 29th of September 2022 and no additional studies were eligible for this review. Figure 1 uses a PRISMA diagram (Liberati et al., 2009) to report the process of identifying studies eligible for this review.

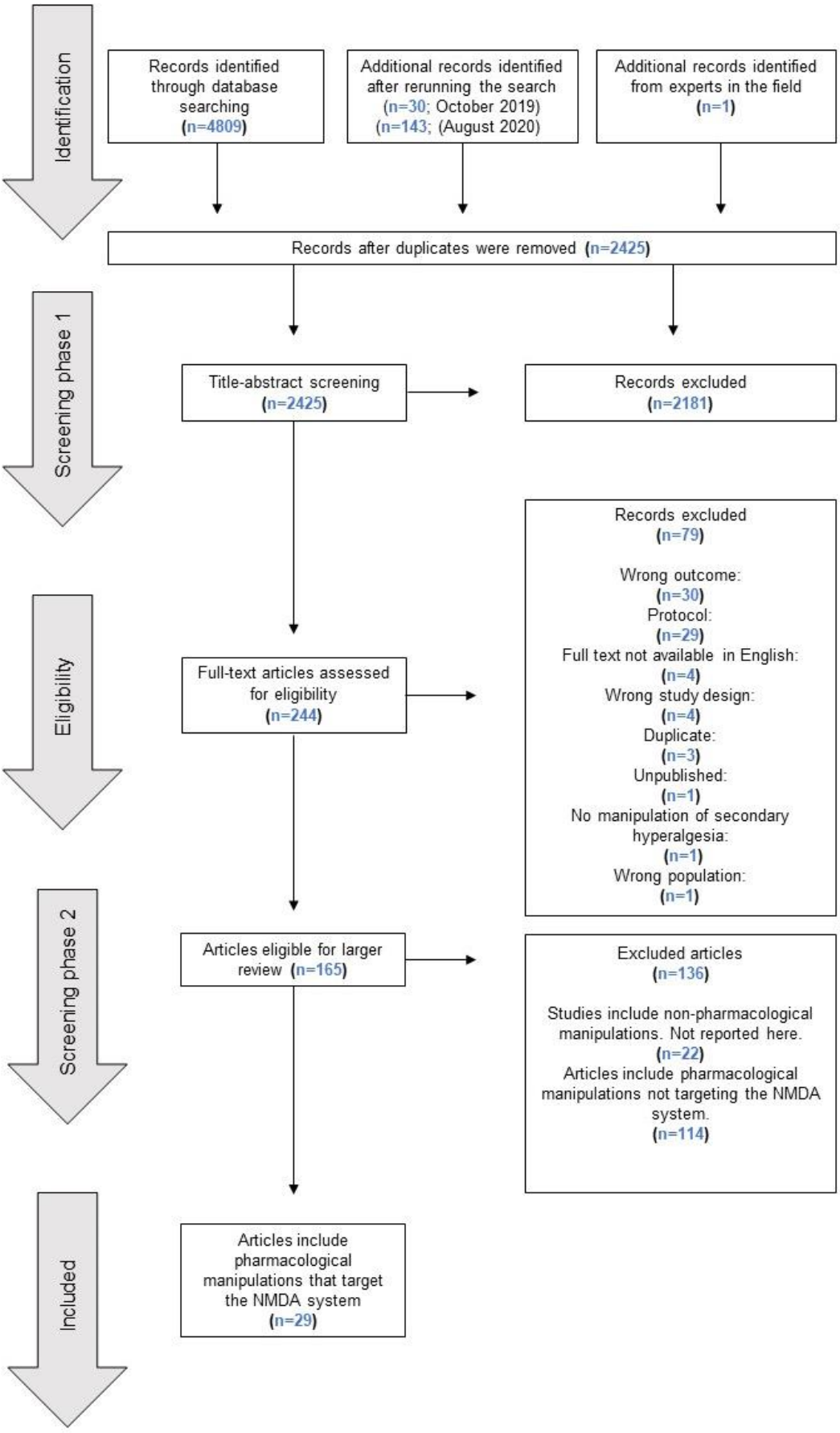


***Figure 1**: The process of selecting studies for this review.*

Included studies

Types of studies

Some records reported multiple manipulation methods, yielding multiple data sets. Here, we refer to each eligible dataset as a 'study'. Of the 29 records included, 16 (Duedahl et al., 2005; Frymoyer, Rowbotham & Petersen, 2007; Ilkjaer et al., 1996; Ilkjaer et al., 1997; Koppert et al., 1999; Koppert et al., 2001; Koppert et al., 2003; Martin et al., 2019; Mathiesen et al., 2006; Mikkelsen et al., 1999; Mikkelsen et al., 2000; Park et al., 1995; Pedersen, Galle & Kehlet, 1998; Pöyhiä & Vainio, 2006; Schulte, Sollevi & Segerdahl, 2004; Sethna et al., 1998) provided data from multiple 'studies'. Therefore, the 29 records yielded 52 datasets or 'studies'. However, it is important to note that some of these 'studies' provide data obtained from the same individuals. For example, a single record could have reported on a crossover design involving three conditions, where two of the between-condition comparisons were relevant to the current review. In this case, we included the two comparisons as separate 'studies'. Study designs included crossover (n=48) and within-subject (n=4) comparisons.

Participants

These 52 studies represented a total of 1007 participants who were deemed healthy and participated in both the experimental and control conditions. In one study (Park et al., 1995) that included 12 participants, only 6 participants received the experimental condition, whereas 12 received the control condition. Of the 52 included studies, 50 reported on sex of participants: 867 participants were identified as male and 122 as female, leaving 18 for whom sex was not reported. In 48 of the 52 records that reported participant sex, males outnumbered females. All participants were generally healthy adults (age ≥ 18 years). Thirteen of 52 studies did not report eligibility criteria (Gottrup, Bach & Jensen, 2004; Klein et al., 2007; Koppert et al., 1999; Mikkelsen et al., 2001; Park et al., 1995; Pedersen, Galle & Kehlet, 1998; Pöyhiä & Vainio, 2006; Wallace, Braun & Schulteis, 2002; Wallace et al., 2002). Thirty-two of the 52 studies reported attrition.

Types of interventions

The included studies used a variety of methods to experimentally induce SH. Capsaicin was used in 21 of the 52 studies and was administered through an intradermal injection (n=16) or topically applied alone (n=1) or with heat (n=4). Burn injury (n=15) was the second most common model, followed by the brief thermal stimulation (n=9), and intradermal electrical stimulation (n=4) models. The least common models were freezing injury (n=2) and high-frequency electrical stimulation (n=1).

Across the eligible studies, we identified five different pharmacological manipulations that likely target the NMDA system. The most commonly used manipulation was ketamine (n=38 of 52 studies), which was given either alone (n=34) or with another pharmacological manipulation (alfentanil (n=2), morphine (n=1), or remifentanil (n=1)). Dextromethorphan (n=10 of 52 studies) was given either alone

(n=8) or combined with morphine (n=2). The remaining pharmacological manipulations were CHF3381 (n=2), magnesium sulphate (n=1), and neramexane (n=1). The manipulations were delivered intravenously (n=28), orally (n=15), subcutaneously (n=5), transdermally (n=2), or intradermally (n=2).

Outcome measures

Results are reported by record (and study); therefore, we have results from 29 records (with 52 studies). Six records (11 of 52 studies) assessed both the magnitude and surface area of SH (Gottrup, Bach & Jensen, 2004; Gottrup et al., 2000b; Gottrup et al., 2000d; Park et al., 1995; Sethna et al., 1998; Pedersen, Galle & Kehlet, 1998). Three records (4 of 52 studies) assessed only magnitude (Klein et al., 2008b; Klein et al., 2007; Pöyhiä & Vainio, 2006;), and 20 records (37 of 52 studies) assessed only the surface area of SH (Andersen et al., 1996; Angst et al., 2003; Duedahl et al., 2005; Frymoyer, Rowbotham & Petersen, 2007; Ilkjaer et al., 1996; Ilkjaer et al., 1997; Koppert et al., 1999; Koppert et al., 2001; Koppert et al., 2003; Martin et al., 2019; Mathiesen et al., 2006; Mikkelsen et al., 1999; Mikkelsen et al., 2000; Mikkelsen et al., 2001; Schulte, Sollevi & Segerdahl, 2004; Wallace, Braun & Schulteis, 2002; Wallace et al., 2002; Warncke, Stubhaug & Jørum, 1997; Warncke, Jørum & Stubhaug, 1997; Warncke, Stubhaug & Jørum, 2000). The summary of ratings scales used to assess magnitude of SH in each study are presented in Table 2. Five records (eight of 52 studies) assessed adverse events.

Table 2: *Rating scales used to assess magnitude of SH in eligible studies*

Study ID for current review	**Study reference** ***(* if assessed both the magnitude and surface area of SH)***	**Rating scale (*scale anchors*)**
105	Gottrup, Hanne et al., 2000*	0; 100 (*no pain; unbearable pain*)
106	Gottrup, Bach & Jensen, 2004*	0 (*no pain*); 100 (*unbearable pain*)
107	Gottrup, H et al., 2000*	0 (*no pain*); 100 (*unbearable pain*)
110	Klein et al., 2008	0 (*non-painful*); 100 (*most intense pain imaginable*)
111	Klein et al., 2007	0 (*not painful*); 100 (*most intense pain imageable*)
120.1-120.2	Park, K. M. et al., 1995*	0; 200 (*not reported*)

121.1-121.2	Pedersen, Galle & Kehlet, 1998*	0; 100 (*not reported*)
122.1-122.2	Pöyhiä & Vainio, 2006	0; 10 (*not reported*)
124.1-124.4	Sethna et al., 1998*	0 (*no pain*); 100 (*worst pain*)

Risk of bias in included studies

The 29 eligible records were assessed for risk of bias. We did not assess risk of bias by dataset; where a record contributed more than one dataset, we assessed the entire record for risk of bias and here, we report those findings. The risk of bias findings are shown in Figure 2.

Study ID and reference	Selection bias	Performance bias	Detection bias	Manipulation veracity	Attrition bias	Measurement bias (magnitude)	Measurement bias (area)	Reporting bias	RoB summary
101 - Andersen et al., 1996						n/m			
102 - Angst et al., 2003						n/m			
103 - Duedahl et al., 2005						n/m			
104 - Frymoyer, Rowbotham & Petersen, 2007						n/m			
105 - Gottrup, Hanne et al., 2000									
106 - Gottrup, Bach & Jensen, 2004									
107 - Gottrup, H et al., 2000									
108 - Ilkjaer, S et al., 1997						n/m			
109 - Ilkjaer, S et al., 1996						n/m			
110 - Klein et al., 2008							n/m		
111 - Klein et al., 2007							n/m		
112 - Koppert et al., 2001						n/m			
113 - Koppert et al., 2003						n/m			
114 - Koppert et al., 1999						n/m			
115 - Martin et al., 2019						n/m			
116 - Mathiesen et al., 2006						n/m			
117 - Mikkelsen, S. et al., 2001						n/m			
118 - Mikkelsen, Soren et al., 1999						n/m			
119 - Mikkelsen, Søren et al., 2000						n/m			
120 - Park, Karen M et al., 1995									
121 - Pedersen, Galle & Kehlet, 1998									
122 - Pöyhiä & Vainio, 2006							n/m		
123 - Schulte, Sollevi & Segerdahl, 2004						n/m			
124 - Sethna et al., 1998									
125 - Wallace, Braun & Schulteis, 2002						n/m			
126 - Wallace et al., 2002						n/m			
127 - Warncke, Jørum & Stubhaug, 1997						n/m			
128 - Warncke, Stubhaug & Jørum, 2000						n/m			
129 - Warncke, Stubhaug & Jørum, 1997						n/m			

***Figure 2**: Risk of bias judgements by domain and study. Green: low RoB; orange: unclear RoB; red: high RoB. If either surface area or magnitude of SH was not measured, measurement bias was not assessed for that domain and is indicated as 'n/m'.*

Selection bias

None of the records was judged to be at low risk or high risk of selection bias. We identified an unclear risk of selection bias in 29 (of 29) records. Four (of 29) records used an enriched design, i.e., participants were excluded if they did not develop SH (Gottrup et al., 2000b; Park et al., 1995) or allodynia (Gottrup et al., 2000a; Gottrup et al., 2000b). Given our research question is interested in the effects of a manipulation on SH, we considered these enriched designs to be an appropriate recruitment strategy. In addition, one study reported recruitment strategy (Martin et al., 2019). All records reported that included participants were pain free; however, none of the records provided the method they used to determine pain status (e.g., self-report using Brief Pain Inventory). All participants had similar baseline demographics in the included records. Thirteen of (29) records used an appropriate method to randomly allocate manipulations. However, 16 of (29) records failed to clearly report random allocation, and eight records within this subset of studies failed to report the method used to generate random allocation of study manipulations.

Performance and Detection bias

None of the records was judged to be at low risk of performance and detection bias. We identified a high risk of performance and detection bias in four records that reported (i) more adverse events at the beginning or shortly after the infusion of dextromethorphan (Duedahl et al., 2005), and (ii) that blinding was impossible due to psychotomimetic side effects of ketamine (e.g., sedation) (Klein et al., 2007), and (iii) blinding was incomplete because (a) almost all participants experienced dissociative effects of ketamine (Koppert et al., 2001) and (b) lack of adverse events were reported by the sham condition (Schulte, Sollevi & Segerdahl, 2004). We identified an unclear risk of performance and detection bias in 25 records (see Figure 2). No study reported attempts to blind analysis staff, and none reported blinding assessments of any kind.

Veracity of manipulation

We considered whether a manipulation check was used to confirm the effectiveness of pharmacological manipulations. All 29 records scored a low risk of bias.

Attrition bias

We judged 14 (of 29) records to have a low risk of attrition-related bias for disclosing and accounting for withdrawals/exclusions. The remaining 15 records had an unclear risk of attrition-related bias because they failed to disclose and account for withdrawals/exclusions.

Measurement bias

All the studies assessing magnitude had a low risk of measurement bias. All the studies assessing surface area had a low risk of measurement bias except for one which had an unclear risk of bias

because they failed to report equipment (e.g., pinpricks) used to assess the surface area of SH (Andersen et al., 1996).

Reporting bias

Nineteen of (29) records had a low risk of reporting bias. Ten (of 29) records were judged to have a high risk of reporting bias (Andersen et al., 1996; Angst et al., 2003; Gottrup, Bach & Jensen, 2004; Koppert et al., 2001; Koppert et al., 2003; Park et al., 1995; Sethna et al., 1998a; Warncke, Stubhaug & Jørum, 1997; Warncke, Jørum & Stubhaug, 1997; Warncke, Stubhaug & Jørum, 2000). One record failed to report study outcomes related to this review (Gottrup, Bach & Jensen, 2004). Six records did not report either funders or conflict of interest (Angst et al., 2003; Park et al., 1995; Sethna et al., 1998a; Warncke, Stubhaug & Jørum, 1997; Warncke, Jørum & Stubhaug, 1997; Warncke, Stubhaug & Jørum, 2000). Three records declared funders but no conflict of interest (Andersen et al., 1996; Koppert et al., 2001; Koppert et al., 2003). None of the records had an unclear risk of reporting bias.

Narrative summary of the effects of manipulations on measures of SH

Given that the 29 records yielded 52 studies, we now report results by studies. Across the 52 included studies, five different manipulations were tested, all of which were anticipated to decrease SH. The manipulations were ketamine, dextromethorphan, magnesium, CH3381 and neramexane. Seven of 52 studies assessed adverse events (Andersen et al., 1996; Ilkjaer et al., 1997; Koppert et al., 1999). Table 3 shows the summary of the included studies.

Table 3: *Summary of included studies*

Manipulation method										
Study ID and reference	**Study design**	**Total sample size** *(male; female)*	**Age of participants** *(years)*	**Secondary hyperalgesia**				**Effect of manipulation**		
				Induction	**Manipulation**	**Outcome measurement (Magnitude)** rating scale, modality, force **(Area)** method, modality, force *(split if both magnitude and surface area were assessed)*		**Hypothesised direction**	**Magnitude**	**Surface area**
Ketamine (n = 38)										
101 - Andersen et al., 1996	crossover	17 (17;0)	mean (range) 24 (21 - 29)	topical capsaicin	ketamine (intravenous)	4 radial lines, pinprick, not reported		↓*		-
102 - Angst et al., 2003	crossover	10 (10;0)	mean (range) 28 (20 - 35)	intradermal electrical stimulation	ketamine (intravenous)	8 radial lines, punctate probe, 160mN		↓*		↓
105 - Gottrup, Hanne et al., 2000	crossover	12(12;0)	not reported	intradermal capsaicin injection	ketamine (subcutaneous)	0 – 100, von Frey filament, 75.86 g	6 radial lines, von Frey filament, 75.86g	↓*	-	-

106 - Gottrup, Bach & Jensen, 2004	crossover	12(12;0)	not reported	intradermal capsaicin injection	ketamine (subcutaneous)	0 – 100, not reported, not reported	6 radial lines, von Frey filament, 75.86g	↓*	-	-
107 - Gottrup, H et al., 2000	crossover	12(12;0)	not reported	intradermal capsaicin injection	ketamine (intravenous)	0 – 100, von Frey filament, 744.9mN	6 radial lines, von Frey filament, 744.9mN	↓*	↓	↓
109.1 - Ilkjaer, S et al., 1996	crossover	19 (19;0)	range 20 - 31	burn injury	ketamine (low dose; intravenous)	4 radial lines, von Frey filament, 1.15 N		↓*		-
109.2 - Ilkjaer, S et al., 1996										↓
109.3 - Ilkjaer, S et al., 1996				brief thermal stimulation	ketamine (high dose; intravenous)					
109.4 - Ilkjaer, S et al., 1996										
111 - Klein et al., 2007	crossover	8 (4;4)	mean (SD) 25 (3)	high-frequency electrical stimulation	ketamine (intravenous)	0 – 100, pinprick, 8, 16, 32, 64, 128, 256 and 512mN		↓*	-	

112 - Koppert et al., 2001	crossover	12 (8;4)	mean (SD) 31 (8)	intradermal electrical stimulation	ketamine (intravenous)	4 radial lines, von Frey filament, 450mN	↓*		↓
113.1 - Koppert et al., 2003	crossover	13 (13;0)	mean (SD) 31.2 (5.3)	intradermal electrical stimulation	ketamine (intravenous)	4 radial lines, von Frey filament, 450mN	↓*		↓
113.2 - Koppert et al., 2003					ketamine and remifentanil (intravenous)				
114.1 - Koppert et al., 1999	within-subject	12 (9;3)	mean (range) 30.8 (23 - 49)	intradermal capsaicin injection	ketamine (low dose; intradermal injection)	6 radial lines, von Frey filament, 450mN	↓*		-
114.2 - Koppert et al., 1999					ketamine (high dose; intradermal injection)				
118.1 - Mikkelsen, Soren et al., 1999	crossover	not reported	not reported	burn injury	ketamine (intravenous)	4 radial lines, von Frey filament, 1.15N	↓		↓
118.2 - Mikkelsen, Soren et al., 1999									

119.1 - Mikkelsen, Søren et al., 2000	crossover	25 (25;0)	not reported	burn injury	ketamine (low dose; oral)	4 radial lines, von Frey filament, 1.15N		↓*		-
119.2 - Mikkelsen, Søren et al., 2000					ketamine (high dose; oral)					
119.3 - Mikkelsen, Søren et al., 2000				brief thermal stimulation	ketamine (low dose; oral)					
119.4 - Mikkelsen, Søren et al., 2000					ketamine (high dose; oral)					
120.1 - Park, Karen M et al., 1995	crossover	12(8;4)	range 20 - 34	intradermal capsaicin injection (before infusion)	ketamine (intravenous)	0 – 200, safety pin, 1 to 17g	8 radial lines, safety pin, 1 to 17g	↓*	-	↓
120.2 - Park, Karen M et al., 1995		12 (7;5)	range 21 - 37	intradermal capsaicin injection (after infusion)					-	↓
121.1 - Pedersen, Galle & Kehlet, 1998	crossover	15 (12;3)	range 26 - 48	burn injury	ketamine (local; subcutaneous)	8 radial lines, von Frey filament, 462mN		↓*	-	-

121.2 - Pedersen, Galle & Kehlet, 1998					ketamine (systemic; subcutaneous)				
122.1 - Pöyhiä & Vainio, 2006	crossover	not reported	mean (SD) 36 (8.6)	intradermal capsaicin injection	ketamine gel at induction site (transdermal)	0-10, von Frey filament, 164.32g	↓*	↓	
122.2 - Pöyhiä & Vainio, 2006					ketamine gel at control site (transdermal)				
123.1 - Schulte, Sollevi & Segerdahl, 2004	crossover	11 (6;5)	mean (range) 36 (22-50)	burn injury	ketamine (intravenous)	8 radial lines, von Frey filament, 45g	↓		↓
123.2 - Schulte, Sollevi & Segerdahl, 2004					ketamine and morphine (intravenous)				-
124.1 - Sethna et al., 1998	crossover	12 (10;2)	range 20-26	intradermal capsaicin injection	ketamine & alfentanil (small dose; intravenous)	0 -100, safety pin, "weighted"	8 radial lines, safety pin, "weighted"	↓*	-	-
124.2 - Sethna et al., 1998					ketamine (large dose; intravenous)					

124.3 - Sethna et al., 1998					ketamine & alfentanil (small dose; intravenous)					
124.4 - Sethna et al., 1998					ketamine & alfentanil (large dose; intravenous)					
125 - Wallace, Braun & Schulteis, 2002	crossover	12 (5;7)	mean (SD) 26.5 (8.9)	intradermal capsaicin injection	ketamine (intravenous)	8 radial lines, von Frey filament, 5.18 (units unclear)		↓*		-
126 - Wallace et al., 2002	crossover	11 (8;3)	mean (SD) 41.4 (13.2)	intradermal capsaicin injection	ketamine (intravenous)	8 radial lines, von Frey filament, 5.18 (units unclear)		↓		↓
127 - Warncke, Jørum & Stubhaug, 1997	crossover	10 (10;0)	median; 22	burn injury	ketamine (subcutaneous)	8 radial lines, von Frey filament, not reported		↓*		↓
128 - Warncke, Stubhaug & Jørum, 2000	crossover	12 (11;1)	median (range) 24 (21 - 29)	burn injury	ketamine (intravenous)	8 radial lines, von Frey filament, 50.6mN		↓*		↓
129 - Warncke, Stubhaug & Jørum, 1997	crossover	12 (10;2)	median (range) 22 (20 - 29)	burn injury	ketamine (intravenous)	8 radial lines, von Frey filament, 50.6mN		↓*		↓
Dextromethorphan (DXM; n =10)										

103.1 - Duedahl et al., 2005	crossover	24(24;0)	mean (range) 27 (21 - 35)	topical capsaicin and heat	DXM (intravenous)	4 radial lines, von Frey filament, 178g/mm	↓*		↓
103.2-Duedahl et al., 2005	crossover	24(24;0)	mean (range) 27 (21 - 35)	brief thermal stimulation	DXM (intravenous)		↓*		↓
104.1 - Frymoyer, Rowbotham & Petersen, 2007	within-subject	22 (8;14)	median (range) 27 (22 - 52)	topical capsaicin and heat	DXM and morphine (oral)	4 radial lines, von Frey filament, 26g	↓*		-
104.2 - Frymoyer, Rowbotham & Petersen, 2007	within-subject	22 (8;14)	median (range) 27 (22 - 52)	brief thermal stimulation			↓*		↓
108.1 - Ilkjaer, S. et al., 1997	crossover	25 (25;0)	mean (range) 24 (21 - 28)	burn injury	DXM 60 mg (oral)	4 radial lines, von Frey filament, 1.15N	↓*		-
108.2 - Ilkjaer, S. et al., 1997	crossover	25 (25;0)	mean (range) 24 (21 - 28)		DXM 120 mg (oral)		↓*		↓
108.3 - Ilkjaer, S. et al., 1997	crossover	25 (25;0)	mean (range) 24 (21 - 28)	brief thermal stimulation	DXM 60 mg (oral)		↓*		-

108.4 - Ilkjaer, S. et al., 1997	crossover	25 (25;0)	mean (range) 24 (21 - 28)		DXM 120 mg (oral)		↓*		↓
115.1 - Martin et al., 2019	crossover	20 (20;0)	not reported	freeze injury	DXM single dose (oral)	6 radial lines, von Frey filament, 588mN	↓*		-
115.2 - Martin et al., 2019	crossover	20 (20;0)	not reported		DXM multiple (oral)		↓*		-
Magnesium (n = 1)									
117 - Mikkelsen, S. et al., 2001	crossover	25 (25;0)	mean (range) 26 (21 - 42)	topical capsaicin application and heat	magnesium sulphate (intravenous)	4 radial lines, von Frey filament, 21,5g	↓		-
CHF3381 (n = 2)									
116.1 - Mathiesen et al., 2006	crossover	27 (27;0)	median (range) 25 (23 - 30)	topical capsaicin application and heat	CHF3381 (oral)	4 radial lines, von Frey filament, 140g	↓*		↓
116.2 - Mathiesen et al., 2006							↓*		↓

Neramexane (n = 1)									
110 - Klein et al., 2008	crossover	18 (7;11)	mean (SD) 23.3 (1.85)	intradermal capsaicin injection	neramexane (oral)	0-100, pinprick, 8, 16, 32, 64, 128, 256, 512mN	↓*	↓	

* Several studies did not explicitly report the hypothesised direction of effect. We hypothesised the direction of effect based on published literature.

↓ Decreased SH

– No effect on SH

Cells coloured grey if SH outcome was not assessed.

Ketamine

Of the 38 studies that used ketamine, 33 used ketamine alone (Andersen et al., 1996; Angst et al., 2003; Gottrup, Bach & Jensen, 2004; Gottrup et al., 2000b; Gottrup et al., 2000a; Ilkjaer et al., 1996; Klein et al., 2007; Koppert et al., 1999; Koppert et al., 2001; Koppert et al., 2003; Mikkelsen et al., 1999; Mikkelsen et al., 2000; Park et al., 1995; Pedersen, Galle & Kehlet, 1998; Pöyhiä & Vainio, 2006; Schulte, Sollevi & Segerdahl, 2004; Sethna et al., 1998; Wallace, Braun & Schulteis, 2002; Wallace et al., 2002; Warncke, Stubhaug & Jørum, 1997; Warncke, Jørum & Stubhaug, 1997; Warncke, Stubhaug & Jørum, 2000); the remaining five combined ketamine with alfentanil (n = 2) (Sethna et al., 1998), morphine (Schulte, Sollevi & Segerdahl, 2004), and remifentanil (Koppert et al., 2003). Twelve studies assessed the magnitude of SH, of which nine found no effect of ketamine on the magnitude of SH (105; 106; 111; 120.1; 120.2; 124.1 and 124.2) and three found a decrease in SH magnitude (122.1 and 122.2). Thirty-five studies assessed the surface area of SH, of which eighteen found no effect of ketamine on the surface area of SH (Table 3: study ID, 105; 106; 111; 124.1; and 124.2) and 17 found a decrease in SH (102; 109.2; 109.3; 109.4; 112; 113.1; 118.1; 118.2; 120.1; 120.2; 123.1; 126; 127; 128; and 129). Of the two studies that combined ketamine and alfentanil (124.3 and 124.4), neither found an effect of the combination on the surface area of SH. One study used ketamine and morphine and found no effect of the combined manipulation on the surface area of SH (study ID: 123.2). In contrast, the only study that used ketamine and remifentanil found a decrease in the surface area of SH (113.2). The combination of ketamine and remifentanil resulted in a decreased area of SH.

Dextromethorphan

Ten studies used dextromethorphan, of which eight used dextromethorphan alone (103.1; 103.2; 108.1; 108.2; 108.3; 108.4; 115.1; 115.2 and two combined dextromethorphan with morphine (104.1; 104.2). None of the studies that used dextromethorphan assessed the magnitude of SH. Of the 8 studies that aimed to manipulate the surface area of SH using dextromethorphan alone, four found no effect of dextromethorphan (115.1; 115.2; 108.1; 108.3) and four found a decrease in the surface area of SH (103.1; 103.2; 108.2; 108.4). Two studies used dextromethorphan and morphine; one found no effect (104.1) and the other found a decrease (104.2) of the surface area of SH.

Magnesium

One study used magnesium and assessed only the surface area of SH. It found no effect of magnesium on the surface area of SH (117).

CH3381

Two studies used CHF3381 alone and assessed only the surface area of SH. Both found that CHF3381 decreased the surface area of SH (116.1 and 116.2).

Neramexane

One study used neramexane alone and assessed only the surface area of SH. It found that neramexane decreased the surface area of SH (110)

Pooled analyses to estimate effect size across studies

We had to consider several factors before pooling data for meta-analysis. First, comparisons were made between conditions or between assessments. Between-condition comparisons compared the experimental (active manipulation) to the sham condition. In contrast, between-assessment comparisons compared between time points. These were typically presented when there was no sham condition. In most cases, the ratings provided after the induction but before the manipulation (baseline assessment) served as the within-subject control time point for the ratings provided after the induction and manipulation (follow-up/experimental time point).

Second, studies assessed and reported SH outcomes at various time points, but we needed data from a single (ideally, common) time point for the meta-analysis. To account for the variation in the reported time points, we followed one of two approaches to select data for pooling. For between-condition data, we extracted data from the time point after the induction and manipulation when the SH outcome reached its peak SH in the control condition. This was intended to maximise our sensitivity to detect an effect of the manipulation, given that this time point is when the effect of the induction was most apparent. For between-assessment data, there was no sham condition, so here, we extracted data from the time point when the effects of the induction were expected to be most apparent, according to prior literature or descriptions provided in the record itself. For example, the effects of intradermal capsaicin injection (as used in (Sethna et al., 1998)) and intradermal electrical stimulation (Koppert et al., 2003) are typically most apparent between 20 and 40 minutes after the induction, respectively – so we selected the study time point best matched to that time window.

Third, some studies reported the (i) average radius (in cm and mm) (Andersen et al., 1996; Koppert et al., 1999) or (ii) used different approaches to report units of the surface area of SH (e.g., mm^2) (Park et al., 1995). For (i), we assumed a circular area and used the formula Area = πr^2 to calculate area from the group mean and standard deviation provided. Where data were in units other than cm^2 (situation ii), we converted them and rounded to two decimal places (see Table 4).

Table 4: *Example of data calculations and conversions*

Study ID and reference	Original data †		Converted data (cm^2)	
	central tendency *(mean)*	**spread indicator or (precision of the estimate)** *SD or (SEM)*	**central tendency** *(mean)*	**spread indicator or (precision of the estimate)** *SD or (SEM)*
101 - Andersen et al., 1996	2.62	0.71	1.63	21.57
114.1 - Koppert, 1999	21.9	14.6	3.6	15.07
114.2 - Koppert, 1999	29.9	22.3	9.51	28.09
120.1 - Park, Karen M et al., 1995	67	(7)	0.07	0.67
120.2 - Park, Karen M et al., 1995	1084	(273)	2.56	10.84

† We have presented data for only the manipulation condition here for illustration purposes.

Fourth, the studies used various measures of central tendency and spread indicators/precision of the estimate to report their findings. We used various formulae to convert the data to a common measure of central tendency (mean) and spread indicator (standard deviation) (Appendix D). We used the '*metacont'* meta-analysis function from the meta package (Balduzzi S, Rücker G & Schwarzer G, 2019) in R to calculate standardised mean differences, including the 95% confidence interval for continuous data. Considering that the included studies have relatively small sample sizes, we used the Hartung and Knapp method for continuous data to calculate the confidence interval around the pooled effect and to adjust test statistics and confidence intervals (IntHout, Ioannidis & Borm, 2014; Knapp & Hartung, 2003). Some studies reported (i) raw ratings of SH, while others expressed SH as a (ii) percentage change from baseline ratings. We pooled studies in subgroups based on (i) and (ii). For example, for the surface area outcome, we pooled the data separately for (i) and (ii), but both forest plots examine the effect of the surface area of SH. Results of the pooled analyses are reported visually, using forest plots. The pooled effect size of provided in terms of Standardised Mean Difference [95% confidence interval].

Lastly, several studies were excluded from the meta-analysis (see Table 5). Initially, if information was missing from the full texts, we planned to contact study authors a minimum of three times over a period of six weeks. However, due to the time limits for the submission of this Masters dissertation, we could only wait three weeks for authors to send us their data. Unfortunately, we had to exclude

these studies from the meta-analysis. Attempts to contact the authors will continue and this work will be revised in preparation for publication.

Table 5: *Studies excluded from meta-analysis*

Magnitude of SH outcome	
Study ID and reference	**Reason for exclusion**
106 - Gottrup, Bach & Jensen, 2004	missing data for meta-analysis
107 - Gottrup, H et al., 2000	missing data for meta-analysis
110 - Klein et al., 2008	only one study using neramexane
111 - Klein et al., 2007	missing data for meta-analysis
124.1: 124.4 - Sethna et al., 1998	missing data for meta-analysis
Surface area of SH outcome	
Study ID and reference	**Reason for exclusion**
106 - Gottrup, Bach & Jensen, 2004	missing data for meta-analysis
107 - Gottrup, H et al., 2000	missing data for meta-analysis
109.1: 109.4 - Ilkjaer, S et al., 1996	missing data for meta-analysis
117 - Mikkelsen, S. et al., 2001	only one study using magnesium sulphate
123.1: 123.2 - Schulte, Sollevi & Segerdahl, 2004	Impossible to estimate the mean and standard deviation from bc.mean.sd formula due to the nature of the data.
124.1: 124.4 - Sethna et al., 1998	missing data for meta-analysis

Primary outcome

Effect of pharmacological manipulations on the magnitude of SH

Crossover studies that used between-condition comparisons

We were able to pool data from seven crossover studies that used between-condition comparisons to shed light on the effect of single-dose ketamine (five studies) and multiple-dose ketamine (two

studies) on the magnitude of experimentally induced SH.

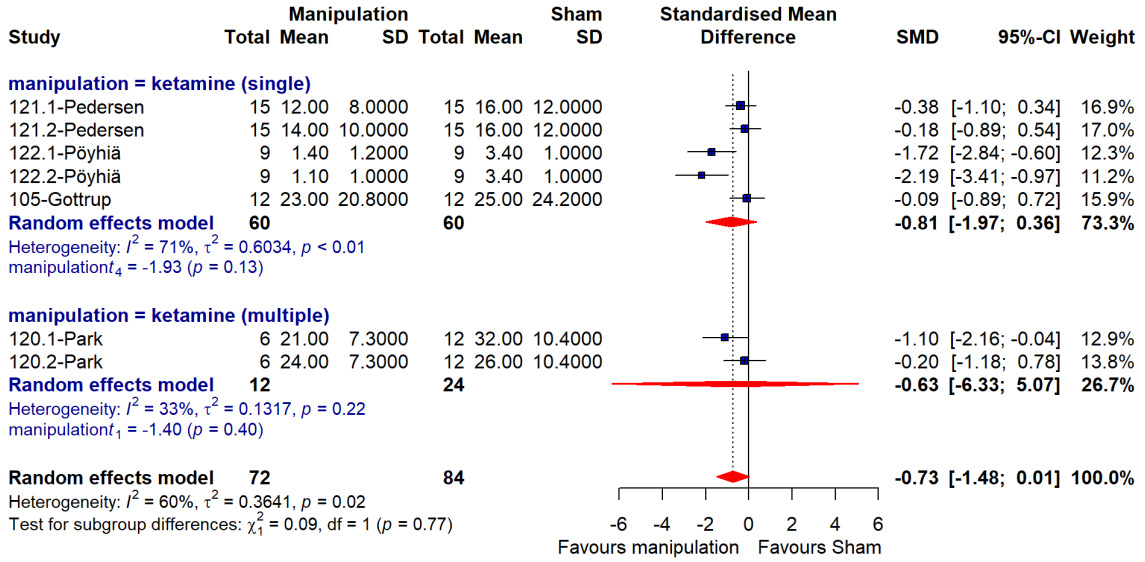


Figure 3 shows that neither single nor multiple doses of ketamine had any effect on the magnitude of SH (single dose: pooled effect estimate = -0.18 [-1.97; 0.36]; p = 0.13; I^2 = 71%; multiple doses: pooled effect estimate = -0.63 [-6.33; 5.07]; p = 0.40; I^2 = 33%).

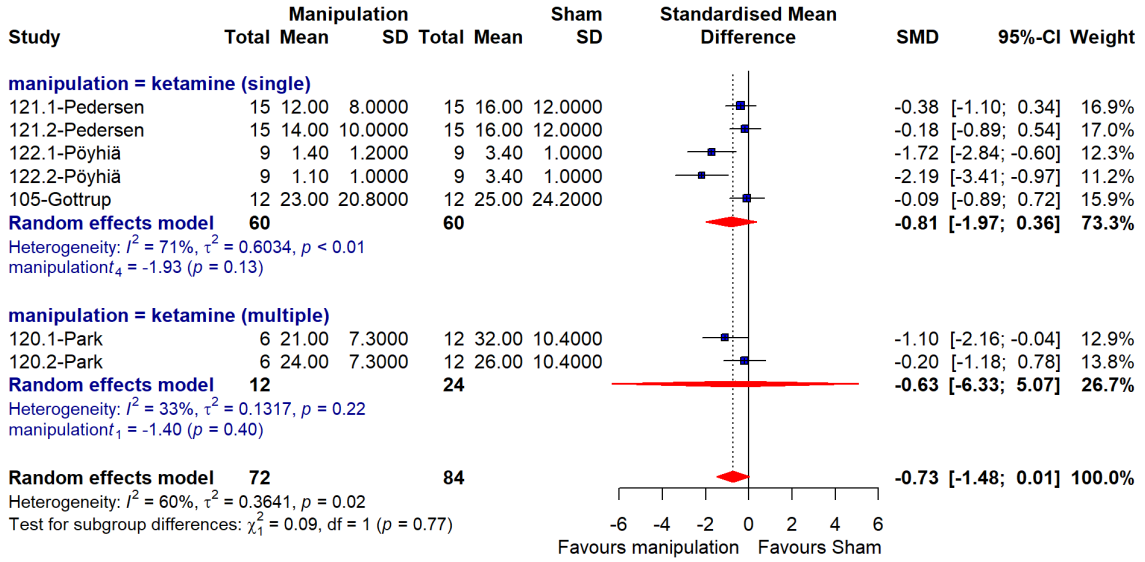


***Figure 3**: Analysis of seven crossover studies that used between-condition comparisons to show the effect of single-dose and multiple-dose ketamine on the magnitude of SH.*

Secondary outcome

Effect of manipulations on the surface area of SH

Studies that expressed surface area as percentage change from baseline

We were able to pool data from eight studies that expressed the surface area of SH as a percentage change from baseline. Six of the studies used a crossover design with between-condition or between-assessment (Figure 4) comparisons to shed light on the effect of single-dose dextromethorphan (n=2), single-dose dextromethorphan and morphine (n=2), including single dose (n=3) and multiple dose (n=1) ketamine on the surface area of SH. The analysis shows that single dose dextromethorphan or coadministration of single-dose dextromethorphan with morphine had no effect on the surface area of SH (dextromethorphan: pooled effect estimate = -0.49 [-3.41; 2.42]; p = 0.28; I^2 = 6%; dextromethorphan and morphine: pooled effect estimate = -0.48 [-1.31; 0.35]; p = 0.09; I^2 = 0%). Single dose ketamine had no effect on the surface area of SH; however, one study showed that multiple dose ketamine reduced the surface area of SH (single dose: pooled effect estimate = -1.37 [-8.57; 5.83]; p = 0.50; I^2 = 94%; multiple doses: pooled effect estimate = -1.62 [-2.76; -0.47]).

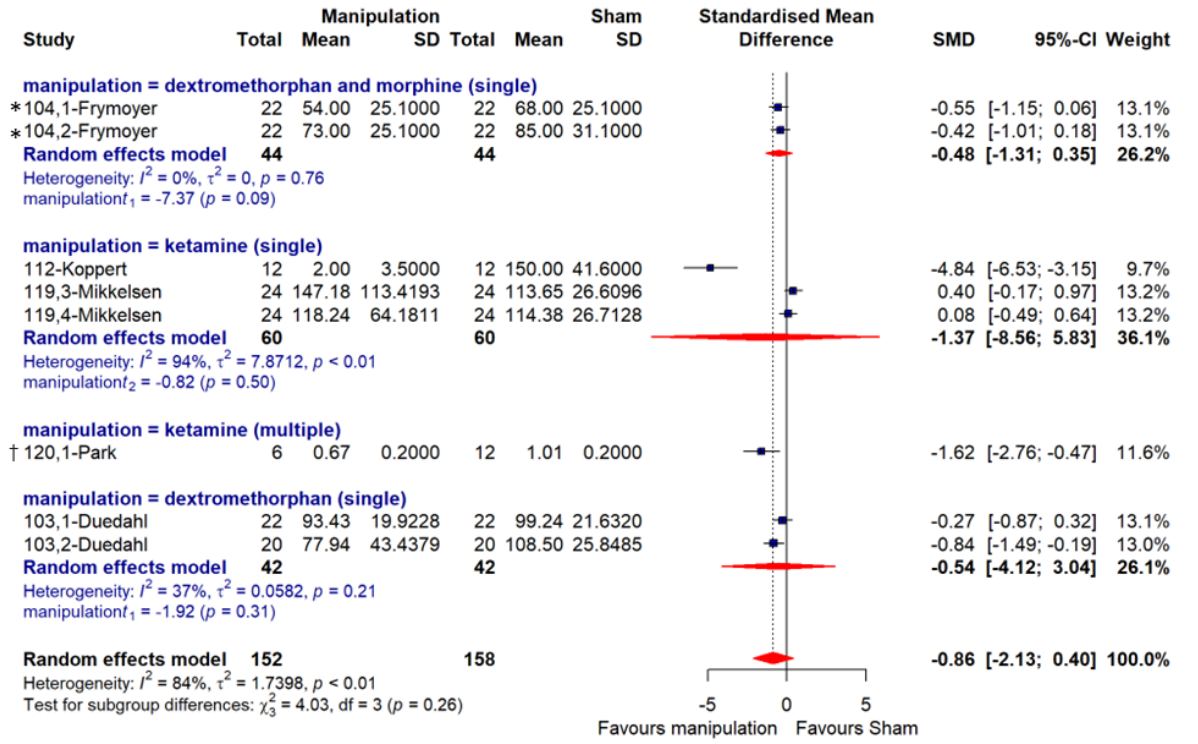


Figure 4: *Analysis of eight crossover and within subject (*) studies designs that used between-condition comparisons showing the effect of dextromethorphan (single dose) dextromethorphan and morphine (single-dose), and ketamine (single and multiple doses) on the surface area of SH. All studies represented here expressed surface area as a percentage change from baseline ratings.*

Studies that reported raw surface area ratings

We were able to pool data from twenty-five studies that reported the raw ratings of the surface area of SH. Twenty-three of the 25 studies were crossover designs that used between-condition comparisons to shed light on the effect of single-dose ketamine (n=11), multiple-dose ketamine (n=4), single dose ketamine and remifentanil (n=1), multiple-dose ketamine and naloxone (n=1), single dose dextromethorphan (n=5), multiple dose dextromethorphan (n=1), and single-dose CHF3381 (n=2). Results are presented below for each manipulation (see Figure 5).

Ketamine

The analysis shows that single dose ketamine, multiple dose ketamine or coadministration of multiple-dose ketamine with naloxone had no effect on the surface area of SH (single-dose ketamine: pooled effect estimate = -0.90 [-2.01; 0.21]; p = 0.10; I^2 = 86%; multiple dose ketamine pooled effect estimate = -1.15 (-2.58; 0.29); p = 0.08; I^2 = 86%; co-administration ketamine with naloxone pooled effect size (-0.31 (-0.89; 0.28)). In one study, the coadministration of single dose ketamine with remifentanil reduced the surface area of SH (pooled effect estimates = -3.19 (-4.40; -1.98)).

Dextromethorphan

The analysis shows that both doses of dextromethorphan had no effect on the surface area of SH (single-dose: pooled effect estimate = -0.13 [-0.51; 0.25]; p = 0.40; I^2 = 6%; multiple dose pooled effect estimate = 0.09 (-0.53; 0.71)).

CHF3381

The analysis shows that single dose CHF3381 had no effect on the surface area of SH (-0.89 [-2.03; 0.26]; p = 0.06; I^2 = 0%).

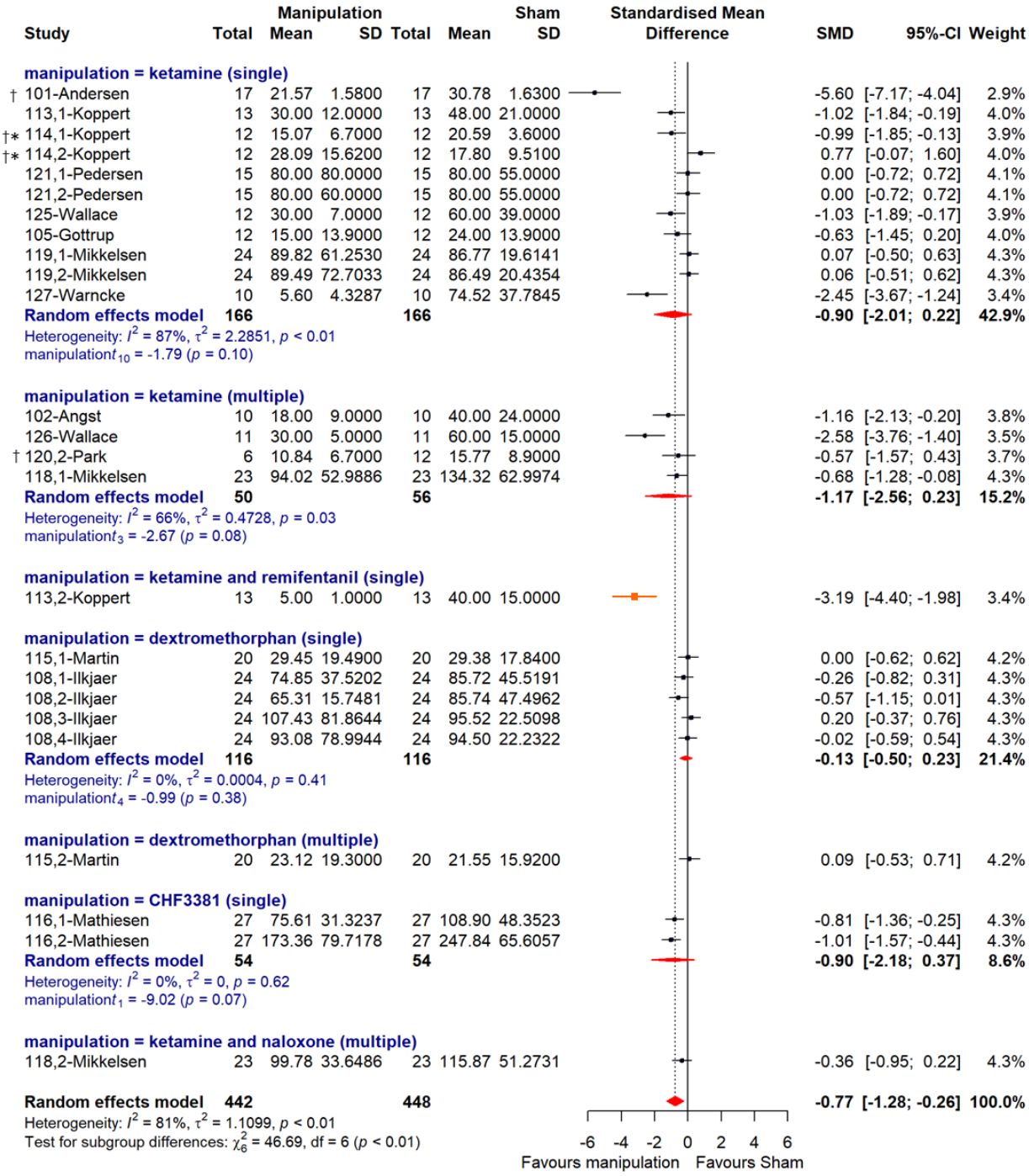


***Figure 5**: Analysis of 25 crossover and within subject (*) studies designs that used between-condition and between assessment (see orange bar) comparisons showing the effect of ketamine (single and multiple doses), ketamine and remifentanil (single dose), ketamine and naloxone (multiple dose), dextromethorphan (single and multiple dose), and CHF3381 (single-dose) on the surface area of SH.*

Publication bias

Magnitude of SH

We included seven studies in the meta-analysis on the effect of manipulations on the magnitude of experimentally induced SH. Initially, we planned to assess publication bias if we had more than ten studies per outcome. However, we opted to assess publication bias anyway for the learning experience. The contour-enhanced funnel plot in Figure 6 suggests that a risk of publication bias may be present. Formal testing confirmed that asymmetry was indeed present (intercept: -6.326, [-9.76; -2.89], p= 0.02), indicating the presence of publication bias.

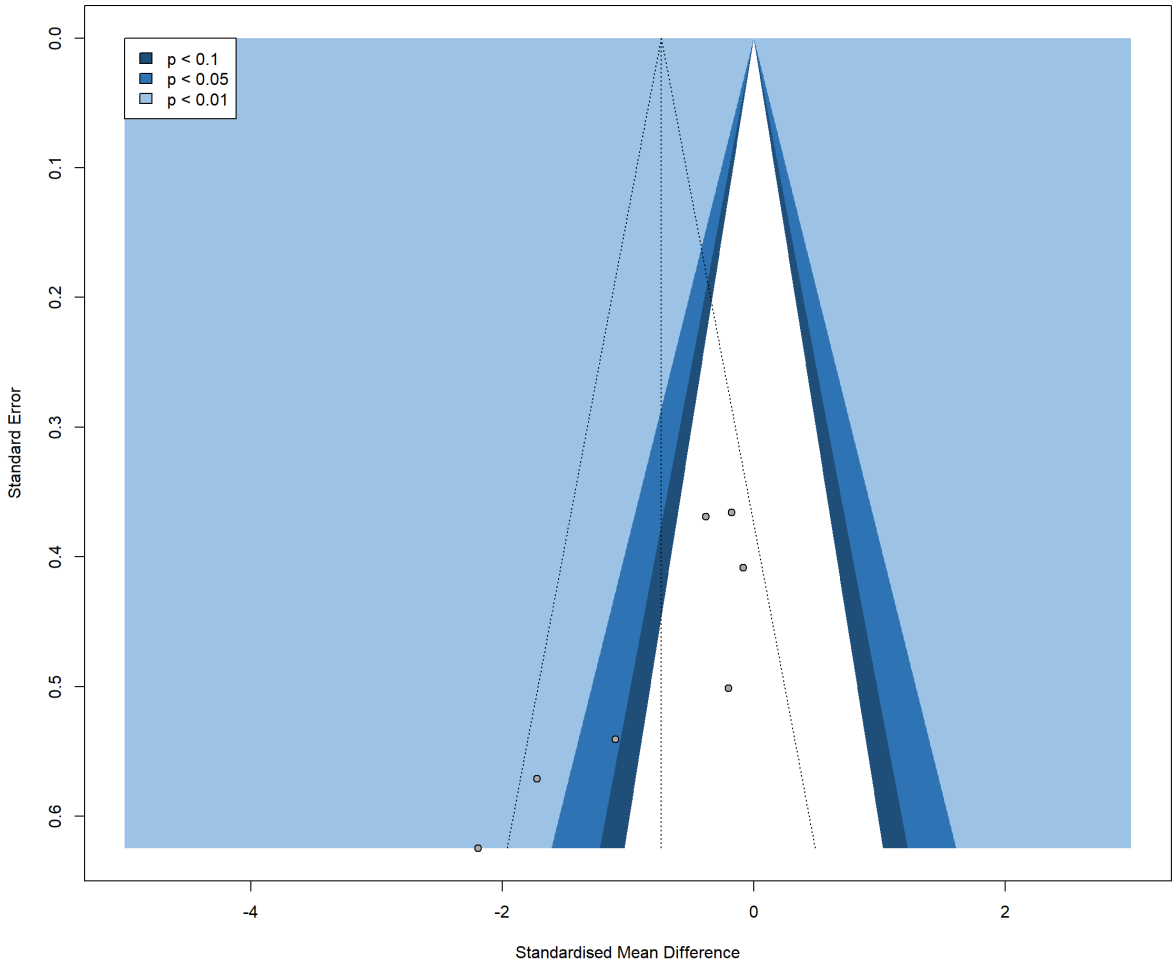


***Figure 6**: Publication bias for studies manipulating the magnitude of SH (n= 7). The contour lines show different thresholds for statistical significance achieved by the studies included in the meta-analysis. Studies falling within the p<0.05 (see blue contour lines) and p<0.01 (see the light blue contour lines) regions have significant effect estimates. Grey dots indicate each study included in the meta-analysis.*

Surface area of SH

We included 33 studies in the meta-analysis on the effect of manipulations on the surface area of experimentally induced SH. The contour-enhanced funnel plot in Figure 7 suggests that the risk of publication bias may be present. Formal testing confirmed that asymmetry was indeed present (intercept: -6.548, [-8.51; -4.59], p= 2.6×10^{-7}), which indicates the presence of publication bias.

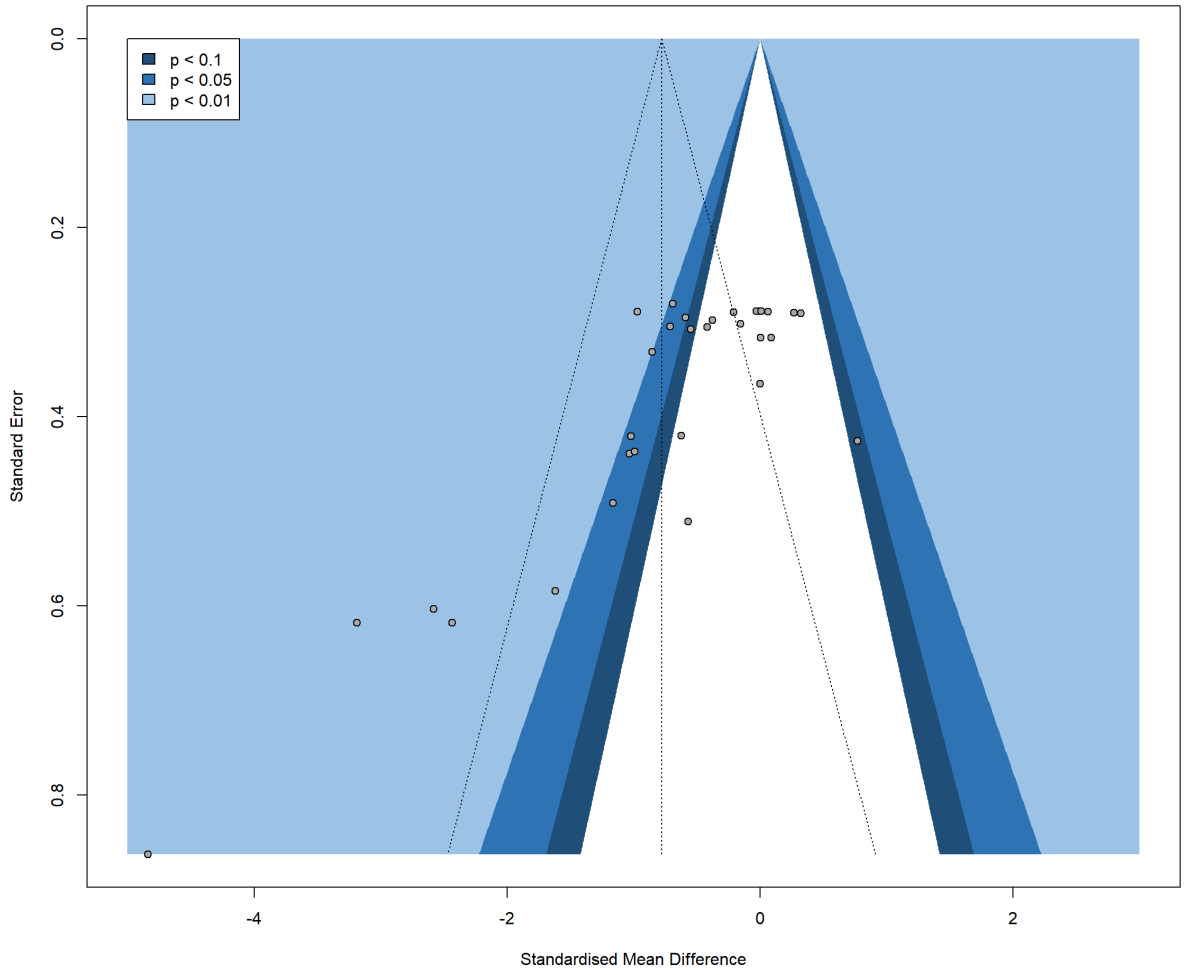


***Figure 7**: Publication bias for studies assessing the surface area of SH (n= 33). The contour lines show different thresholds for statistical significance achieved by the studies included in the meta-analysis. Each grey dot indicates each study included in the meta-analysis.*

Summary of findings and assessment of the certainty of the evidence

The GRADE criteria were used to assess the quality of evidence on the effect of manipulations on the SH outcomes. Below we report GRADE findings by outcome and manipulation.

What is the effect of ketamine on the magnitude of SH?

For single and multiple-dose ketamine, the studies were rated as having a 'serious' risk of bias because all had an unclear risk of bias. Consequently, the risk of bias was downgraded by one. For both doses of ketamine, inconsistency and indirectness were 'not serious'. Imprecision was 'not serious' for single-dose ketamine; however, multiple-dose ketamine had a 'serious' imprecision limitation due to the small sample size. Consequently, imprecision was downgraded by one. Overall, the certainty of evidence that single-dose or multiple-dose ketamine can reduce the magnitude of experimentally induced SH was rated as 'moderate' and 'low' certainty of evidence, respectively (see Table 6).

Table 6: Effect of ketamine on the magnitude of SH

What is the effect of ketamine on the magnitude of SH in the experimental laboratory?									
Number of studies	**Study design**	**Certainty assessment**				**Number of participants**		**Effect**	**Certainty**
		Risk of bias	**Inconsistency**	**Indirectness**	**Imprecision**	**Ketamine**	**Sham**	**Absolute (95% CI)**	
Single-dose ketamine									
5	crossover	serious[a]	not serious	not serious	not serious	60	60	SMD **0.81 lower** [1.97 lower to 0.36 higher]	moderate[c]
Multiple-dose ketamine									
2	crossover	serious[a]	not serious	not serious	serious[b]	12	24	SMD **0.63 SD lower** [6.33 lower to 5.07 higher]	low[d]

[a] The score for risk of bias was downgraded by one because the overall risk of bias was unclear for all studies.
[b] The score for risk of bias was downgraded by one because the sample size is small (n =2), and the CI are wide.
[c] Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.
[d] Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.
SMD = standardised mean difference; CI = confidence interval

What is the effect of ketamine on the surface area of SH?

For single and multiple-dose ketamine, the studies were rated as having a 'serious' risk of bias because all had an unclear risk of bias. Consequently, the risk of bias was downgraded by one. For both doses, inconsistency was a 'serious' limitation and was downgraded by one. For both doses, indirectness was 'not serious'. For both doses, imprecision was 'serious' due to high heterogeneity across studies. Consequently, imprecision was downgraded by one. Overall, the certainty of evidence that single-dose or multiple-dose ketamine can reduce the surface area of experimentally induced SH was rated as 'low' certainty of evidence (see Table 7).

***Table 7**: Effect of ketamine on the surface area of SH*

What is the effect of ketamine on the surface area of SH in the experimental laboratory?									
Number of studies	**Study design**	**Certainty assessment**				**Number of participants**		**Effect**	**Certainty**
		Risk of bias	**Inconsistency**	**Indirectness**	**Imprecision**	**Ketamine**	**Sham**	**Absolute (95% CI)**	
Single dose ketamine									
14	crossover (n = 12) and within-subject (n =2)	serious[a]	serious[b]	not serious	serious[c]	226	226	SMD **0.97 SD lower** [2.03 lower to 0.09 higher]	low
Multiple dose ketamine									
5	crossover	serious[d]	serious[e]	not serious	serious[f]	12	24	SMD **1.23 SD lower** [2.21 lower to 0.25 lower]	low

[a] The score for risk of bias was downgraded by one because 4 of (14) studies scored a low RoB, 1 (of 14) study scored a high RoB, and 9 of (14) studies scored an unclear RoB

[b] The score for inconsistency was downgraded by one because studies 101, 114.1, and 114.2 reported the surface area in different units of measure; we converted all to cm^2
[c] The score for imprecision was downgraded by one because there was high heterogeneity across studies (88%)
[d] The score for risk of bias was downgraded by one because 4 of (5 studies) reported an unclear RoB
[e] The score for inconsistency was downgraded by one because studies 120.1, and 120.2, reported the surface area in different units of measure; we converted all to cm^2
[f] High heterogeneity across studies (59%)
SMD = standardised mean difference; CI = confidence interval

What is the effect of CHF3381 on the surface area of SH?

Single dose CHF3381 was rated as having a 'serious' imprecision limitation because of the small sample size. Consequently, imprecision was downgraded by one. Otherwise, all other domains were not a serious limitation. Overall, the certainty of evidence that single-dose CHF3381 can reduce the surface area of experimentally induced SH was rated as a 'moderate' certainty of evidence (see Table 8).

***Table 8**: Effect of CHF3381 on the surface area of SH*

What is the effect of CHF3381 on the surface area of SH in the experimental laboratory?									
Number of studies	**Study design**	**Certainty assessment**				**Number of participants**		**Effect**	**Certainty**
		Risk of bias	**Inconsistency**	**Indirectness**	**Imprecision**	**CHF3381**	**Sham**	**Absolute (95% CI)**	
Single-dose CHF3381									
2	crossover	not serious	not serious	not serious	serious[a]	54	54	SMD **0.91 SD lower** [3.12 lower to 1.3 higher]	moderate

[a] The score imprecision was downgraded by one because the sample size is small (n =2)
SMD = standardised mean difference; CI = confidence interval

What is the effect of dextromethorphan on the surface area of SH?

Single-dose dextromethorphan, including dextromethorphan and morphine, were rated as having a 'serious' risk of bias limitation because six (of seven) and two (of two) studies scored an unclear RoB. Consequently, the risk of bias was downgraded by one. For dextromethorphan and morphine, the score

imprecision was downgraded by one because the sample size was small. Otherwise, all other domains were not a serious limitation. Overall, the certainty of evidence that (i) single-dose dextromethorphan or (ii) coadministration of dextromethorphan and morphine can reduce the surface area of experimentally induced SH was rated as (i) 'moderate' and (ii) 'low' certainty of evidence (see Table 9).

***Table 9**: Effect of dextromethorphan and coadministration of dextromethorphan and morphine on the surface area of SH*

What is the effect of dextromethorphan on the surface area of SH in the experimental laboratory?									
Number of studies	**Study design**	**Certainty assessment**				**Number of participants**		**Effect**	**Certainty**
		Risk of bias	**Inconsistency**	**Indirectness**	**Imprecision**	**Dextromethorphan**	**Sham**	**Absolute (95% CI)**	
Single-dose dextromethorphan									
7	crossover	serious[a]	not serious	not serious	not serious	158	158	SMD **0.23 SD lower** [0.59 lower to 0.12 higher]	moderate
Single-dose dextromethorphan and morphine									
2	within-subject	serious[b]	not serious	not serious	serious[c]	44	44	SMD **0.48 SD lower** [1.31 lower to 0.35 higher]	low

[a] The score for risk of bias was downgraded by one because 6 (of 7) studies scored an unclear RoB
[b] The score for risk of bias was downgraded by one because all studies scored an unclear risk of bias
[c] The score imprecision was downgraded by one because the sample size is small (n =2)
SMD = standardised mean difference; CI = confidence interval

Adverse events associated with the induction

Five of the 29 records assessed adverse events associated with the induction. Of the five, three records reported no adverse events associated with the burn injury (Warncke, Stubhaug & Jørum, 1997; Warncke, Stubhaug & Jørum, 2000), including topical capsaicin application and heat (Duedahl et al., 2005) models. However, the other two records reported second-degree burns (Mikkelsen et al., 1999) (n =1), including blisters (~ 20% of burn injuries within 24 hours) and slight colour changes persisting three weeks after the induction after using a burn injury model (Pedersen, Galle & Kehlet, 1998) (~ 25% of burn injuries). Twenty-four of the (29) records failed to assess adverse events associated with the induction.

Adverse events associated with the manipulation

Three of (the 29) records failed to assess and report adverse events associated with the manipulations. Of the three records, adverse events were not reported for ketamine (Andersen et al., 1996) and dextromethorphan (Ilkjaer et al., 1997) manipulations. Twenty-six (of the 29) records assessed and reported adverse events associated with the manipulations.

Discussion

The aim of this systematic review and meta-analysis was to understand the effect of NMDA-targeting pharmacological manipulations on experimentally induced SH in healthy human adults without clinical pain. We identified 29 eligible records (yielding 52 studies) that used pharmacological NMDA-targeting manipulations that were expected to influence the magnitude (primary outcome) and/or surface area (secondary outcome) of SH. It is assumed that targeting the NMDA system is important in treating pain underpinned by CS. The proposed mechanism of analgesia via the NMDA system requires an NMDA antagonist to block excitatory synaptic transmission and enhanced synaptic efficiency, all of which contribute to SH. We wanted to know if targeting the NMDA system can reduce the magnitude and surface area of SH, a common feature of most pain conditions. We found conflicting findings between our narrative analysis and meta-analyses.

Effect of NMDA antagonists on the magnitude of SH

In twelve studies that used ketamine to manipulate the magnitude of SH, only three reported a reduction in magnitude, and the other nine found no effect. Similarly, findings from our meta-analysis suggest that single and multiple-dose ketamine has no effect on the magnitude of SH.

Effect of NMDA antagonists on the surface area of SH

Our narrative analysis showed a somewhat equal distribution of findings on the effect of ketamine on the surface area of SH. Almost half the studies (18 of 35) found no effect of ketamine on the surface area of SH, and the rest found a decrease (17 of 35) in the surface area of SH. We generated two meta-

analyses for the surface area outcome (see Figure 4 and Figure 5) to account for the differences in how data were reported in the included studies. However, here, in the discussion, we compare manipulations across all studies that manipulated the surface area of SH. Our findings suggest that a single dose of CHF3381, or dextromethorphan (single and multiple doses), including coadministration of dextromethorphan with morphine, had no effect on the surface area of SH. Similarly, single-dose ketamine, including co-administration of ketamine with naloxone, had no effect on the surface area of SH. Similarly, multiple-dose ketamine had little (see Figure 4, when n=1) to no effect (Figure 5) on the surface area of SH. However, the coadministration of ketamine with remifentanil (n =1) resulted in a decreased surface area of SH – but in only one study.

Our findings suggest that targeting the NMDA system had no effect on SH. In our study, multiple-dose ketamine had a small effect on the surface area but had no effect on the magnitude of SH. These findings suggest that different mechanisms may contribute to generation of the magnitude of SH compared with the surface area of SH.

Mechanisms underlying the development of the magnitude and surface area of SH

Our findings suggest that different mechanisms are likely involved in developing the magnitude and surface area of experimentally induced SH. First, the strengthening of synapses in the CNS following SH induction may reflect the magnitude of induced SH. Second, the effects of the induction 'spreading' to adjacent neural pathways (i.e., beyond the induction site) may reflect the surface area of SH. Ketamine has a low to moderate affinity for NMDA receptors with moderate blocking properties (~86%). Further, ketamine induces NMDA receptor inhibition-dependent synaptic plasticity. It is possible that inhibition may change the connectivity between neurons by reducing the trasmission of information between synapses to decrease connectivity and activity between neighbouring synapses. If true, this may explain the unique role of ketamine in reducing the surface area of SH. These findings suggest that the treatment of SH may require understanding the mechanisms underlying each SH outcome rather than regarding SH as a single feature.

Limitations

First, there was high heterogeneity in the studies included and missing data. Therefore, the pooling of studies into meta-analysis was limited to the available data. Second, the included studies had various study designs, resulting in our conducting multiple meta-analyses for each study design. Third, most data were extracted from figures when no raw data were reported. While data extraction was done in duplicate, it is possible that some data points extracted may be inaccurate. This emphasises the need for researchers to make their raw data publicly available, as recommended by the Organization for Economic Cooperation and Development guidelines (Pilat & Fukasaku, 2007) using a platform such as Open Science Framework (https://osf.io/). Fourth, several calculations were done to convert data from several studies to cm^2. Future studies should consider using standardised units when measuring

SH. Fifth, the studies included had a very low female participation rate, which limits the generalisability of our findings. Therefore, future studies should balance the participation of females and males (Casale et al., 2021). Lastly, the overall risk of bias in the included studies was unclear. The risk of detection, performance, measurement, and reporting bias is important for our research question. The risk of detection and performance bias was largely unclear or high. Some studies failed to report outcomes, placing them at a high risk of detection bias. Therefore, the poor risk of detection, performance and reporting bias weakens our confidence in our findings. In contrast, all studies had a low risk of measurement bias, which increases our confidence in the techniques used to measure SH outcomes.

Conclusion

The findings of this study show that NMDA antagonists had no effect on the surface area or magnitude of experimentally induced SH. Conducting a meta-analysis provides better clarity by pooling the results of multiple studies to diminish the risk of drawing false conclusions that may arise from individual study results. These findings have important implications for clinical practice, as they show the limited effectiveness of NMDA antagonists in modulating SH.

Chapter 2 aimed to elucidate the physiological mechanisms underlying SH in humans to inform more targeted therapeutic approaches. Surprisingly, the bulk of evidence presented in Chapter 2 suggests that NMDA antagonists have no effect on SH, particularly within an experimental model of SH. This finding is notable considering the widespread use of NMDA antagonists for treating neuropathic pain conditions, of which SH is a characteristic feature (Goldberg et al., 2005; Harbut & Correll, 2002; Kiefer et al., 2008; Lin et al., 1998; Takahashi et al., 1998; Ushida et al., 2002). Perhaps NMDA antagonists address other features of neuropathic pain without affecting SH. However, it is worth noting that despite their popularity in clinical practice (Correll et al., 2004; Huge et al., 2010; Leung et al., 2001; Nelson et al., 1997; Pud et al., 1998; Schwartzman et al., 2009) NMDA antagonists are not among the first-line pharmacotherapies for neuropathic pain. Instead, tricyclic antidepressants and serotonin-noradrenaline reuptake inhibitors are pharmacotherapies with the strongest evidence for improving neuropathic pain (Finnerup et al., 2015), and were developed to primarily target distress. These findings raise an intriguing question: could distress be a proximal contributor to SH, potentially acting through the same pathways influenced by these antidepressants? We concurrently conducted a systematic review examining the available evidence on this question (which is not reported in this thesis). Chapter 3 reports on an experimental study we conducted to examine whether distress acts an upstream driver of SH in PWH.

CHAPTER 3: Does distress predict central sensitisation in people living with HIV?

Introduction

The relationship between PSD and CS in individuals living with HIV is not clearly understood. PSD includes psychological and emotional factors such as anxiety and depression and is recognised as a potential contributor to the development and maintenance of CS (Moriki et al., 2022). Understanding the association between PSD and CS in the context of HIV is for informing the management and treatment of pain in this population. This study focused on the relationship between PSD and CS in PWH by drawing on an experimental model of SH, which is a recognised indicator of CS. We aimed to investigate whether PSD severity predicts the severity of SH in PWH. By elucidating this relationship, we can improve our understanding of the underlying mechanisms which can improve pain management and overall quality of life for PWH.

HIV-related pain is poorly understood and difficult to manage. People living with HIV report persistent pain more frequently than their HIV-negative peers (Sabin et al., 2018). In a range of populations, pain has been linked to psychosocial distress (PSD) (Duric et al., 2016), which includes symptoms of depression and anxiety. People living with HIV in sub-Saharan Africa frequently report depression (ranging from 13% - 50%) (Farley et al., 2010; Kagee & Martin, 2010; Kaharuza et al., 2006; Lawler et al., 2011; Monahan et al., 2009; Myer et al., 2008) and anxiety (ranging from 11% - 29%) (Nel & Kagee, 2013; Pefura-Yone et al., 2013; Yunusa, Njoku & Obembe, 2014). Notably, PSD and pain share similar adverse health outcomes, such as impaired functioning (Deshmukh, Borkar & Deshmukh, 2017; Sabin et al., 2018) and reduced quality of life (Deshmukh, Borkar & Deshmukh, 2017; Marcus et al., 2000).

Pain and PSD seem to have a bidirectional relationship. Pain is linked to distress (Edwards et al., 2016; Turk et al., 2016), and distress is commonly linked to pain (Lagana et al., 2002; Miaskowski et al., 2011; Richardson et al., 2009; Rosenfeld et al., 1996; Rotheram-Borus, 2000). Several authors have proposed mechanisms by which PSD could influence pain (Crofford, 2015; Gatchel et al., 2007; Merlin, 2015; Miller, Halkitis & Durvasula, 2019; Woo, 2010). However, these ideas remain speculative without supporting or negating evidence. PSD may be particularly relevant to PWH because PSD is associated with HIV-related stigma (Zhang et al., 2018). Further, PSD perpetuates persistent pain when PSD is inadequately treated (Woo, 2010). Therefore, it is plausible that PSD contributes to persistent pain and the mechanisms underlying this relationship may reflect structural and functional changes within the central nervous system - characterised as central sensitisation (IASP, 2021d).

The HFS human surrogate model is useful for understanding pain mechanisms underpinned by CS. This model induces a short-lived SH that mimics the increased responsiveness of central neurons reported in clinical persistent pain (Klein et al., 2005; Klein et al., 2004). This model generates nociceptive and non-nociceptive input that bypass the receptors to activate A-delta and C fibres directly (31). A-delta activation elicits a pricking pain sensation, whereas C-fibre activation elicits a hot or burning pain sensation (Hansen et al., 2007), all of which are features common in pain underpinned by CS (e.g., neuropathic pain (Finnerup, Kuner & Jensen, 2020)). The various pain sensations elicited by the HFS model reflect the usefulness of this model in understanding pain mechanisms related to SH.

Aim

To investigate the relationships between self-reported PSD, SH (surface area and magnitude) and persistent pain in people with HIV who reported persistent pain (i.e., those who would be expected to have some CS) or who reported having no pain. The surface area (**primary outcome**) and magnitude (**secondary outcome)** of SH were the outcomes of this study.

Objectives

1) Assessed PSD;
2) Induced SH;
3) Assessed whether PSD was associated with the surface area of experimentally induced SH (**primary analysis**) and explored if this relationship was different between groups;
4) Assessed whether PSD was associated with the magnitude of experimentally induced SH (**secondary analysis**) and explored if this relationship was different between groups.
5) Assessed whether the surface area and magnitude of SH differed between groups.

Our primary hypothesis was that we would observe a positive relationship between PSD and SH outcomes (surface area and magnitude) in both groups. Our secondary hypothesis was that the persistent pain group would have greater SH than the no pain group.

METHODS

Study overview

We conducted a single-blinded, case-controlled experimental study with PWH who reported either persistent pain (as cases) or no pain (as controls). This study was linked to a larger parent study that gathered self-report data on pain status (using the Brief Pain Inventory; BPI (Cleeland & Ryan, 1994)) and PSD (Hopkins Symptom Checklist-25; HSCL-25 (Ventevogel et al., 2007)). Researcher 1 collected data in the parent study, and Researcher 2 collected data in this present study. Researcher 1 (a clinical research assistant) screened participants for eligibility and assessed pain and PSD self-

report to support the blinding of Researcher 2 to participants' pain status and PSD severity. Participants' self-reported pain status was used to allocate them into one of two equally sized groups (i.e., pain or no pain). Researcher 2 (LM) induced and assessed SH and performed data analysis. Ratings of pinprick stimuli were given on an electronic vertical visual analogue scale (VAS) before and after the induction. The procedure could be undertaken in either English or isiXhosa, at the participant's choice. A third researcher (VJM) recoded participant ID codes and group membership data to allow Researcher 2 to perform data analysis while maintaining blinding to participant identity and group membership. After the analysis and its interpretation had been finalised, Researcher 2 was unblinded to group membership to allow completion of the current chapter.

Participants, study setting and recruitment

Participants from the parent study (Madden et al., 2022) were invited to participate in the present study as an additional opt-in procedure. Participants were told that the purpose of the study was to understand how sensitivity, distress, pain, and the immune system interact in PWH. The parent study recruited participants with well-controlled HIV from a community health clinic providing HIV services in Khayelitsha, Cape Town, South Africa. Khayelitsha is a peri-urban area in Cape Town, which has a high prevalence of HIV (Lund et al., 2020). Some residents in Khayelitsha experience financial barriers (e.g., unemployment rate is estimated to be 38%) and have access to fewer resources (~55% live in informal housing) (Statistics & Department, 2011). Financial barriers and lack of resources are often associated with distress in PWH (Catz, Gore-Felton & McClure, 2002; Zeng et al., 2019), thus making the population of this area compatible with our study. Therefore, we invited eligible participants from the parent study and intended to classify them into groups defined by pain status. The inclusion of participants was subject to the eligibility criteria specified in Table 10.

***Table 10**: Inclusion and exclusion criteria*

Inclusion criteria
Aged 18 to 65 years
Fluent in spoken English or isiXhosa
HIV positive with evidence of viral suppression *(test showing viral load < 50 copies/ml within preceding three months)*
With persistent pain or without pain (participant-reported). Persistent pain was defined as pain on most days for more than three months (Blyth et al., 2001)
Exclusion criteria
Pregnant or suspects pregnancy

Has an electrical implant, or metal implant in the forearm to be tested
Known neurological, cardiovascular, or acute psychiatric* problems
Has been advised by a medical practitioner to avoid stressful situations
Problems with sensation on the forearm to be tested
Tattoos on the site to be tested

*Excluded individuals in need of acute mental health care. The parent study actively screened volunteers for acute psychosis or suicidality.

Sample size

The sample size was determined *a priori*, based on pragmatism, considering the sample size of the parent study and the additional eligibility criteria for and potential barriers to willingness to participate in the current study. Previous studies of experimentally induced SH without between-group comparisons have used samples of 7-20 (Klein et al., 2004). We calculated that a sample size of n = 53 would provide 85% power to detect a 'small to medium' effect size of r = 0.40 when alpha is 0.05, using a two-tailed Pearson's correlation for the primary analysis of the relationship between PSD and SH (Appendix E). We wished to avoid under-powering, and it seemed feasible to recruit 30 participants per group. Therefore, considering the possibilities of mid-procedure withdrawal and potential minor loss of power due to non-parametric data distribution, we aimed to recruit 60 participants to the study. Recruitment of participants commenced on the 1st of February 2021. Data were collected from the 9th of February to the 24th of November 2021.

Grouping variable: persistent clinical pain

The parent study used an adapted version of the Brief Pain Inventory (BPI) self-report measure (Cleeland & Ryan, 1994) to classify participants into groups (Appendix F). The main BPI has 11 different questions that are grouped into the pain severity subscale (4 questions) and pain interference subscale (7 questions) (Goodin et al., 2018). The BPI is a reliable self-report measure used to study several populations with HIV and persistent pain in South Africa and has been translated and validated in isiXhosa (Parker, Jelsma & Stein, 2017). The opening question from the BPI asked participants to disclose their pain status: *"Throughout our lives, most of us have had pain from time to time (such as minor headaches, sprains, and toothaches). Have you had pain other than these everyday kinds of pain on most days for the past 3 months?"*. A positive response to this first question was followed by a second question on pain frequency and duration. Both questions were used to determine participant eligibility and group membership.

Ethical approval

The study was approved by the Faculty of Health Sciences Human Research Ethics Committee of the University of Cape Town under the parent study protocol number 764/2019 and the current study

277/2021. The study was conducted according to the Declaration of Helsinki (WorldMedical, 2019). We enrolled only eligible, informed and consenting individuals.

Method of induction of SH

We used high frequency electrical stimulation (HFS) and specialised electrodes on the skin (see Figure 8) to induce SH. The cathode consisted of 10 blunt-ended steel pins arranged in a circle, closely resembling those used in other studies (Klein et al., 2008a; Klein et al., 2004). The cathode was attached to the anterior surface of one forearm, approximately 8 cm distal to the cubital fossa, using a double-sided adhesive sticker. The anode was a fabric surface electrode strapped around the participant's upper arm. High-frequency electrical stimulation was delivered in five one-second 100 Hz trains, each followed by a 9-second break, at a current of 10 times the individual's detection threshold for a single electrical stimulus (Klein et al., 2008a). Affect5 software (updated from Affect4 – (Spruyt et al., 2010)) controlled the electrical stimulation, which was produced by a constant current electrical stimulator (DS7A; Digitimer Limited, Hertfordshire, UK) set to maximum voltage (400 V), 2000μs pulse width and a square pulse shape. The effect of the HFS is typically centred at the cathode at the forearm.

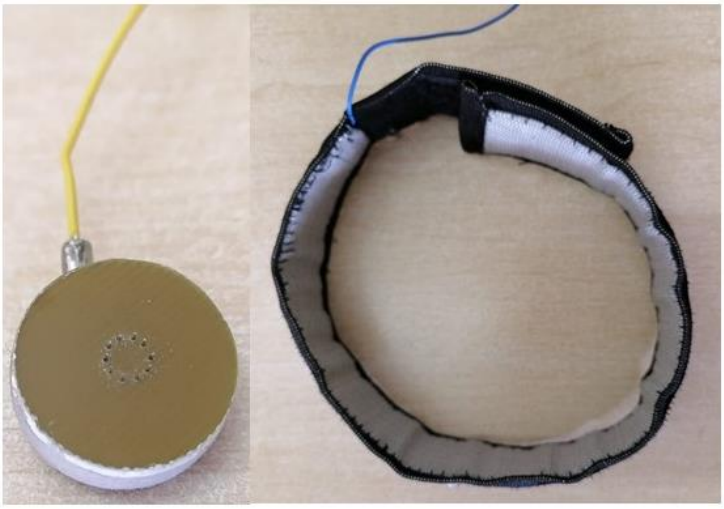


Figure 8*: Cathode (left) and anode (right)*

Arm receiving HFS

One of the parent study's procedures included venepuncture. Given that venepuncture can result in local sensitivity, we induced and assessed SH on the contralateral arm whenever possible. Some participants had contraindications to HFS (e.g., electrical implant) on the contralateral arm or had had venepuncture bilaterally due to difficulties with the blood draw. In those cases, the current study procedure was delayed for a 7-day 'washout period', and the pain free arm was used for the induction and assessment of SH.

Outcome measure: Electronic Visual Analogue Scale (VAS)

All participants used a touch-based electronic vertical VAS to rate each sensory stimulus during sensory testing and HFS induction (Appendix G). This scale was anchored at 0 with “no pain” (bottom) and 100 with “pain ” (top). The anchors were translated into isiXhosa for Xhosa-speaking participants. The VAS was explained to the participant, who was asked to record their rating on the touch screen after being tested with each sensory modality. To give a rating, the participant swiped a stylus pen horizontally across a vertical bar shown (with its anchors) on the screen. In response, the bar filled up with red to the point of the swipe, thus providing visual feedback of the rating. The rating was electronically recorded, and the computer converted the analogue feedback into a numerical rating between 0 (lower anchor) and 100 (upper anchor). The VAS is a validated and reliable measure that requires brief training to be administered and is appropriate for determining pain intensity in diverse adult populations (Melzack & Katz, 2006).

Outcomes

Exploratory predictor variable: psychosocial distress

Psychosocial distress was assessed using the Hopkins Symptom Checklist-25 (HSCL-25, see Appendix H) (Ventevogel et al., 2007). The HSCL-25 is a validated self-report measure consisting of 25 items that assess symptoms of depression (15 items) and anxiety (10 items) (11). Participants indicate the degree to which each symptom of depression (e.g., 'poor appetite', 'feeling lonely' etc.) and anxiety (e.g., trembling, feeling fearful etc.) has applied to them within the preceding month on a four-point Likert scale (1 = 'not at all' to 4 = 'extremely'). The final scores were calculated by dividing the sum of the scores of all the items by 25 to yield a final score ranging between 1.00 and 4.00. The HSCL-25 has been used in South Africa (Kagee & Health, 2005) and with PWH (Kagee, Saal & Bantjes, 2017). To our knowledge, the HSCL-25 has not been formally validated in isiXhosa; we translated it with a formal forward-and-back-translation process (Brislin, 1970) to optimise the local relevance of the language.

Primary outcome: surface area of SH

Our pilot testing of the established method of assessing the surface area of induced SH revealed that (1) some participants do not develop a noteworthy area of SH, yet (2) the conventional, 8-radial-lines approach to assessing the surface area of SH can still falsely ‘identify’ an area of SH in the same participants. The reason appears to be that participants who have not experienced the distinct difference in sensitivity typical of SH will affirm a change in sensation that is quite vague, and the researcher may interpret this report as confirmation of SH. To avoid this problem, we included a brief screening process to identify when individuals had actually developed an area of SH before we assessed the size of that area. This solution maximises specificity at the risk of a loss of sensitivity.

The surface area of SH was screened for and assessed using a von Frey filament (VFF: Marstock nervtest, Germany), exerting a force of 128mN on eight radial lines (Madden et al., 2019). A template was used to mark eight radial lines on the anterior surface of the participant's forearm. The radial lines originated from the stimulation site and were arranged at 45 degrees to each other (Figure 9). The surface area of SH was assessed 30, 45, and 60 minutes after the induction of SH.

Surface area screening

Researcher 2 used the von Frey filament to screen for an area of higher sensitivity by providing two successive stimulations with the von Frey filament and asking, "Does it feel different if I touch you here [1st stimulus, distal to the cubital fossa (see the area marked 'A' on Figure 9)] and here [2nd stimulus, adjacent to the electrode (see the area marked 'A0' on Figure 9)]?" If the participant confidently reported a *distinct* difference in sensation, the surface area of SH was assessed. If the participant reported no distinct difference in sensation between the two points on radial line 'A', we screened from 'E' to 'E0'. If there was also no distinct difference on radial line 'E', we recorded the surface area of SH as "ND" – no difference (equivalent to 0).

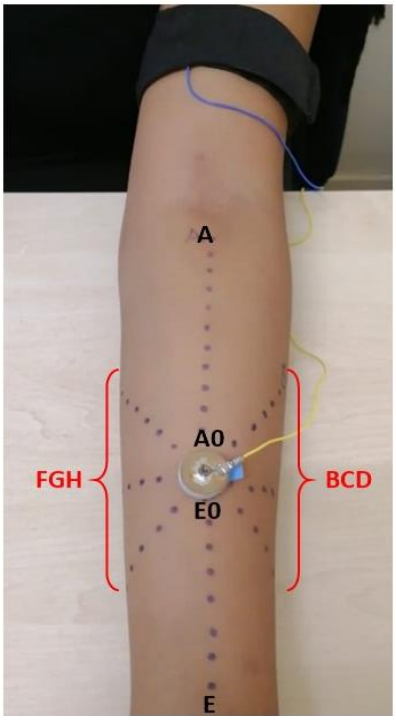


***Figure 9**: The eight lines on the skin indicate radial lines that are 45° apart, and the black dots forming the radial lines are 1cm apart. A to A0 and E to E0 indicates the points for SH screening. No screening was done over radial lines BCD and FGH.*

We acknowledge the limitations of such a screening method, especially the risk of missing an area of SH on the lateral (see 'BCD' in Figure 9) and medial ('FGH' in Figure 9) sides of the forearm where we did not screen for SH. We aimed to prioritise clarity, and procedural requirements prevented immediate estimation of SH magnitude. A previous study found that the area of SH is also typically greater in the proximal-distal axis than in the medial-lateral axis (lines perpendicular to the centre of the HFS application) (Filbrich et al., 2020). In contrast, another study found no differences in the area of SH between the proximal-distal and medial-lateral axes (Della Porta et al., 2022). The two studies differ with respect to the HFS protocol: Filbrich et al. (2020) used an HFS protocol resembling ours, whereas the HFS protocol by Della Porta et al. (2022) induced SH using 12 HFS trains, each of 100 Hz, with each train lasting for 0.42 seconds. Considering the difference in evidence between the two studies, when selecting this assessment approach, we chose to err on the side of underestimating the surface area of SH rather than overestimating it.

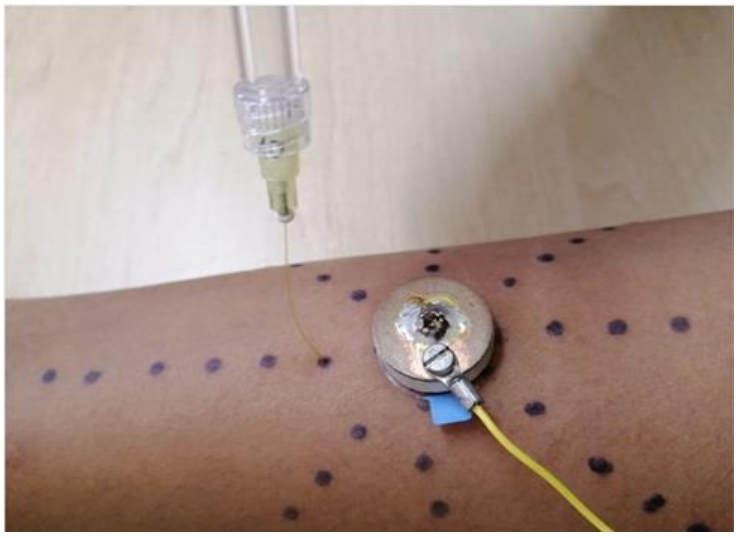


Figure 10*: Surface area estimation. The 128mN von Frey filament was applied starting at the most distal point and moving proximal to the electrode.*

Surface area assessment

The surface area of SH was estimated by applying the von Frey filament along eight radial lines arranged radially around the electrode in steps of 1 cm at intervals of ~1s, starting at the most distal points and moving towards the electrode (Figure 10). During this procedure, the participant was asked to keep their eyes shut and verbally indicate when they felt a distinct difference in sensation between a preceding stimulus (delivered further from the electrode) and the succeeding stimulus (closer to the electrode). The transition point was recorded as the border of SH for each radial line. We estimated the surface area of the eight triangles formed by the radial lines using the formula, surface area = ½ab (sin 45°), where a and b are the lengths of the sides of a triangle adjacent to the 45-degree angle. We used

the trapezoidal integration rule, represented by the '*area_under_curve*' function in R (Signorell), to estimate each participant's area under the line across the three time points.

Secondary outcome: magnitude of SH

The magnitude of SH was assessed using two blunt-ended metal "PINPRICK" rods (MRC systems, Heidelberg, Germany) that exert a force of 128 mN or 256 mN when applied parallel to gravity. Researcher 2 stimulated the area around the electrode thrice with each weight, with each stimulation lasting ~1 second. The stimulation was not repeated in the same area for the 3 stimuli to avoid potential sensitisation of cutaneous nociceptors by repeated stimulation of the skin. The participant was asked to give an average VAS rating for the three stimuli. These ratings were collected before (baseline test), and at 35, 50, and 65 minutes (follow-up tests) after the induction of SH (see Figure 11). Baseline tests represented the within-subject control time point for follow-up tests. We calculated the mean ratings for the two pinprick weights (128 mN and 256 mN) at baseline and for each follow-up time point. We expressed ratings from each follow-up point as a *change of the mean ratings* obtained at baseline, so each participant had three estimates of SH magnitude. In the protocol, we initially stated that follow-up ratings would be expressed as a *percentage of the mean ratings* obtained at baseline. However, we realised that this approach overestimates the magnitude of SH (see Protocol deviation: secondary analysis). We visually represented the change in the magnitude of SH across the three time points. We used the trapezoidal integration rule to estimate each participant's area under the line across the three time points.

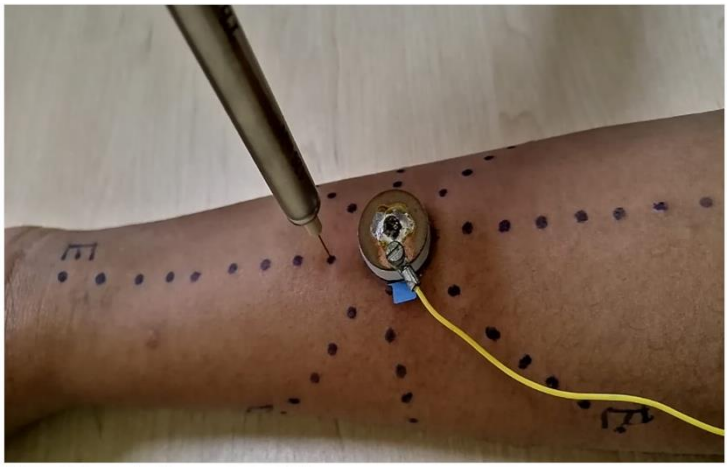


Figure 11*: Assessing the magnitude of SH.*

Exploratory outcomes

Manipulated variable: how painful was the SH induction?

We intended to manipulate skin sensitivity using HFS to induce SH. HFS is typically painful, and we wanted to gauge the painfulness of the induction. Participants received 5 HFS trains and rated their

painfulness on the electronic VAS. We report these results descriptively (see Results: Induction check).

Temporal summation, secondary allodynia, and primary hyperalgesia

Temporal summation was assessed prior to the induction. First, participants rated a single 256 mN pinprick stimulation delivered on a flat area on the forearm. Second, participants rated the 10th of ten 256 mN pinprick stimuli. All 11 stimuli were applied to exactly the same skin site. We also assessed three additional outcomes at 35, 45, and 55 minutes after the induction. Participants were asked to give a VAS rating of stimulation with (i) a 32 mN von Frey filament (VFF: Marstock nervtest, Germany), (ii) a light brush (Prime Art Bianco Flat 16), and (iii) a single electrical stimulus with the same settings as used for the HFS. The VFF and brush reflect static and dynamic secondary allodynia, respectively, which is pain due to a stimulus that does not usually provoke pain (Baumgärtner et al., 2002; IASP, 2021c), whereas the single electrical stimulus reflects primary hyperalgesia, which is enhanced pain sensitivity to a usually painful stimulus at the injured site (Treede, 2016). None of the outcomes listed here were used in the current study.

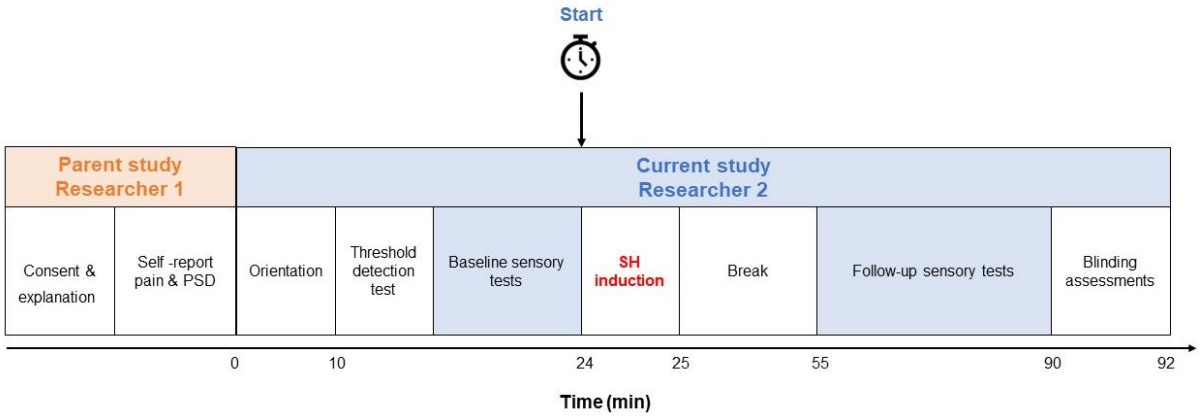


***Figure 12**: Study procedures. Note that the current study assessments were separated from parent study assessments by either a short break or up to 7 days. The orange box indicates assessments performed in the parent study by Researcher 1. The blue box indicates assessments performed in the current study by Researcher 2. Data collection in the current study was conducted in a separate session and room with Researcher 2, and all SH assessments were performed by the same researcher to avoid inter-observer variability (Meesters et al., 2020). Baseline sensory tests (before SH induction) served as the within-subject control time point for the follow-up sensory tests (after SH induction). The figure is not true to the time scale.*

Procedure

Consent and explanation of procedures

Researcher 1 assessed participants for eligibility and completed the informed consent procedures. Participants chose to communicate in English or isiXhosa. Researcher 1 explained the study information, and participants received an opportunity to ask questions before signing informed consent. The study information sheet and consent form are available in Appendix I. Consent procedures followed the explain-and-explain-back approach to confirm understanding and detect cognitive difficulties preventing participants from understanding the study procedures. None of the participants struggled to understand the study procedures. Participants who signed informed consent permitted the parent study to draw a blood sample and participate in the assessments listed below. Participants who completed the procedure were compensated R150[1] for their participation, and those who withdrew were compensated pro-rata. The participants were aware that participation was voluntary: the decision to participate in the present study had no relationship with their clinical care, and they were free to withdraw without negative consequences.

Self-report pain and PSD

Participants completed pain screening (using BPI) and PSD assessment (HSCL-25) with Researcher 1 (see Figure 12).

Orientation to assessments

The participant sat on a comfortable chair and placed the arm to be tested on the table while facing Researcher 2. The researcher followed a script to ensure consistency in all experimental procedures (Appendix J). For all procedures, the researcher and participants wore masks, and frequently used alcohol-based sanitiser on the skin. A perspex screen (dimensions: 800W X 600H) was placed on the table between the researcher and the participant to reduce airborne transmission of the SARS-CoV-2 virus. Researcher 2 briefly orientated the participant to the equipment (electrodes, stylus, etc.), assessments (e.g., baseline tests, HFS trains, etc.), and sensory modalities (128mN, 256mN, single electrical stimulus, brush and VFF), and the participants were given an opportunity to ask questions. The pinprick stimulators (128mN and 256mN) were cleaned with a sterilising fluid after being used on each participant. When the participants verbally confirmed their previously-given written consent, study assessments commenced. First, the participant's forearm was marked as in Figure 9, and the electrical current was individually calibrated for each participant (see Threshold detection). After this, Researcher 2 demonstrated sensory testing, first on her own alcohol-cleaned forearm and then on the participant's alcohol-cleaned forearm, to show the different sensations evoked by the different sensory modalities. Finally, the electronic touch VAS was explained to the participant, who received an

[1] Approximately US$9.71 at the time of data collection

opportunity to practice giving ratings on the touch-VAS for each sensory modality. All practice ratings were recorded but excluded from the analysis.

Threshold detection test

The threshold at which the participant could detect a single electrical stimulus was identified by delivering repeated electrical stimuli to the anterior side of the forearm used for testing, with the same electrodes as used for the HFS. Three steps were used to calibrate the electrical current for each participant. In step 1, the electrical current gradually increased, from 0mA, in steps of 0.10mA, until the participant reported first feeling the electrical stimulus. In step 2, the electrical current was decreased by steps of 0.5mA until the participant reported that they could no longer feel it. In step 3, the electrical current was increased by steps of 0.2mA until the participant could again feel it. The value obtained in step 3 was recorded as the threshold for the participant's detection of a single electrical stimulus, which was multiplied by 10 to yield the stimulation intensity used to induce SH and all single electrical stimuli during sensory testing (Klein et al., 2004).

Baseline sensory tests

The participant received three rounds of baseline testing before the induction of SH. The different sensory modalities were used in the following order: 128mN pinprick, 256mN pinprick, single electrical stimulus, brush, and VFF. The pinprick stimulations were presented first because the ratings of these modalities reflected the magnitude of SH (secondary outcome), thus making them of greater interest than the ratings of the other stimulus modalities. For the punctate mechanical stimuli (pinpricks and VFF), the participant was stimulated thrice, with each stimulation lasting ~1 second, for each rating. For the single electrical stimulus, the participant was stimulated once, with each stimulation lasting 2ms. For the brush, the participant was stimulated once, with each stimulation lasting ~1s. Every rating was recorded on the electronic VAS. The three rounds of baseline testing took approximately 6 minutes. Ratings were averaged across the three rounds for each modality. Finally, temporal summation was assessed once.

SH induction

The participant received five HFS trains, as described above (see Stimuli). As the first train was delivered, Researcher 2 started a stopwatch to ensure the accurate timing of follow-up assessments. As a safety precaution, the participant was informed that the bail-out word "stop" or "yeka" (isiXhosa) could be used if they needed to stop the trains and withdraw from the study. In addition, Researcher 2 placed her finger on the safety switch of the electrical simulator so that, if the participant used the bail-out word, Researcher 2 could immediately flick the switch to deactivate the stimulator. The participant was asked to rate each HFS train using the touch-VAS. The SH induction yielded one VAS rating for each of the 5 HFS trains from each participant.

Break

There was a 30-minute break between the SH induction and follow-up tests, to allow for the development of SH (Biurrun Manresa et al., 2018; Cayrol et al., 2018): the effect of the induction takes 20-30 mins to become apparent (Pfau et al., 2011; Quesada et al., 2021). During the break, participants were given magazines to view.

Follow-up sensory tests

Surface area estimation

The surface area of SH was assessed at 30, 45 and 60 minutes after the onset of the HFS trains (see figure 12). Therefore, we obtained three estimations of the surface area of SH, for each participant.

Sensory testing

The magnitude of SH was assessed at 35, 50, and 65 minutes after the onset of the HFS trains (see figure 12). Therefore, we obtained three rounds of follow-up testing for each sensory modality, for each participant.

Blinding

We developed strategies to support the blinding of Researcher 2 and the participants. Researcher 1 allocated group membership (based on pain status) and assessed the severity of PSD. This supported the blinding of Researcher 2 to pain status and PSD to minimise detection bias during the assessments of SH. During analysis, the group names ‘pain’ or ‘no pain’ were coded as ‘a’ and ‘b’ on the data files to which Researcher 2 had access so that the researcher was unaware of which group each participant belonged. To support the blinding of participants, we withheld information on the study aims and hypotheses, including the classification of participants into different groups to mitigate response bias.

We assessed unblinding for Researcher 2 and the participants at the end of the procedure. Researcher 2 answered two questions on the study computer. First, she guessed each participant's group membership in a forced choice between the two alternatives of 'pain' or 'no pain'. Second, she rated her confidence in her guess about group membership on a five-point Likert scale of "not at all confident", "not confident", "neutral", "confident", and "extremely confident". Next, Researcher 2 assessed for the unblinding of participants by asking them what they thought the purpose of the study was. Participants uncertain about the aim of the study were asked to guess, and Researcher 2 followed up on any indication that blinding may have been broken. Researcher 2 used conservative criteria to judge whether the participant remained blinded to the study aim and hypotheses (i.e., if uncertain, Researcher 2 judged blinding as broken, rather than intact).

Statistical analysis

Demographic data, PSD and pain status data were recorded on RedCap (https://redcap.uct.ac.za/). All data on SH surface area and magnitude were entered into and recorded by Affect5, directly into soft copy. Additional procedure notes were made manually and later transcribed into an Excel sheet. All data from RedCap, Affect5, and Excel were imported to R for analysis. We used the R 4.2.0 (The R Foundation for Statistical Computing) and R Studio (Integrated Development Environment for R) versions.

We posted and locked the study protocol (Appendix K) (on the 30 March 2022, updated; 06 April 2022) and pilot analysis script (30 March 2022) online, before formal analysis on Open Science Framework. Pilot and final analysis scripts used for data analysis are available as supplementary information in the appendices (Appendix L). The following packages were used for analysis: tidyverse (Wickham et al., 2019b), readxl (Wickham et al., 2019a), gridExtra (Auguie & Antonov, 2017), here (Müller, 2017), kableExtra (Zhu, 2019), ggstatsplot (Patil, 2021), pracma (Borchers, 2019), BI (Schwartz & Mercaldo, 2022), dplyr (Wickham et al., 2022), readr (Wickham, Hester & Bryan, 2022), arsenal (Heinzen et al., 2021), bayestestR (Makowski, Ben-Shachar & Lüdecke, 2019), DescTools (Signorell, 2022), rcompanion (Mangiafico, 2022), performance (Lüdecke et al., 2021), ggrepel (Slowikowski, 2021), formatR (Xie, 2022). Demographic data were reported descriptively and presented in tables. Before statistical tests were conducted, data were visualised to assess distribution assumptions and hypothesised relationships. All data were not normally distributed. Kendall's correlation test was used to test hypothesised relationships because our data were not normally distributed, and we had a small sample size. In general, a p-value of less than 0.05 was used as the statistical significance threshold.

Consideration of confounding variables

We anticipated that three variables could confound our results. First, the current used for the SH induction was calibrated for each participant, and could influence the area and/or magnitude of SH. Second, the painfulness of the HFS induction typically varies between individuals, and could influence area and/or magnitude of SH. For these two potential confounding variables, we tested for an association between the potential confounding variable and each SH outcome and planned to include the potential confounding variable as a covariate in the analyses of the SH outcomes if a significant association was found. Third, the delay between the assessment of self-reported PSD and the induction and assessment of SH (which affected only a few participants) could result in the PSD values poorly representing the participant's state at the time of the SH induction and assessment. Therefore, we wanted to assess the influence of days between PSD and SH assessments on the study results. For this, we assessed whether including the days between assessments as a covariate improved model fit in

either the primary or secondary analysis. Relationships between the confounders and each study outcome were formally tested using Kendall's test due to the distribution of our data and the small sample size (Rupinski & Dunlap, 1996).

Induction check

How painful was the induction?

We plotted and described the VAS (0 - 100) ratings for the five HFS trains, including exploring for group differences. These data were used to determine whether the induction was effective or not.

Primary and secondary analyses

The primary analysis used data on surface area, and the secondary analysis used data on magnitude of SH. For the surface area we plotted the (i) raw surface area data across the three time points. For the magnitude, we plotted the raw ratings across the four time points (baseline and three follow-up time points). Next, we visually represented the change in the magnitude of SH across the three time points. We used the area under the line of each SH outcome data in scatter-and-line plots to visualise the relationship between PSD and SH outcomes, first for the whole sample, and then by group. To establish whether the experimentally induced SH was associated with the severity of PSD, we (1) tested whether the PSD (independent variable) predicted each outcome (dependent variable) using the data from both groups. If this model indicated that each outcome was predicted by PSD, we planned to use the plots to guide decisions to potentially (2) test for an effect of group membership on each outcome by including group membership as a second independent variable, and (3) test for a difference in this relationship between PSD and group membership, by including an interaction between PSD and group in the model. However, we expected to interpret steps 2 and 3 as exploratory, since we would be underpowered, especially to statistically detect an interaction effect.

Assessment of model fit

Residual plots are used to validate assumptions about the regression model. Model assumptions are critical because when assumptions of a regression model are met, we can confirm that the model fits our data (Field, Miles & Field, 2012). We visualised model assumptions using the 'check model' function in R, which generates plots to check the following assumptions: posterior predictive check, linearity, homogeneity of variance, influential observations, normality of residuals and collinearity (where applicable). When an appropriate formula was available, we formally tested if the assumptions were met. Generally, if more than one assumption is violated, the result of the regression model may be unreliable (Field, Miles & Field, 2012). Below we explain each model assumption and how we determined whether the assumption was met.

Posterior predictive check

The posterior predictive check assesses how closely a simulation of the data, based on the fitted model, resembles the actual (observed) data (Gelman & Hill, 2006). Therefore, it identifies systematic discrepancies by plotting the model-predicted data with the observed data (Gelman et al., 2021). We checked if this assumption was met by generating a plot of the 'density' (y-axis) against the 'outcome of interest,' in our case, the surface area or magnitude of SH (x-axis) for each dataset (model-predicted vs observed). The assumption of the posterior predictive check is considered to have been met if the lines of the model-predicted data resemble the lines of the observed data.

Linearity

The assumption of linearity is that there is a linear relationship between the independent variable(s) and the dependent variable (Field, Miles & Field, 2012). We checked if this assumption was met by generating a scatterplot of the 'residuals' (y-axis) against the 'fitted values' (x-axis). A residual is a numeric difference between the model-fitted and observed values. If the relationship is linear, the reference line on the scatterplot should be straight and horizontal.

Homogeneity of variance

Homogeneity of variance means that at each level of the independent variable(s), the residuals have the same variance (homoscedasticity) (Field, Miles & Field, 2012). Heteroscedasticity (i.e., a violation of the assumption of homogeneity of variance) increases the variance of the regression coefficient estimates (beta-values). However, the model calculations need to account for this problem, which may result in some independent variables appearing to be statistically significant when they are not. In this way, the regression model becomes unreliable; the p-values are underestimated in the unstable model. We checked this assumption by generating a scatterplot of the 'square root of the standardised residuals' for the model (y-axis) against the 'fitted values' (x-axis). Homoscedasticity is deemed to be present if (i) the reference line on the scatterplot is straight and horizontal and (ii) there is no discernible pattern in the dots on the scatterplot (e.g., no branching, funnelling, or clustering). We formally checked for heteroscedasticity using the 'check heteroscedasticity' formula in R: a p-value less than 0.05 would indicate that our model violates the assumption of homoscedasticity.

Influential observations

Influential observations are outliers and/or data points that have high leverage (i.e., they exert undue influence over the model to 'pull' it in a certain direction) (Field, Miles & Field, 2012). Outliers have large residuals. High leverage typically occurs in observations with an unusual combination of independent variables. Observations with outliers and high leverage influence the regression model, and when these observations are removed, they change the estimates of the regression coefficients. We checked this assumption by plotting the 'standardised residuals' (y-axis) against the 'leverage' (x-

axis). The assumption of influential observations is met if the dots on the plot fall within the contour lines. We formally checked for influential observations using the 'check outliers' formula and Cook's distance method in R.

Normality of residuals

The normality of residuals means that the residuals of the model are random and normally distributed (Field, Miles & Field, 2012). The differences between the model-fitted and observed data should have a mean of approximately 0. We checked this assumption by generating a Q-Q plot of the 'sample quantiles' (y-axis) against the 'standard normal distribution quantiles '(x-axis). The assumption of normality of residuals was met if the dots on the plot roughly resembled the linear diagonal reference line. Data points above the diagonal reference line have a positive residual, whereas those below the diagonal reference line have a negative residual. We formally checked the normality of residuals using the 'check normality' formula in R: a p-value less than 0.05 indicated that our model does not meet this assumption.

Collinearity

Collinearity is assumed when one or more independent variable(s) in the regression model are strongly correlated (Field, Miles & Field, 2012). This causes the estimates of the regression coefficients to become unreliable. Collinearity is relevant to multiple regression because simple regression requires one independent variable. We checked this assumption by plotting the 'variance inflation factor' (y-axis) against the 'independent variable', in our case, PSD and group (x-axis). Three descriptors characterised collinearity: low (if the data point is less than 5), moderate (less than 10) and high (equal to 10). Therefore, collinearity was met if the variance inflation factor was low or moderate. We formally checked collinearity using the 'check collinearity' formula in R: output generated by the formula returns collinearity results for each independent variable included in the regression model.

Unplanned exploratory analyses

After seven participants withdrew from the study citing the painfulness of the induction procedure, we wished to ascertain whether this had biased our results. Therefore, we plotted PSD by completion status to shed initial light on the possibility that we had lost results from a particularly distressed subgroup.

Blinding analysis

We calculated and reported the percentage of participants who correctly guessed the aim of the study and the number of participants for whom Researcher 2 accurately guessed group membership. Due to a technical error, the 'neutral' option was missing from the assessor blinding assessment. Therefore, blind assessment was performed on a four-point Linkert scale of "not at all confident", "not

confident", "confident", and "extremely confident". Given the two-alternative forced choice design, 50% accuracy in Researcher 2's guesses about participants' group membership would represent chance accuracy, not blinding failure (55). However, no recommendations could be found to guide an acceptable threshold for the variance around this 50% value, outside which the data should be interpreted as indicating broken blinding. Therefore, instead of using a purely descriptive method to handle these data, we used two statistical methods to assess the blinding of Researcher 2 to group membership and planned to conduct sensitivity analyses in the event of broken blinding.

James' blinding index (BI) and the Chi-square goodness of fit test were used to assess blinding to group membership. The James' BI is a variation of the kappa coefficient, which is sensitive to the degree of 'disagreement' and not 'agreement' and prioritises 'do not know' responses (56). It is designed for use in studies that use a 3-alternative choice design including a 'don't know' option. However, our study lacked that option because we wanted to force Researcher 2 to give her best guess of each participant's group membership, rather than allowing a convenient 'don't know' response. Therefore, in our analysis, the number of 'don't know' responses were assigned a weight of 0. Accurate guesses were assigned a weight of 0, and inaccurate guesses were assigned a weight of 0.5, as previously described (Bang, Ni & Davis, 2004). If the BI is equal to 0, all guesses by Researcher 2 were accurate (Bang, Ni & Davis, 2004). If the BI is equal to 0.5, half of the guesses made by Researcher 2 were accurate, while half were inaccurate, which would mean random guessing (Bang, Ni & Davis, 2004). If the upper limit of the two-sided confidence interval is less than 0.5, then Researcher 2 was unblinded. Otherwise, we concluded that there was insufficient evidence to validate the blinding of Researcher 2 (Bang, Ni & Davis, 2004).

Given the limited match between our two-alternative forced choice design and the 3-alternative design anticipated by James' Blinding index, we supplemented that analysis with a Chi-square goodness of fit test. We used the Chi-square goodness of fit test to compare the observed distribution to the expected distribution by random chance (at 50% for each group) given by Researcher 2's guesses to participants group membership (Tezel et al., 2021). Together, these two tests would provide information on the breaking of blinding at the group level, but they shed no light on the potential influence on the study results if blinding were broken for a small number of participants. To test for this possible influence, we took a very conservative approach and identified participants for whom Researcher 2 has accurately guessed the group *and* also reported confidence of 4 (i.e., confident) or 5 (i.e. extremely confident), and conducted sensitivity analyses to examine the influence of cases in which blinding was possibly broken on the findings of the primary and secondary analyses. To achieve this, we repeated the primary and secondary analyses with data from these participants omitted.

RESULTS

Participants

Sixty-three participants identifying as male and female (17 males; 46 females) were enrolled in this study and screened for eligibility. Participants could be excluded before or after participation, according to the eligibility criteria. Three participants were excluded before the analysis because of data unavailability resulting from technical errors (n=2), and the procedure was aborted because the participant struggled to give ratings to the HFS trains in time (n=1). Two participants were regrettably included in the study despite reporting epilepsy but given that the epilepsy exclusion criterion was motivated by participant safety, rather than by anticipated effect on the findings, the data from these participants were retained. Two participants were excluded after participation because (i) the arm for SH induction had also been used for venepuncture without a 7-day washout period (n=1) and (ii) the participant was screened into the persistent pain group but was ineligible to participate because they had acute pain at baseline (as per the parent study timeline) (n=1). Seven participants withdrew from the study because the SH induction was too painful. We obtained complete datasets for 51 participants; however, blinding assessment data were lost for four participants due to technical errors. In the protocol we stated that we would recruit 60 participant (30 in each group); however, due to multiple errors of inclusion we end up with fifty-one participants (14 males, 37 females) aged 30 - 64 years (mean (SD) = 42.82 (8.7)) who completed the study. Here, we report only the data from these 51 participants (see Table 11), except for an exploratory analysis to check whether PSD differed between those who did and did not withdraw (see Results Unplanned exploratory analyses). All participants reported that they had not taken pain medication in the 24 hours prior to the induction of SH.

Table 11: Demographic characteristics of participants taken to analysis

	no pain (N=29)	pain (N=22)	Total (N=51)	p value
Sex				0.543
- female	22 (75.9%)	15 (68.2%)	37 (72.5%)	
- male	7 (24.1%)	7 (31.8%)	14 (27.5%)	
Age (years)				0.794
- mean	43.103	42.455	42.824	
- SD	9.116	8.245	8.671	
- median	41.000	42.000	42.000	
- range	31.000 - 64.000	30.000 - 58.000	30.000 - 64.000	
Calibrated current (mA)				0.509
- mean	0.140	0.128	0.135	
- SD	0.058	0.069	0.062	
- median	0.150	0.120	0.130	
- range	0.040 - 0.260	0.040 - 0.270	0.040 - 0.270	
PSD severity (scored using HSCL-25)				< 0.001
- mean	1.439	2.085	1.718	
- SD	0.447	0.628	0.618	
- median	1.360	2.100	1.480	
- range	1.000 - 2.520	1.080 - 3.360	1.000 - 3.360	

Induction check

Due to technical issues, almost half of the HFS train ratings were missing: 126 ratings were missing; 129 ratings were available. Ratings to HFS trains ranged between 0 and 100 on the VAS, and the median HFS rating was 45. There were no differences in HFS ratings between groups (Figure 13).

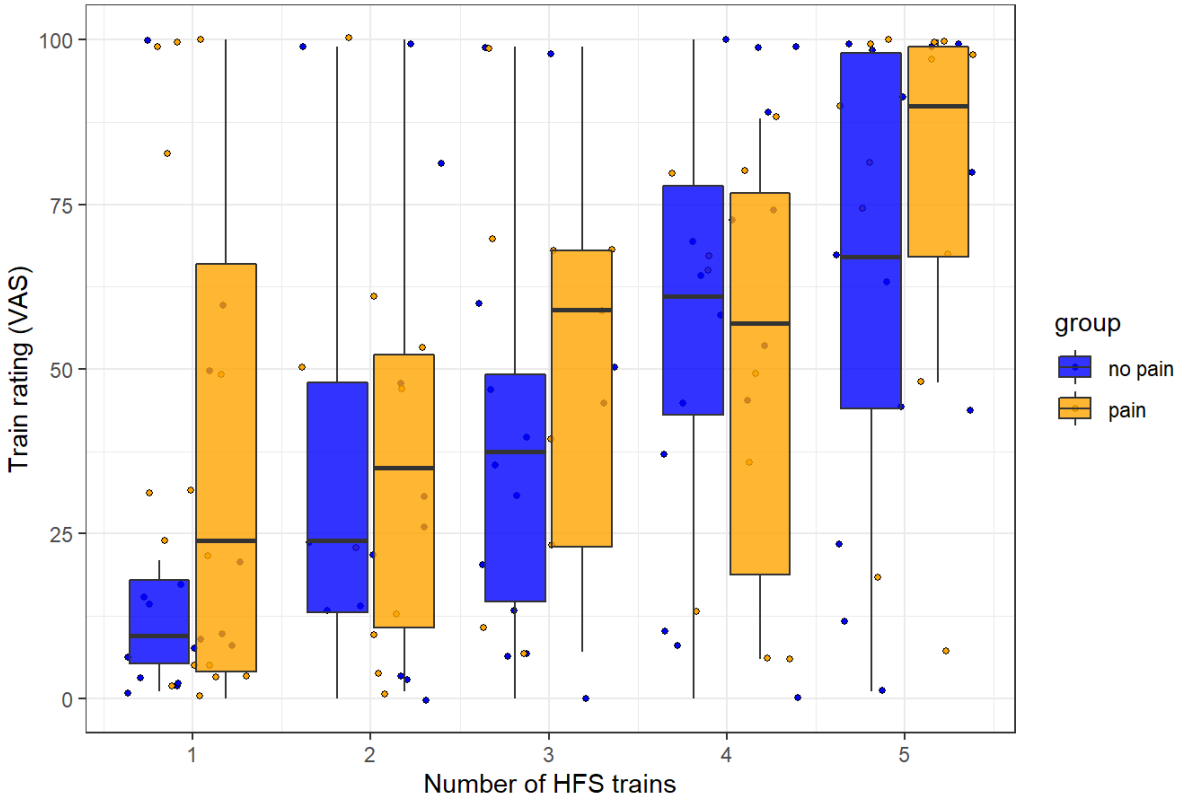


***Figure 13**: VAS ratings during each HFS train plotted by group (n=51)*

Blinding check assessment

Due to an unexpected power failure, blinding check data were not saved for one participant. Therefore, blinding data were available for 50 participants, all of whom (100%) were judged to be blinded to the aim of the study. Researcher 2 accurately guessed the group membership of 23 out of 50 participants. The James' BI was 0.55 (upper limit: 0.69) which confirms the blinding of Researcher 2 to participant group membership and that the researcher accurately guessed group membership for almost half of the participants meaning that guessing was random. This finding is also supported by the Chi-squared goodness of fit test conducted for both groups (X^2: no pain = 0.03; pain: X^2= 1.19). Researcher 2 accurately and confidently guessed group membership for five participants (three in the no pain group). Researcher 2 inaccurately but confidently guessed group membership for 11 participants (six in the no pain group).

Possible confounding

We anticipated three possible confounding variables: (i) calibrated current, (ii) painfulness of the HFS induction, and (iii) the number of days between PSD and SH assessments. The correlation tests showed no statistically significant relationship between either of the confounding variables and the surface area of experimentally induced (in (i, calibrated current) tau = -0.09, p = 0.36, and (ii, painfulness of the HFS induction) tau = 0.13, p = 0.24. Similarly, the correlation test showed no

relationship between either confounding variable and the magnitude of experimentally induced SH and (in (i, calibrated current) tau = -0.004, p = 0.97, and (ii, painfulness of the HFS induction) tau = 0.06, p= 0.60). Therefore, neither of these covariates was included in the analysis of the surface area and magnitude of SH. Finally, (iii, the number of days between PSD and SH assessments) was carried forward to the primary and secondary analyses as a potential covariate.

Primary and secondary analyses

Table 13 shows the descriptive statistics for SH outcomes. We used three models to formally test the relationships between the independent variables (PSD, group (i.e., pain status)) and dependent variables (SH outcomes). For each outcome, we chose one model (where applicable) that fits our data, and we used that model to test for the influence of (iii), the delay in days between PSD assessment and SH induction and assessment because we anticipated that the delay would influence relationships between PSD and SH outcomes. Finally, we provided the estimated marginal means and assessments of model fit for the chosen model. In all boxplots, black dots indicate individual participant scores, and horizontal lines indicate the 25th, 50th (median) and 75th percentile. Boxplot whiskers represent the maximum and minimum values.

Table 12: *Descriptive statistics for secondary hyperalgesia outcomes*

	No pain group (N=29)	Pain group (N=22)	Total (N=51)	p value
Surface area (AUL)				0.235
- number of participants	29	22	51	
- mean	894.948	1224.418	1037.072	
- SD	906.715	1047.365	973.798	
- median	450.622	1282.654	806.997	
- range	0.000 - 3010.039	0.000 - 3587.307	0.000 - 3587.307	
Magnitude (AUL)				0.702
- number of participants	29	22	51	
- mean	154.009	120.568	139.583	
- SD	281.739	337.860	304.423	
- median	37.500	18.750	30.000	
- range	-76.250 - 1132.500	-312.500 - 857.500	-312.500 - 1132.500	

Primary analysis

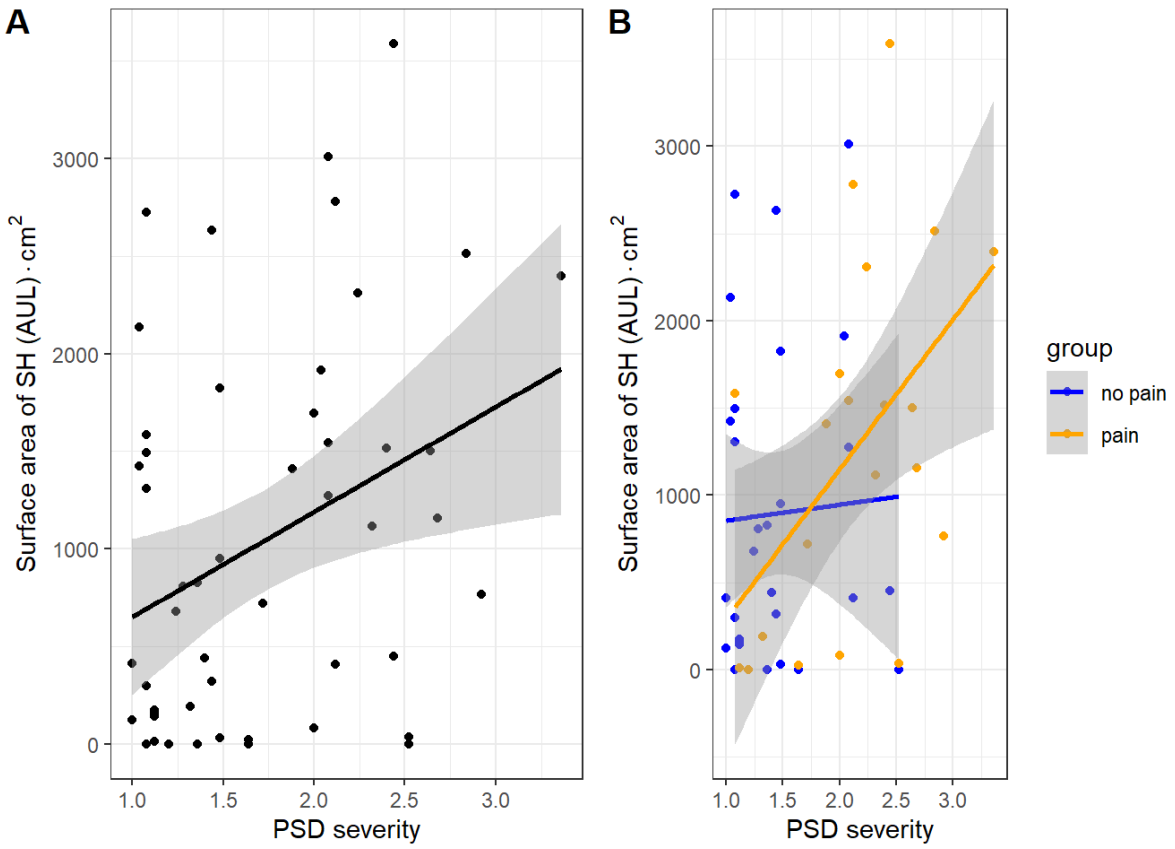


Figure 14*: LOESS curve showing the relationship between the surface area of SH and PSD severity, expressed as (A) total sample and (B) by group. Grey shadow represents 95% confidence intervals.*

Figure 14A suggests a positive relationship between the surface area and PSD. Formal modelling of the relationship between surface area and PSD (Model 1) confirmed that the surface area was indeed predicted by PSD ($F(1,49) = 6.49$, $p=0.01$). Model 1 predicted that a one-unit increase in PSD is associated with a 538.7 cm^2 increase in the surface area of SH, where PSD could range between one and three.

Next, we wanted to know whether this relationship was different between groups. Figure 14B shows the relationship between the surface area of SH and PSD between groups: pain group showed a stronger positive association between surface area and PSD than the no pain group. Formal modelling of the relationship between surface area, PSD and group (Model 2) confirmed that the surface area was still predicted by PSD ($p = 0.03$) but not group ($p = 0.93$), ($F(2,48) = 3.18$, $p=0.05$). The inclusion of 'group' as an independent variable in Model 2 did not improve the model fit (ANOVA comparing models 1 and 2 $Pr(>F) = 0.93$).

Given the suggestion of a between-group difference in the relationship between PSD and surface area suggested by Figure 14B, we formally tested for an interaction effect between PSD and group (Model 3). Model 3 confirmed that the surface area of SH was not predicted by PSD (p=0.82), group (p=0.15) or the interaction between PSD and group (p=0.13), (F(3,47) = 2.95, p=0.04). Including the interaction effect between PSD and group did not improve the model fit (ANOVA comparing models 1 and 3 p=0.32). Therefore, Model 1 was deemed the final model i.e., relationship between surface area and PSD.

Finally, we wanted to test if including the delay in days between PSD assessment and SH induction as a covariate improved the fit of model 1 in the current analysis (Model 4). Model 4 confirms that the surface area was still predicted by PSD (p = 0.01) but not by the delay between PSD and SH induction and assessment (p = 0.29), (F(2,48) = 3.83, p=0.03). The inclusion of the 'delay between PSD and SH induction' as an independent variable in Model 4 did not improve the model fit (ANOVA comparing models 1 and 4 Pr(>F) = 0.29). Therefore, Model 1 is still deemed the final model. Table 14 shows the estimated marginal means for the predicted effect of PSD on the surface area of SH when using Model 1.

***Table 13**: Estimated marginal means: predicted effect of PSD on the surface area of SH in Model 1.*

PSD	Predicted surface area	standard error	CI(lower bound)	CI(upper bound)
1.0	650.4603	199.4820	259.4827	1041.438
1.2	758.2046	169.5329	425.9262	1090.483
1.4	865.9489	145.8367	580.1141	1151.784
1.6	973.6932	131.8103	715.3498	1232.037
1.8	1081.4375	130.6070	825.4525	1337.423
2.0	1189.1818	142.5520	909.7851	1468.579
2.2	1296.9261	164.8108	973.9030	1619.949
2.4	1404.6704	193.8627	1024.7065	1784.634
2.6	1512.4147	227.1157	1067.2761	1957.553
2.8	1620.1590	262.9810	1104.7258	2135.592
3.0	1727.9033	300.5247	1138.8858	2316.921
3.2	1835.6476	339.1899	1170.8476	2500.448
3.4	1943.3919	378.6333	1201.2844	2685.500

In our sample, 76% of participants developed an identifiable surface area of SH (no pain: 83%; pain: 68%). The surface area of SH is shown by group (i) as an area under the line in Figure 15A and (ii) over time in Figure 15B. There was no clear difference in the surface area of SH between the groups.

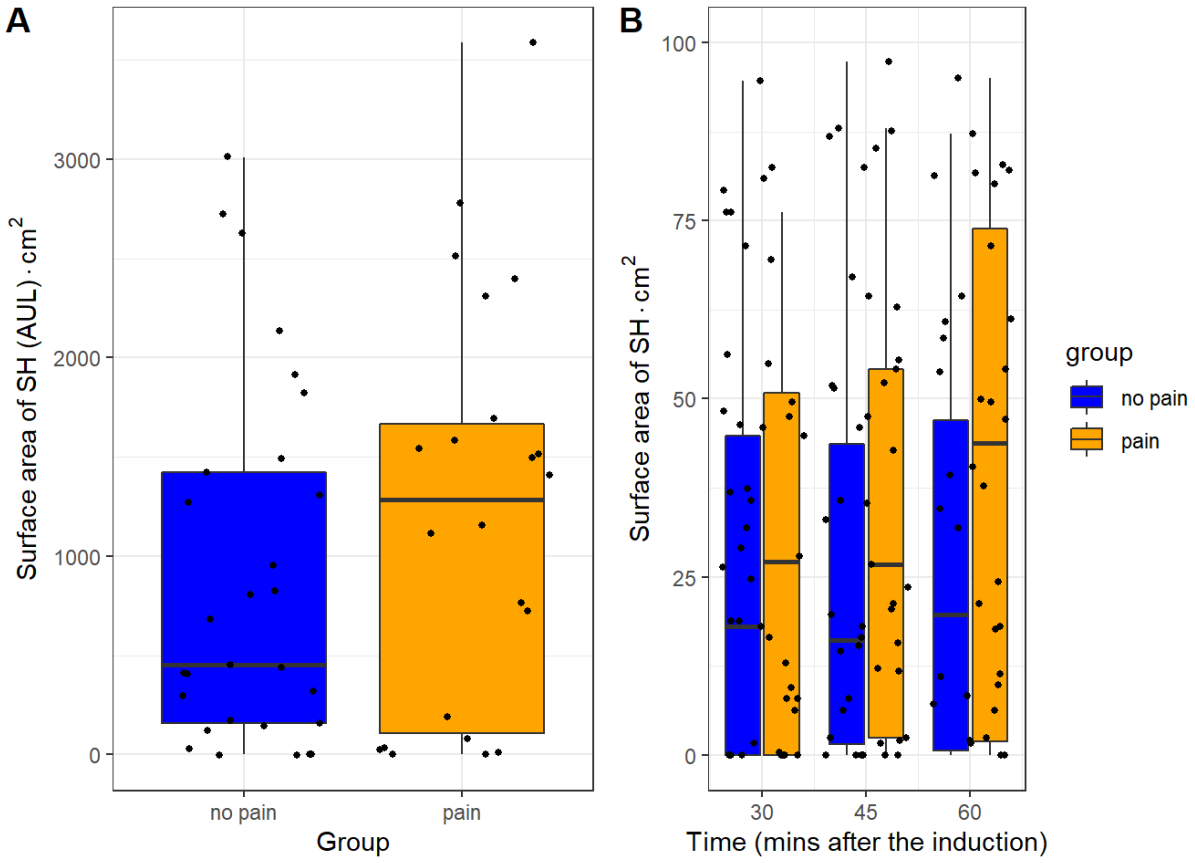


Figure 15*: Surface area of SH coloured by group and expressed as (A) area under the line and (B) timepoint after the induction.*

Assessment of model 1 fit for the primary analysis

Figure S1 (see supplementary information; top left) shows the posterior predictive check assumption, with a close resemblance between the model-fitted and observed data. Although the observed data formed a small cluster to the left, it was deemed acceptable for the sample size, indicating the assumption was met. In Figure S1 (top right), the scatter plot shows some deviation of the fitted values from the reference line, but it was considered acceptable for this sample, meeting the linearity assumption.

Regarding the homogeneity of variance assumption (Figure S1, middle left), the scatter plot exhibited acceptable deviation, lacking a discernible pattern. Formal testing confirmed meeting this assumption ($p = 0.46$). In terms of influential observations (Figure S1, middle right), no outliers or high leverage points were identified, and observations fell within the contour lines, satisfying the assumption (Cook's distance: 0.7). However, Figure S1 (bottom left) illustrates that the dots did not consistently

resemble the linear diagonal reference line, suggesting a potential violation of the normality of residuals assumption. Formal testing confirmed non-normality (p = 0.02).

Secondary analysis

Figure 17A suggests a small, positive relationship between the magnitude of SH and PSD. Formal modelling of the relationship between magnitude and PSD (Model 1) showed that the magnitude was not statistically predicted by PSD ($F(1,49) = 2.84$, $p=0.09$). Model 1 violated four (posterior predictive check, linearity, homogeneity of variance, and normality of residuals) of five assumptions, and the assumption of influential observations was deemed to have been met (see supplementary information; Figure S2).

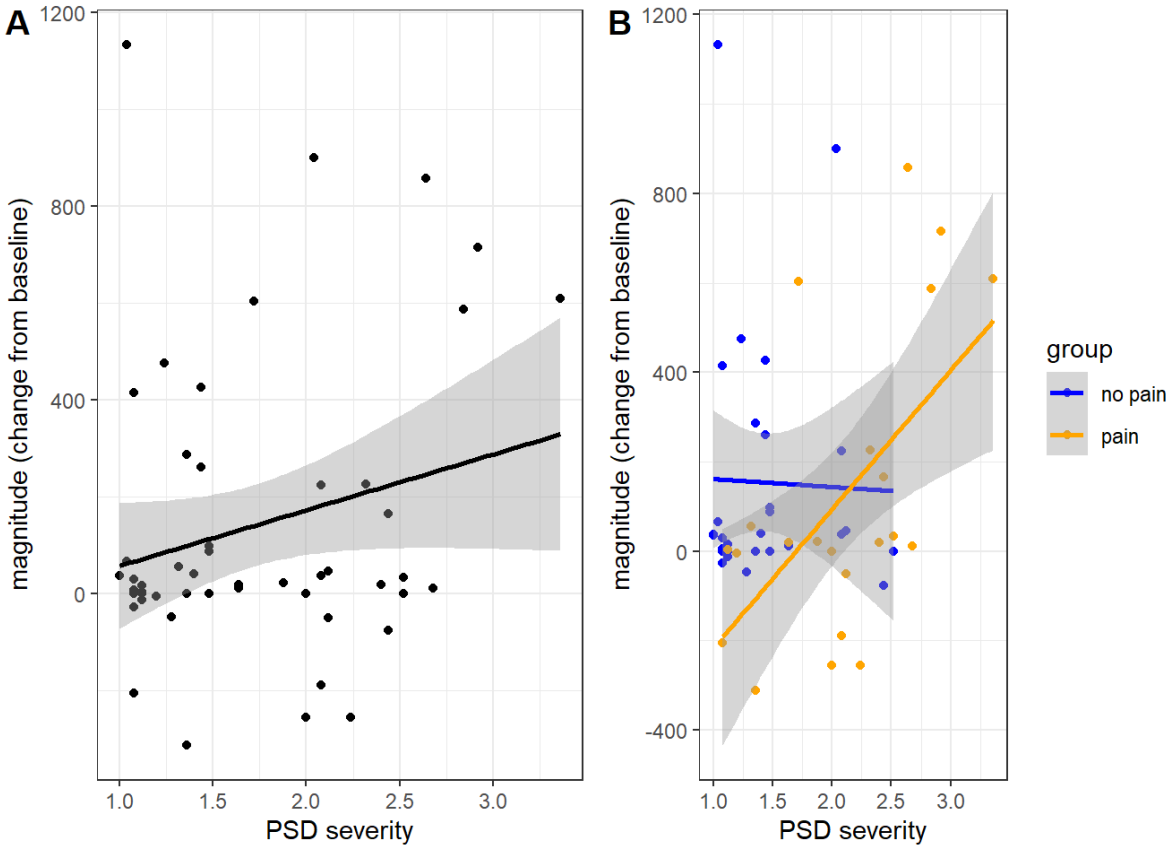


***Figure 16**: LOESS curve relationship between the magnitude of SH and PSD severity, expressed as (A) total sample and (B) by group. Grey shadow represents 95% confidence intervals.*

In Figure 19A, in the pain group, we noticed clustering of the magnitude data (see black oval). We investigated the possibility that the clustering of magnitude data was an artefact of our calculation process and found that the clustering was present in the raw data (see Figure 19C). No current modelling technique can adequately and appropriately account for the clustering effect observed in the pain group.

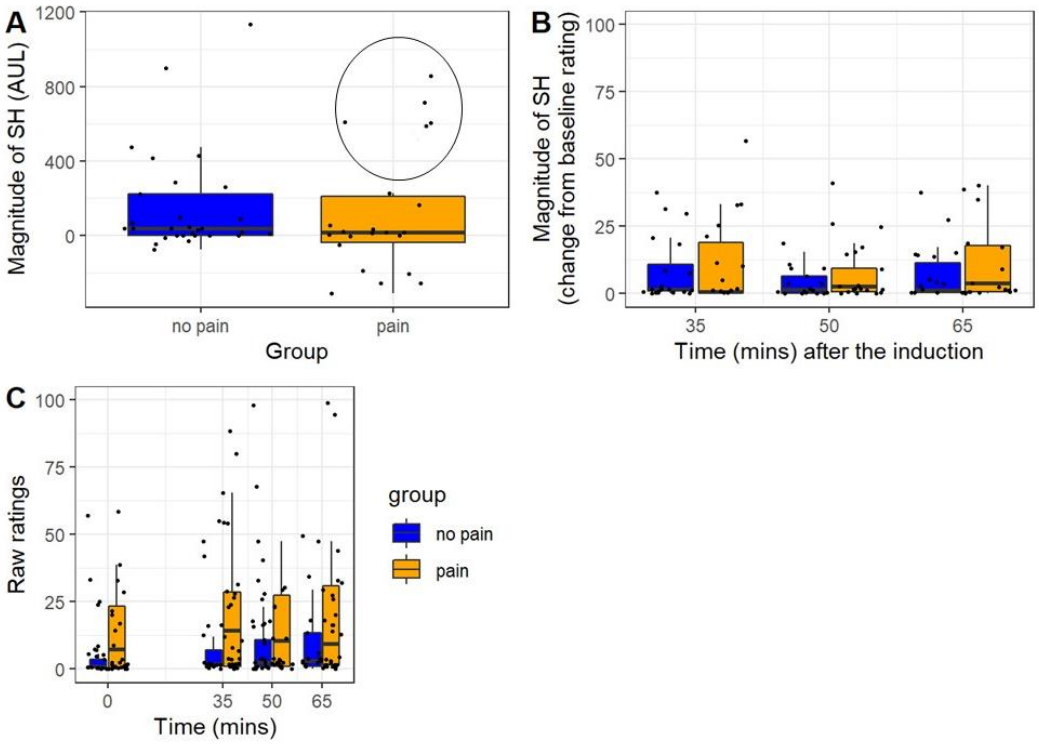


Figure 17*: Magnitude of SH coloured by group and expressed as (A) area under the line. The black oval shows the clustering effect observed, (B) change from baseline rating after the induction, (C) raw ratings.*

Initially, we planned to test the relationship between the magnitude of SH and (i) PSD and group (Model 2) and (ii) the interaction effect between PSD and group (Model 3) if Model 1 confirmed that PSD predicted the magnitude of SH. However, we opted to test Models 2 and 3 anyway because Figure 17B suggested a possible difference in the relationships between magnitude and PSD between groups: the pain group showed a stronger positive relationship between magnitude and PSD; however, this relationship was not observed in the no pain group. Given the learning priority of the current research project, we opted to proceed with testing Models 2 and 3 to investigate the pattern seen in Figure 17B, although we realised the model assumptions would likely be problematic.

Model 2 confirms that the magnitude was predicted by PSD ($p = 0.03$) but not group ($p = 0.14$), ($F(2,48) = 2.61$, $p=0.08$). The inclusion of 'group' as an independent variable in Model 2 did not improve the model fit (ANOVA comparing models 1 and 2 $Pr(>F) = 0.14$). Model 2 violated three (posterior predictive check, homogeneity of variance, and normality of residuals) of six assumptions. The assumptions of linearity, collinearity and influential observations were deemed to have been met (see Figure S3).

Model 3 reflected what was seen in Figure 17B: the magnitude was predicated by group (0.02) and the interaction between PSD and group (0.04) but not predicated by PSD (0.88), ($F(3,47) = 3.34$, $p=0.03$). Including the interaction effect between PSD and group improved the model fit (ANOVA comparing models 2 and 3 $p=0.04$). The estimated marginal means (see Table 15) for Model 3 support noteworthy differences in the relationship between the magnitude of SH and PSD between the groups. However, this is likely strongly influenced by the four values circled in Figure 19A. Unsurprisingly, Model 3 violated five (posterior predictive check, homogeneity of variance, linearity, collinearity and normality of residuals) of six assumptions, and the assumption of influential observations was deemed to have been met (see Figure S4). Overall, it was impossible to generate an acceptable statistical model of the relationships between magnitude and (i) PSD and (ii) group for these data.

***Table 14**: Estimated marginal means: predicted effect of PSD on the magnitude of SH in Model 3*

Group	PSD severity	Predicted magnitude	CI(lower bound)	CI(upper bound)
no pain	1.0	161.76481	15.13773	308.39189
pain	1.0	-216.62713	-458.80683	25.55257
no pain	1.2	158.22818	40.15175	276.30462
pain	1.2	-154.49734	-363.72509	54.73040
no pain	1.4	154.69156	50.53658	258.84653
pain	1.4	-92.36755	-271.03023	86.29513
no pain	1.6	151.15493	40.61922	261.69064
pain	1.6	-30.23776	-182.16972	121.69419
no pain	1.8	147.61830	13.26169	281.97492
pain	1.8	31.89203	-99.50489	163.28894
no pain	2.0	144.08168	-24.28972	312.45307
pain	2.0	94.02182	-26.25220	214.29584
no pain	2.2	140.54505	-67.08400	348.17411
pain	2.2	156.15161	34.96830	277.33491
no pain	2.4	137.00843	-112.66014	386.67699
pain	2.4	218.28140	84.40156	352.16123
no pain	2.6	133.47180	-159.82434	426.76795
pain	2.6	280.41119	124.90849	435.91388
no pain	2.8	129.93518	-207.96205	467.83241
pain	2.8	342.54098	159.62780	525.45415
no pain	3.0	126.39855	-256.73343	509.53054
pain	3.0	404.67077	190.77305	618.56848
no pain	3.2	122.86193	-305.93799	551.66185
pain	3.2	466.80056	219.68500	713.91611
no pain	3.4	119.32530	-355.45074	594.10135
pain	3.4	528.93035	247.15237	810.70832

.

In our sample, 41% of participants developed an identifiable magnitude of SH as indicated by a change in ratings to pinprick stimulation (no pain group: 26%; pain group: 59%). The magnitude of SH is shown by group (i) as an area under the line in Figure 17A and (ii) over time in Figure 17B. There was no clear difference in the magnitude of SH between the groups.

Unplanned exploratory analyses

Figure 22 illustrates PSD severity plotted by completion status. PSD severity ranged between 1.12 to 2 for those who withdrew. Therefore, the withdrawal of these participants did not bias our results.

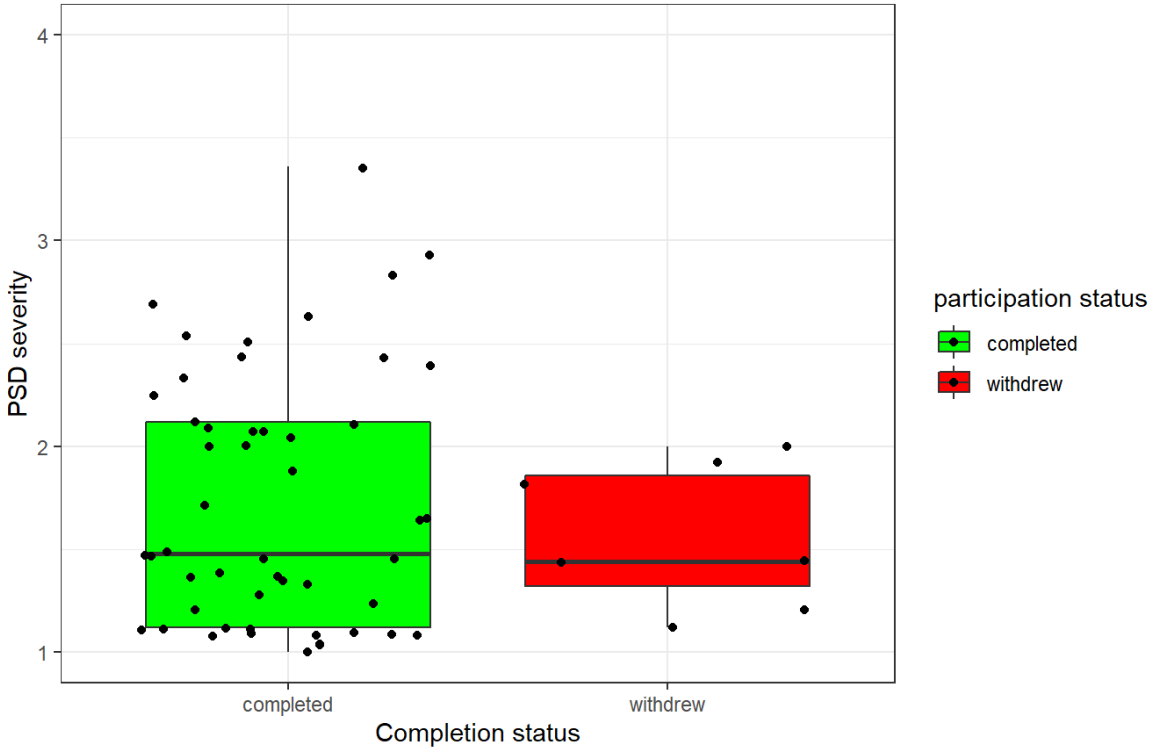


***Figure 18**: PSD severity from participants who completed the study (n=51) and those who withdrew (n=7) plotted by group.*

DISCUSSION

We aimed to investigate the relationship between the severity of self-reported PSD and SH (surface area and magnitude) in people with HIV who reported persistent pain or no pain. Our primary hypothesis was that we would observe a positive relationship between PSD and SH outcomes in both groups. Our secondary hypothesis was that the persistent pain group would have a greater surface area and magnitude of experimentally induced SH than the no pain group. To achieve our aim, we collected PSD severity, manipulated skin sensitivity using HFS and assessed SH outcomes in both groups. The induction check confirmed that the induction was comparable across groups. Participants were blinded

to the study aims, and Researcher 2 was blinded to the pain status of each participant. Our primary hypothesis was only partially upheld in that PSD severity was positively associated with the surface area of experimentally induced SH but not with the magnitude of SH. Similarly, our secondary hypothesis was partially upheld in that participants with pain developed a significantly *greater* magnitude of SH than those without pain. In contrast, those with pain developed a significantly *smaller* surface area of SH than those without pain.

Relationship between PSD and surface area of SH

In this study, participants who reported pain developed SH over a smaller *surface area* than those without pain. We speculate that this could be of indicate more efficient descending inhibition as a result of 'training' when coping with clinical pain. Evidence suggests that PWH have demonstrated resilience despite challenges, although this is not specific to pain-related resilience (Dale et al., 2014; Emlet et al., 2017). A study focusing on pain-specific resilience found that individuals with higher levels of resilience displayed reduced sensitivity to experimental pain stimuli during quantitative sensory testing (Slepian et al., 2016). Resilient individuals also tend to employ adaptive pain coping strategies, such as pain acceptance (McCracken, 1998), which contribute to their ability to effectively cope with pain (Sturgeon & Zautra, 2010). Additionally, adaptive strategies like distraction and positive coping self-statements have been shown to enhance adaptation to chronic pain and reduce sensitivity to experimental pain stimuli (Malloy & Milling, 2010; Roditi, Robinson & Litwins, 2009; Verhoeven et al., 2011). It is plausible to hypothesise that individuals with effective pain-specific coping strategies may possess more efficient descending inhibition, potentially limiting the spread of SH—which could explain why our participants who reported pain developed SH over a smaller surface area than those without pain. However, this speculation contrasts with existing literature reporting *less* efficient descending inhibition in people with pain (de Tommaso et al., 2011; Smith et al., 2008), although those studies did not consider potential moderation by coping. We could not investigate this directly in the current study because we did not have data on our participants' coping or resilience, but it would be helpful if future studies did collect such data, so as to investigate whether pain-related coping or resilience has an influence over experimentally induced SH.

Relationship between PSD and magnitude of SH

In our secondary hypothesis, we hypothesised that those reporting pain would report a greater SH. Our findings align with this: the proportion of participants reporting pain had a more pronounced *magnitude* of SH compared to those without pain. This finding suggests that individuals with pre-existing hyper-responsive systems may be more susceptible to SH on induction. In other words, clinical pain and heightened responsiveness to SH induction may be 'parallel consequences' of a hypersensitive system. In contrast, we did not find a direct correlation between reporting pain and the *surface area* of SH. However, it is possible that individuals with more widespread pain or multiple

pain locations may exhibit a larger surface area of SH, while those with higher pain intensity may demonstrate a greater magnitude of SH. Investigating these relationships is part of the ongoing parent study, as our current group classification method did not allow us to explore this specific question using the available data.

LTP-like changes have been proposed to underlie experimentally induced SH (Klein et al., 2004). However, evidence suggests that NMDA receptors are implicated in long-term potentiation in the spinal cord (Chen et al., 2014; Ikeda, Asai & Murase, 2000; Randic, Jiang & Cerne, 1993; Zhou et al., 2012). The current systematic review indicates that NMDA antagonists have no significant effect on SH in the experimental models studied raises questions about the "LTP-like" nature of the underlying physiology of SH in those particular experimental models. Specifically, it is possible that the different experimental models may replicate a short-lived version of the same physiological processes observed in clinical SH. However, further investigation is needed to fully understand the neurophysiological mechanisms of SH and its relationship to 'LTP-like' processes.

Short-lived SH

Our participants developed a short-lived SH lasting a few minutes. Other studies have developed a long-lasting SH using similar HFS protocols. Below we propose several reasons for the short-lived SH.

Reason 1: electrode configuration and distance between cathode and anode

Other studies have produced robust and long-lasting SH (~ one to five hours) in healthy participants using high frequencies (100 Hz), as used in the current study (Broeke & Mouraux, 2014; Klein, Magerl & Treede, 2006; Klein et al., 2004; van den Broeke et al., 2019; Vo & Drummond, 2013). A study using an HFS protocol closely resembling ours showed evidence of SH in healthy people (Cayrol et al., 2018). The main difference between that study and ours is the (i) electrode configuration and (ii) distance between the cathode and anode. For example, previous studies used an anode which was a stainless-steel reference electrode located inside the cathode, to achieve spatial summation within the receptive field of the spinal cord neurons. In contrast, the electrode configuration we used was a cathode with ten steel pins, and the anode was a fabric electrode strapped around the participant's arm resulting in a greater distance between the anode and cathode. Therefore, the difference in the electrode configuration, including the distance between the cathode and anode used in our study, may have resulted in the short-lived SH.

Reason 2: different time to effect of induction in the current population

The initial SH assessment was conducted 30 minutes after the induction. We chose this time point based on evidence that the effect of the HFS induction is apparent between 20 to 30 minutes after the induction (Pfau et al., 2011; Quesada et al., 2021). Notably, this evidence was based on studies in

healthy participants without pain. To the best of our knowledge, the current HFS protocol has not been used in PWH reporting pain or no pain. Therefore, it is plausible that the duration of SH development may differ in this population, leading to a difference in time to effect. This could explain the short-lived SH observed in both groups; however, we could not confirm this because we only assessed SH 30 minutes after induction.

Reason 3: descending inhibitory pain pathway limits spinal plasticity

Evidence suggests that the descending pain modulation pathway can limit spinal plasticity. This type of inhibition arises from the brain and descends to the spinal cord to influence pain perception by reducing the transmission of nociceptive inputs (Kwon et al., 2014). Animal studies on pain show that noxious stimuli inhibit the response of neurons in the spinal cord after receiving noxious and innocuous stimuli to electrical stimulation and noxious heat (Le Bars, Dickenson & Besson, 1979). This inhibition is induced by pain and modulated by the descending inhibitory pathway by inhibiting neurons in the dorsal horn that were not involved in the initial nociceptive stimulus (Ossipov, Dussor & Porreca, 2010). Consequently, this prevents the nociceptive signal from spreading. A previous study that used HFS to induce SH in healthy participants investigated the link between painful mechanical stimuli on the forearm and forehead analgesia (Vo & Drummond, 2013). This study found that HFS concurrently induced SH and pain inhibition. Therefore, descending pain inhibition may limit spinal plasticity, thereby affecting the development of SH, which could explain the short-lived SH observed in both groups.

Reason 4: lack of randomisation of sensory modalities, the quantity of pinpricks and frequency of pinprick stimulation

During orientation, we introduced five sensory modalities to participants. We told participants that each modality evoked different sensations and that every time they rated a new modality, they shouldn't compare it to the previous modality. A previous study randomised modalities within each time point to decrease the predictability of modalities and support the collection of accurate ratings (Bedwell et al., 2022). This study differs from ours in that we consistently presented the pinpricks first because these modalities were relevant to the SH outcomes. Further, the modalities were given in a fixed order (128mN, 256mN, single electrical stimulus, brush and VFF) within each timepoint, which might have made it easy for the participant to predict which modality they were being stimulated with. Therefore, the fixed order of modalities could explain the short-lived SH across our sample.

Another study using an HFS protocol similar to ours developed a clear and apparent magnitude of SH in healthy participants (Klein et al., 2007). The obvious difference between this study and ours is that they recruited pain-free participants, whereas our participants live with HIV and reported pain or no pain. Other differences include (i) the quantity of pinprick weights and (ii) the frequency of stimulation with each pinprick around the induction site when assessing SH. This study assessed the

magnitude of SH using seven different pinprick weights (with forces of 8, 16, 32, 64, 128, 256 and 512mN), which were stimulated *once* around the induction site, and a rating was given each time. In contrast, we assessed SH using two pinprick weights (128mN and 256mN), which were stimulated *thrice* around the induction site and an average rating of the three stimuli were given. We did not repeat the stimulation in the same area for the three stimuli because we wanted to avoid the potential sensitisation of cutaneous nociceptors by repeated stimulation of the skin. When using pinpricks, stimuli should be delivered rapidly and briefly for adequate synchronous activation of A delta-fibres (Le Bars, Gozariu & Cadden, 2001). We may have failed to generate sufficient activation of A-delta fibres because of the technique we used to assess SH which could explain the apparent short-lived SH.

Reason 5: influence of attention on sensory testing

Ratings of quantitative sensory testing can be influenced by errors from decreased attention by the participant (Shy et al., 2003). In the current study, we identified two participants (PID 202 and 544) for whom Researcher 2 reported that loud music was playing from the houses near the research site. A recent study reported that spatial attention to one forearm could modify the development of SH in healthy participants (Filbrich et al., 2020). This study showed that SH was larger on the forearm where attention was directed towards than the arm where attention was diverted away. This evidence supports our findings on participant 202 who developed a poor magnitude of SH, which was only apparent 65 minutes after the induction. However, a recent replication study showed that spatial attention might not affect the development of SH (Della Porta et al., 2022). Surprisingly, this evidence also supports our findings on participant 544 who developed a clear magnitude of SH, which was apparent at all time points. Evidence of the influence of attention on sensory testing is conflicting; however, some evidence presented here could explain the poor development of SH in some participants.

SH outcomes mediated by different mechanisms

It is plausible that different mechanisms mediate SH outcomes. First, the development and expression of SH differed between groups. Second, we plotted the SH outcomes for each time point after the induction by each group to check if the peak of the surface area corresponded temporally with the peak of the magnitude of SH (see Appendix L, formal analysis, Unplanned analysis 1). Our plot suggests that the surface area and magnitude of the SH peak did not consistently correspond temporally in both groups. Third, eight participants displayed hypoalgesia, reduced pain in response to a normally painful stimulus (IASP, 2021j) (Figure 23). Three of (the 8 participants) were in the no pain group (232, 481, and 591), and five of (the 8 participants) were in the pain group (156, 201, 457, 561 and 563). Notably, of the eight participants who developed hypoalgesia, seven developed an apparent surface area of SH, which was relatively small in the participants of this subset who reported no pain but large in those that reported pain. These findings suggest that the degree of spread of neural

activity to nearby fibres (surface area of SH) is unrelated to the degree of strengthening of synapses (magnitude of SH) following HFS induction. This difference in SH outcomes suggests that different mechanisms likely contribute to developing the magnitude and surface area of experimentally induced SH, which supports findings from Chapter 2 (see Discussion).

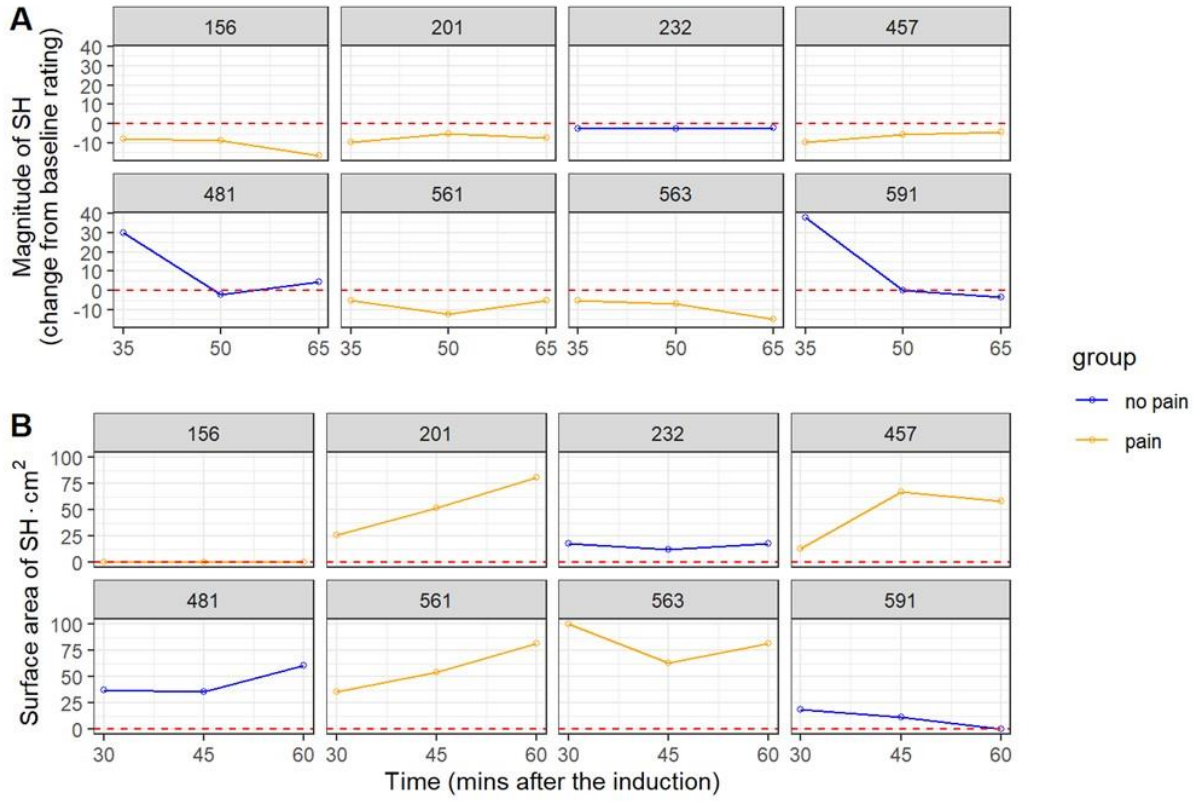


***Figure 19**: Eight participants experienced hypoalgesia. Red dotted lines indicate no hyperalgesia.*

Limitations

Participants in the pain group were included if they reported persistent pain (pain on most days for more than three months) based on pain frequency and duration, with no regard to the severity, location, or presumed cause of their pain. In two participants, the pain reported during T0 when they were recruited to the parent study was inconsistent with the group assigned during screening. When participants came for screening, they only completed three questions on pain screening. In contrast, during T0 assessments, participants were asked to complete the full BPI self-report, which included questions on pain severity and interference. Therefore, for the two participants we identified conflicting pain reports, we used the group reported at T0 because we thought it was more accurate (the full BPI) than the group reported at screening, using three questions. This is a potential source of response bias, negatively impacting the data interpretation.

In the protocol, we stated that follow-up ratings of pinprick stimuli would be expressed as a percentage of the mean pinprick ratings obtained at baseline. However, we realised that this approach yields different estimates of the magnitude of SH. Table 16 shows a hypothetical scenario with baseline and follow-up ratings from three healthy participants who would have received HFS to induce SH. The change in the ratings for Participant 1 is the same as the change in rating for Participant 2, if we assume that the change in rating represents a consistent amount of change regardless of the range of the scale in which the change is reported. Participant 3 showed about a third of the change in Participants 1 and 2. Applying commonly used methods that divide follow-up ratings by baseline ratings (methods 1 and 2) yields very different values for the magnitude of SH for Participants 1 and 2. In contrast, subtracting the baseline from follow-up ratings (method 3) yielded the same magnitude of SH for Participants 1 and 2. Therefore used decided to use method 3 to estimate the magnitude of SH.

Table 15: *Hypothetical scenario with baseline and follow-up ratings and various methods of calculating the magnitude of SH assessed using VAS*

Healthy participants	**Pinprick ratings**		**Various methods of calculating % change**		
	Baseline (B)	**Follow-up (FU)**	**Method 1 (FU-B)/B***	**Method 2 (FU/B)***	**Method 3 (FU-B)**
Participant 1	0	30	301	301	30
Participant 2	5	35	6	7	30
Participant 3	0	10	100	101	10

*Conventionally, 0.1 is added to all raw ratings in this method to avoid losing zero values during the calculations.

Conclusion

Our experimental study showed a positve relationship between PSD and the surface area of SH in individuals living with HIV, regardless of pain status. These findings suggest that interventions aimed at reducing PSD may hold therapeutic potential in addressing SH. It is important for future studies to distinguish between experimental SH and clinical SH, as existing literature has shown promising results for distress-reducing therapies in reducing experimental SH (You et al., 2014). While our study focused on experimental SH, further exploration of the effects of distress-reducing therapy on clinical SH would be valuable.

CHAPTER 4: CONCLUSION

Research objectives

The aim of this research project was to understand the mechanisms of SH. We thought that research on manipulating SH will give us insight into the mechanisms of SH. To achieve our aim, we conducted two studies, (i) a systematic review according to PRISMA guidelines on the existing evidence that targeting the NMDA system alters experimentally induced SH, in healthy human participants (*Chapter 2*) and (ii) an experimental study to investigate whether distress predicted CS in PWH (*Chapter 3*). In Chapter 2, we investigated whether using NMDA antagonists modulates experimentally induced SH. This research not only clarifies the efficacy of NMDA antagonists but provides a better understanding of the mechanisms underlying SH. In Chapter 3, we investigated whether PSD was associated with SH outcomes (*primary hypothesis*) and if those with pain would report a greater SH than those without pain (*secondary hypothesis*). In this chapter, we will summarise the main findings for Chapters 1 and 2, including the strengths and limitations of this research project. Lastly, we will provide recommendations for future research.

Main findings

Systematic review

The effect of NMDA antagonists on experimentally induced SH was assessed by change in magnitude of SH in 13 (of 52) studies and assessed by the surface area of SH in 37 (of 52) studies. Nine (of 52) assessed both the change in magnitude and surface area of SH in response to NMDA antagonist administration. Here, findings were described narratively, and a meta-analysis was conducted for each SH outcome.

In the narrative summary, twelve studies manipulated the magnitude of SH using ketamine alone (n=10) or combined with alfentanil (n=2). Seven (of 10) studies reported no effect, and three found that ketamine reduced the magnitude of SH. Coadministration of ketamine with alfentanil had no effect on the magnitude of SH. Together, these studies suggest that antagonising the NMDA system does not consistently reduce the magnitude of experimentally induced SH. Thirty-five studies manipulated the surface area of SH using ketamine alone (n=31) or combined with alfentanil (n=2), morphine (n=1) or remifentanil (n=1). Of those using ketamine alone, fifteen (of 31) reported no effect, and sixteen (of 31) found that ketamine reduced the surface area of SH. Coadministration of ketamine with alfentanil or morphine had no effect on the surface area of SH. In contrast, the coadministration of ketamine and remifentanil reduced the surface area of SH. Ten studies manipulated the surface area of SH using dextromethorphan alone (n=8) or combined with morphine (n=2). Four of those using dextromethorphan alone reported no effect, and the rest reported a decrease in the surface area of SH. Coadministration of dextromethorphan with morphine had no effect (n=1) or

decreased (n=1) the surface area of SH. Studies that used CH3381 (n=1) or neramexane (n=1) reported a decrease in the surface area of SH. One study that used magnesium sulphate found no effect of magnesium on the surface area of SH. Together, these studies suggest that antagonising the NMDA system does not consistently reduce the surface area of experimentally induced SH.

In the meta-analysis, we included forty studies (seven of which assessed the magnitude of SH). These pooled data suggested that ketamine had no effect on the magnitude of SH (n=7). Similarly, ketamine (n=19),CHF3381 (n=2) and dextromethorphan (10 studies) had no effect on the surface area of SH. In contrast, multiple-dose ketamine reduced the surface area of SH (n=1). Further, the coadministration of ketamine with remifentanil reduced the surface area of SH (n =1).

Our findings suggest that NMDA antagonists had no significant effect on modulating the mechanisms underlying SH. This implies that the involvement of NMDA receptors in SH mechanisms may be limited or secondary to other factors. It is important to consider alternative pathways or neurochemical systems that may play a more prominent role in the development and modulation of SH. Further research is needed to explore these alternative mechanisms and identify potential targets for interventions aimed at reducing SH.

Experimental study

In the experimental study, our primary hypothesis that PSD severity would be associated with SH outcomes was partially upheld. PSD severity was associated with the surface area but not the magnitude of SH. However, the association between the surface area of SH and PSD was not different between those who reported pain and those without pain, suggesting that pain status did not predict the surface area of SH. Our plots suggested an association between PSD and the magnitude of SH, which seemed stronger in those with pain than those without. However, we couldn't confirm this because we couldn't find a suitable statistical model.

Similarly, our secondary hypothesis was partially upheld in that those with pain developed a significantly greater magnitude of SH than those without pain. Further, those with pain developed a smaller surface area of SH than those without pain. We think that the increase in pain-specific resilience in those who reported pain could possibly explain why they developed a smaller surface area of SH compared to those without pain. Interestingly, we identified eight participants who developed hypoalgesia, of which seven developed an apparent surface area of SH. In fact, the development and expression of SH differed between groups. These findings support the idea proposed in Chapter 2 that different mechanisms underlie the development of each SH outcome.

The participants in our sample developed a short-lived SH, lasting for a matter of minutes. Five reasons were proposed to explain the short-lived SH observed in this sample: (i) electrode configuration and distance between cathode and anode; (ii) different time to effect of induction in the current population; (iii) descending inhibitory pain pathway which limits spinal plasticity; (iv) lack of

randomisation of sensory modalities, including the quantity of pinpricks and frequency of pinprick stimulation; and (v) the influence of attention on sensory testing. Therefore, future research should consider standardising these factors when using HFS to induce SH.

Strengths

Systematic review

To the best of our knowledge, no systematic review has investigated the efficacy of NMDA antagonists in modulating experimentally induced SH. This review is useful for understanding the mechanisms of SH and how NMDA antagonists can modulate SH. The research question of this review was clear to ensure accurate inclusion of records. Our search strategy was comprehensive, and records were included based on pre-defined eligibility criteria. Data analysis was guided by the Cochrane Handbook of Systematic Reviews for Interventions to ensure that the methods used for the meta-analysis were robust.

Experimental study

This is the first study investigating the relationships between PSD and SH outcomes in PWH in South Africa. South Africa has a high prevalence of HIV, and studying the relationships between PSD and SH in this population provides insights into the experiences and challenges faced by PWH in South Africa. Additionally, investigating PSD and SH contributes to the existing body of research conducted in different geographical regions and different human populations. Our findings can help identify potential variations or similarities in the relationships between PSD and SH across diverse populations. Further, this experimental study comes with several methodological strengths. First, we developed strategies to blind the participants, researcher performing the SH induction and assessment, and data analyst. In addition, we included a blinding assessment to minimise the risk of performance and detection bias. Second, we used a validated technique and reporting scale to assess SH. The researcher performing the SH induction and assessment received training prior to data collection to support accurate data collection.

Limitations

Systematic review

The systematic review encountered limitations associated with methodological heterogeneity, missing data, unreported data requiring extraction of estimates from plots, low female inclusion rates, and potential biases (see Chapter 2, Limitations). These limitations should be considered when interpreting the findings and future research in this field.

Experimental study

The experimental study had several limitations that should be considered when interpreting the results and generalising the findings. We identified potential response bias suggested by inconsistent pain reports from two participants. Further, we deviated from the protocol by proposing an alternative method of computing the magnitude of SH (see Chapter 3, Limitations). Additionally, some data for ratings of high-frequency stimulation trains were lost, preventing a comprehensive analysis of the ratings of trains. It is important to note that a statistical model to fit the data on the magnitude of SH could not be identified, thus preventing the investigation of the relationships between PSD and the magnitude of SH.

Recommendations for future research

Systematic review

NMDA antagonists showed no effect in modulating mechanisms underlying SH. However, this review does not rule out that NMDA antagonists could influence SH. The influence of NMDA antagonists on SH could be explained by other factors, such as genetics. For example, a study investigated the expression levels of NMDA receptor subunits (GluN2A, GluN2D, GluN3A, and GluN3B) in human blood lymphocytes as a potential biomarker for online computer game addiction (Sadat-Shirazi et al., 2018). Game addicts had higher mRNA expression of NR2A and NR2B subunits than control subjects. By investigating the expression of NMDA receptors in individuals, researchers can identify subgroups of individuals who may be more likely to benefit from NMDA antagonist treatment.

To improve the quality of experimental pain research, there are several recommendations. First, we had to exclude several studies from the meta-analysis due to missing data. Further, when no raw data were reported, most data were extracted from figures when no raw data were reported. Researchers should make their raw data publicly available to facilitate transparency and reproducibility. Platforms like Open Science Framework (https://osf.io/) can be used to share raw data (Soderberg, 2018). Second, several studies reported the surface area in units other than cm^2. Therefore, the units used to report the surface area of SH should be standardised to avoid performing conversion calculations that could potentially confound data. Third, female participants are often excluded from experimental pain studies because research suggests that females show greater sensitivity to pain across multiple stimulus modalities than compared to males. In addition, there are concerns about variable pain sensitivity through a menstrual cycle. However, statistical models can be developed to manage this variation (Bartley & Fillingim, 2013; Melchior et al., 2016). In addition, most of the included studies used small samples. Therefore, the efficacy of NMDA antagonists should be reproduced in larger sample sizes to support the generalisability of the research. Finally, the overall risk of bias in the included studies was unclear, which reduced our confidence in the research. Therefore, future

researchers should pay careful attention to the risk of bias domains when conducting research as we did in Chapter 3.

Experimental study

This study provides evidence for the association between PSD and the surface area of SH, independent of pain report. However, to enhance the quality and reliability of future studies there are several recommendations. There should be standardised techniques for computing the magnitude of SH to ensure consistent and reliable results across studies. Additionally, we observed a potential response bias in inconsistent pain reports from two participants, indicating the need for measures to minimise bias in future studies, such as rigorous participant screening and vigilant monitoring of pain reports. Additionally, we could not find a suitable statistical model for the magnitude of SH, and we noticed a clustering of the magnitude data, possibly due to our limited sample size. Therefore, future studies should conduct replication studies with larger sample sizes to reduce the risk of this clustering-type effect identify an appropriate statistical model and conduct a more comprehensive investigation of the relationships between PSD, SH and pain. These improvements will support the advancement of knowledge in this area and contribute to future research.

Conclusion

The current data highlight the ongoing gaps in our understanding of the underlying mechanisms of SH. The influence of NMDA-targeting pharmacological manipulations on experimentally induced SH in healthy human adults without clinical pain remains a topic of ongoing investigation. The present systematic review adds to the existing literature by demonstrating limited effects of NMDA antagonists on experimentally induced SH. Further investigation into other pharmacological or non-pharmacological therapies is required to evaluate their potential in modulating SH. Future studies should explore these therapies while using rigorous methodologies and considering the limitations of this study. This will help refine our knowledge and gain valuable insights into the mechanisms that underlie SH.

Our experimental study found a strong positive relationship between PSD and the surface area of SH, independent of the presence of pain, in individuals living with HIV. These results imply a potential therapeutic role of PSD-reducing treatments in addressing SH and other pain syndromes that share mechanisms with SH. Future studies should also acknowledge the distinction between experimental SH and clinical SH. While our study focused on experimental SH, it would be valuable to explore the effects of distress-reducing therapy on clinical SH. Existing experimental work has demonstrated the potential for distress-reducing therapy to reduce experimental SH (You et al., 2014). Therefore, a meta-analytical review examining the effects of both pharmacological and non-pharmacological distress-reducing therapies on clinical SH would be beneficial. Additionally, considering the high rates of distress, challenging living circumstances, and poor health conditions experienced by the study

population, it is plausible that the role of PSD may be particularly pronounced in this context. It is essential to recognise the influence of psychosocial factors when assessing and managing pain in individuals with HIV, as addressing PSD may play a role in reducing CS and improving SH outcomes. However, it is also important to acknowledge the limitations of our study. The research is based on a specific population of individuals living with HIV, and the generalisability of the findings to other populations may be limited. Again, further research is needed to explore the underlying mechanisms that link PSD to SH, including potential mediators and moderators of this relationship.

.

REFERENCES

ANDERSEN, O.K., FELSBY, S., NICOLAISEN, L., BJERRING, P., JENSEN, T.S. & ARENDT-NIELSEN, L. 1996. The effect of Ketamine on stimulation of primary and secondary hyperalgesic areas induced by capsaicin--a double-blind, placebo-controlled, human experimental study. *Pain.* 66(1):51-62. DOI:10.1016/0304-3959(96)02995-8.

ANGST, M.S., KOPPERT, W., PAHL, I., CLARK, D.J. & SCHMELZ, M. 2003. Short-term infusion of the mu-opioid agonist remifentanil in humans causes hyperalgesia during withdrawal. *Pain.* 106(1-2):49-57. DOI:10.1016/s0304-3959(03)00276-8.

ARENDT-NIELSEN, L., MORLION, B., PERROT, S., DAHAN, A., DICKENSON, A., KRESS, H., WELLS, C., BOUHASSIRA, D. et al. 2018. Assessment and manifestation of central sensitisation across different chronic pain conditions. 22(2):216-241.

AUGUIE, B. & ANTONOV, A. 2017. gridExtra: miscellaneous functions for "grid" graphics. *R package version.* 2(1).

BACHE, S.M. & WICKHAM, H. 2022. *magrittr: A Forward-Pipe Operator for R*. Available: https://CRAN.R-project.org/package=magrittr [Available: https://CRAN.R-project.org/package=magrittr].

BALDUZZI, S., RÜCKER, G. & SCHWARZER, G. 2019. How to perform a meta-analysis with R: a practical tutorial. *Evid Based Ment Health.* 22(4):153-160. DOI:10.1136/ebmental-2019-300117.

BANG, H., NI, L. & DAVIS, C.E. 2004. Assessment of blinding in clinical trials. *Control Clin Trials.* 25(2):143-156. DOI:10.1016/j.cct.2003.10.016.

BARTLEY, E.J. & FILLINGIM, R.B. 2013. Sex differences in pain: a brief review of clinical and experimental findings. *British journal of anaesthesia.* 111(1):52-58.

BAUMGÄRTNER, U., MAGERL, W., KLEIN, T., HOPF, H.C. & TREEDE, R.D. 2002. Neurogenic hyperalgesia versus painful hypoalgesia: two distinct mechanisms of neuropathic pain. *Pain.* 96(1-2):141-151. DOI:10.1016/s0304-3959(01)00438-9.

BEDWELL, G.J., LOUW, C., PARKER, R., VAN DEN BROEKE, E., VLAEYEN, J.W., MOSELEY, G.L. & MADDEN, V.J. 2022. The influence of a manipulation of threat on experimentally-induced secondary hyperalgesia. *PeerJ.* 10:e13512. DOI:10.7717/peerj.13512.

BIURRUN MANRESA, J., KÆSELER ANDERSEN, O., MOURAUX, A. & VAN DEN BROEKE, E.N. 2018. High frequency electrical stimulation induces a long-lasting enhancement of event-related potentials but does not change the perception elicited by intra-epidermal electrical stimuli delivered to the area of increased mechanical pinprick sensitivity. *PloS one.* 13(9):e0203365. DOI:10.1371/journal.pone.0203365.

BLYTH, F.M., MARCH, L.M., BRNABIC, A.J.M., JORM, L.R., WILLIAMSON, M. & COUSINS, M.J. 2001. Chronic pain in Australia: a prevalence study. *Pain.* 89(2). Available: https://journals.lww.com/pain/Fulltext/2001/01000/Chronic_pain_in_Australia__a_prevalence_study.4.aspx.

BORCHERS, H.W. 2019. Pracma: practical numerical math functions. *R package version.* 2(9):519.

BRENNUM, J., KAISER, F. & DAHL, J.B. 2001. Effect of naloxone on primary and secondary hyperalgesia induced by the human burn injury model. *Acta Anaesthesiologica Scandinavica.* 45(8):954-960. DOI:https://doi.org/10.1034/j.1399-6576.2001.450806.x.

BRESINK, I., DANYSZ, W., PARSONS, C.G. & MUTSCHLER, E. 1995. Different binding affinities of NMDA receptor channel blockers in various brain regions--indication of NMDA receptor heterogeneity. *Neuropharmacology.* 34(5):533-540. DOI:10.1016/0028-3908(95)00017-z.

BRISLIN, R.W. 1970. Back-Translation for Cross-Cultural Research. *Journal of Cross-Cultural Psychology.* 1(3):185-216. DOI:10.1177/135910457000100301.

BROEKE, E.N.v.d. & MOURAUX, A. 2014. High-frequency electrical stimulation of the human skin induces heterotopical mechanical hyperalgesia, heat hyperalgesia, and enhanced responses to nonnociceptive vibrotactile input. *Journal of Neurophysiology.* 111(8):1564-1573. DOI:10.1152/jn.00651.2013.

BURNS, J.W., GERHART, J.I., BRUEHL, S., PETERSON, K.M., SMITH, D.A., PORTER, L.S., SCHUSTER, E., KINNER, E. et al. 2015. Anger arousal and behavioral anger regulation in everyday life among patients with chronic low back pain: Relationships to patient pain and function. Health Psychol. 34(5):547-555. DOI:10.1037/hea0000091.

CAMPBELL, J.N. & MEYER, R.A. 2006. Mechanisms of neuropathic pain. *Neuron.* 52(1):77-92. DOI:10.1016/j.neuron.2006.09.021.

CARPENTER, K.J. & DICKENSON, A.H. 1999. Translational vignette: NMDA receptors and pain--hopes for novel analgesics. *Regional Anesthesia and Pain Medicine.* 24(6):506.

CARLSSON, K., HOEM, N., MOBERG, E. & MATHISEN, L. 2004. Analgesic effect of dextromethorphan in neuropathic pain. Acta Anaesthesiologica Scandinavica. 48(3):328-336.

CASALE, R., ATZENI, F., BAZZICHI, L., BERETTA, G., COSTANTINI, E., SACERDOTE, P. & TASSORELLI, C. 2021. Pain in Women: A Perspective Review on a Relevant Clinical Issue that Deserves Prioritization. *Pain and therapy.* 10(1):287-314. DOI:10.1007/s40122-021-00244-1.

CASTILLO, R.C., WEGENER, S.T., HEINS, S.E., HAYTHORNTHWAITE, J.A., MACKENZIE, E.J. & BOSSE, M.J. 2013. Longitudinal relationships between anxiety, depression, and pain: results from a two-year cohort study of lower extremity trauma patients. Pain. 154(12):2860-2866. DOI:10.1016/j.pain.2013.08.025.

CATERINA, M.J., SCHUMACHER, M.A., TOMINAGA, M., ROSEN, T.A., LEVINE, J.D. & JULIUS, D. 1997. The capsaicin receptor: a heat-activated ion channel in the pain pathway. *Nature.* 389(6653):816-824. DOI:10.1038/39807.

CATERINA, M.J., LEFFLER, A., MALMBERG, A.B., MARTIN, W.J., TRAFTON, J., PETERSEN-ZEITZ, K.R., KOLTZENBURG, M., BASBAUM, A.I. et al. 2000. Impaired nociception and pain sensation in mice lacking the capsaicin receptor. *Science.* 288(5464):306-313. DOI:10.1126/science.288.5464.306.

CATZ, S.L., GORE-FELTON, C. & MCCLURE, J.B. 2002. Psychological distress among minority and low-income women living with HIV. *Behavioral medicine.* 28(2):53-60.

CAYROL, T., PITANCE, L., ROUSSEL, N., MOURAUX, A. & VAN DEN BROEKE, E.N. 2018. No evidence of widespread mechanical pressure hyperalgesia after experimentally induced central sensitization through skin nociceptors. *Pain reports.* 3(6):e691-e691. DOI:10.1097/PR9.0000000000000691.

CHASSAING, C., SCHMIDT, J., ESCHALIER, A., CARDOT, J.M. & DUBRAY, C. 2006. Hyperalgesia induced by cutaneous freeze injury for testing analgesics in healthy volunteers. *Br J Clin Pharmacol.* 61(4):389-397. DOI:10.1111/j.1365-2125.2006.02582.x.

CHEN, S.-R., ZHOU, H.-Y., BYUN, H.S., CHEN, H. & PAN, H.-L. 2014. Casein kinase II regulates N-methyl-D-aspartate receptor activity in spinal cords and pain hypersensitivity induced by nerve injury. Journal of Pharmacology and Experimental Therapeutics. 350(2):301-312.

CLEELAND, C.S. & RYAN, K.M. 1994. Pain assessment: global use of the Brief Pain Inventory. *Ann Acad Med Singap.* 23(2):129-138.

CORRELL, G.E., MALEKI, J., GRACELY, E.J., MUIR, J.J. & HARBUT, R.E. 2004. Subanesthetic ketamine infusion therapy: a retrospective analysis of a novel therapeutic approach to complex regional pain syndrome. Pain Medicine. 5(3):263-275.

CROFFORD, L.J. 2015. Chronic Pain: Where the Body Meets the Brain. *Transactions of the American Clinical and Climatological Association.* 126:167-183. Available: https://pubmed.ncbi.nlm.nih.gov/26330672 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4530716/.

DALE, S.K., COHEN, M.H., KELSO, G.A., CRUISE, R.C., WEBER, K.M., WATSON, C., BURKE-MILLER, J.K. & BRODY, L.R. 2014. Resilience among women with HIV: Impact of silencing the self and socioeconomic factors. Sex roles. 70:221-231.

DAVIS, J.B., GRAY, J., GUNTHORPE, M.J., HATCHER, J.P., DAVEY, P.T., OVEREND, P., HARRIES, M.H., LATCHAM, J. et al. 2000. Vanilloid receptor-1 is essential for inflammatory thermal hyperalgesia. *Nature.* 405(6783):183-187. DOI:10.1038/35012076.

DELLA PORTA, D., VILZ, M.-L., KUZMINOVA, A., FILBRICH, L., MOURAUX, A. & LEGRAIN, V. 2022. No evidence for an effect of selective spatial attention on the development of secondary hyperalgesia: A replication study. *Frontiers in Human Neuroscience.* 16. DOI:10.3389/fnhum.2022.997230.

DESHMUKH, N.N., BORKAR, A.M. & DESHMUKH, J.S. 2017. Depression and its associated factors among people living with HIV/AIDS: Can it affect their quality of life? *Journal of family medicine and primary care.* 6(3):549-553. DOI:10.4103/2249-4863.222016.

DE TOMMASO, M., FEDERICI, A., SANTOSTASI, R., CALABRESE, R., VECCHIO, E., LAPADULA, G., IANNONE, F., LAMBERTI, P. et al. 2011. Laser-evoked potentials habituation in fibromyalgia. The journal of pain. 12(1):116-124.

DITRE, J.W., ZALE, E.L., LAROWE, L.R., KOSIBA, J.D. & DE VITA, M.J. 2018. Nicotine deprivation increases pain intensity, neurogenic inflammation, and mechanical hyperalgesia among daily tobacco smokers. *Journal of abnormal psychology.* 127(6):578.

DUEDAHL, T.H., DIRKS, J., PETERSEN, K.B., ROMSING, J., LARSEN, N.-E. & DAHL, J.B. 2005. Intravenous dextromethorphan to human volunteers: relationship between pharmacokinetics and anti-hyperalgesic effect. *Pain.* 113(3):360-368. DOI:10.1016/j.pain.2004.11.015.

DURIC, V., CLAYTON, S., LEONG, M.L. & YUAN, L.-L. 2016. Comorbidity Factors and Brain Mechanisms Linking Chronic Stress and Systemic Illness. *Neural plasticity.* 2016:5460732-5460732. DOI:10.1155/2016/5460732.

EDWARDS, R.R., DWORKIN, R.H., SULLIVAN, M.D., TURK, D.C. & WASAN, A.D. 2016. The Role of Psychosocial Processes in the Development and Maintenance of Chronic Pain. *The journal of pain.* 17(9 Suppl):T70-92. DOI:10.1016/j.jpain.2016.01.001.

EGGER, M., DAVEY SMITH, G., SCHNEIDER, M. & MINDER, C. 1997. Bias in meta-analysis detected by a simple, graphical test. *Bmj.* 315(7109):629-634. DOI:10.1136/bmj.315.7109.629.

EMLET, C.A., SHIU, C., KIM, H.-J. & FREDRIKSEN-GOLDSEN, K. 2017. Bouncing back: resilience and mastery among HIV-positive older gay and bisexual men. The Gerontologist. 57(suppl_1):S40-S49.

FARLEY, J., MILLER, E., ZAMANI, A., TEPPER, V., MORRIS, C., OYEGUNLE, M., LIN, M., CHARURAT, M. et al. 2010. Screening for hazardous alcohol use and depressive symptomatology among HIV-infected patients in Nigeria: prevalence, predictors, and association with adherence. *Journal of the International Association of Physicians in AIDS Care.* 9(4):218-226.

FATTORI, V., HOHMANN, M.S., ROSSANEIS, A.C., PINHO-RIBEIRO, F.A. & VERRI, W.A. 2016. Capsaicin: Current Understanding of Its Mechanisms and Therapy of Pain and Other Pre-Clinical and Clinical Uses. *Molecules.* 21(7). DOI:10.3390/molecules21070844.

FELSBY, S., NIELSEN, J., ARENDT-NIELSEN, L. & JENSEN, T.S. 1996. NMDA receptor blockade in chronic neuropathic pain: a comparison of ketamine and magnesium chloride. Pain. 64(2):283-291.

FENSELAU, H., HEINKE, B. & SANDKÜHLER, J. 2011. Heterosynaptic long-term potentiation at GABAergic synapses of spinal lamina I neurons. Journal of Neuroscience. 31(48):17383-17391

FIELD, A., MILES, J. & FIELD, Z. 2012. *Discovering statistics using R.* Sage publications.

FILBRICH, L., VAN DEN BROEKE, E.N., LEGRAIN, V. & MOURAUX, A. 2020. The focus of spatial attention during the induction of central sensitization can modulate the subsequent development of secondary hyperalgesia. *Cortex.* 124:193-203. DOI:https://doi.org/10.1016/j.cortex.2019.11.014.

FILLINGIM, R.B., LOESER, J.D., BARON, R. & EDWARDS, R.R. 2016. Assessment of Chronic Pain: Domains, Methods, and Mechanisms. *The journal of pain.* 17(9 Suppl):T10-T20. DOI:10.1016/j.jpain.2015.08.010.

FINCH, P.M., KNUDSEN, L. & DRUMMOND, P.D. 2009. Reduction of allodynia in patients with complex regional pain syndrome: A double-blind placebo-controlled trial of topical ketamine. Pain. 146(1-2):18-25. DOI:10.1016/j.pain.2009.05.017.

FINNERUP, N.B., ATTAL, N., HAROUTOUNIAN, S., MCNICOL, E., BARON, R., DWORKIN, R.H., GILRON, I., HAANPÄÄ, M. et al. 2015. Pharmacotherapy for neuropathic pain in adults: a systematic

review and meta-analysis. Lancet Neurol. 14(2):162-173. DOI:10.1016/s1474-4422(14)70251-0.
FINNERUP, N.B., KUNER, R. & JENSEN, T.S. 2020. Neuropathic Pain: From Mechanisms to Treatment. *Physiological Reviews.* 101(1):259-301. DOI:10.1152/physrev.00045.2019.
FRYMOYER, A.R., ROWBOTHAM, M.C. & PETERSEN, K.L. 2007. Placebo-controlled comparison of a morphine/dextromethorphan combination with morphine on experimental pain and hyperalgesia in healthy volunteers. *J Pain.* 8(1):19-25. DOI:10.1016/j.jpain.2006.05.010.
FUNDYTUS, M.E. 2001. Glutamate receptors and nociception: implications for the drug treatment of pain. CNS Drugs. 15(1):29-58. DOI:10.2165/00023210-200115010-00004.
GALER, B., TWILLING, L., HARLE, J., CLUFF, R., FRIEDMAN, E. & ROWBOTHAM, M. 2000. Lack of efficacy of riluzole in the treatment of peripheral neuropathic pain conditions. Neurology. 55(7):971-975.
GATCHEL, R.J., PENG, Y.B., PETERS, M.L., FUCHS, P.N. & TURK, D.C.J.P.b. 2007. The biopsychosocial approach to chronic pain: scientific advances and future directions. 133(4):581.
GELMAN, A. & HILL, J. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models.* Cambridge: Cambridge University PressDOI:DOI: 10.1017/CBO9780511790942.
GELMAN, A., CARLIN, J., STERN, H., DUNSON, D., VEHTARI, A. & RUBIN, D. 2021. *Bayesian Data Analysis Third edition (with errors fixed as of 15 February 2021).* 3. Available: http://www.stat.columbia.edu/~gelman/book/BDA3.pdf.
GOHEL, D. & SKINTZOS, P. 2022. *flextable: Functions for Tabular Reporting*. Available: https://CRAN.R-project.org/package=flextable [Available:].
GOLDBERG, M.E., DOMSKY, R., SCARINGE, D., HIRSH, R., DOTSON, J., SHARAF, I., TORJMAN, M.C. & SCHWARTZMAN, R.J. 2005. Multi-day low dose ketamine infusion for the treatment of complex regional pain syndrome. Pain Physician. 8(2):175.
GONZALEZ, C.E., OKUNBOR, J.I., PARKER, R., OWENS, M.A., WHITE, D.M., MERLIN, J.S. & GOODIN, B.R. 2019. Pain-Specific Resilience in People Living With HIV and Chronic Pain: Beneficial Associations With Coping Strategies and Catastrophizing. *Frontiers in Psychology.* 10. DOI:10.3389/fpsyg.2019.02046.
GOODIN, B.R., OWENS, M.A., WHITE, D.M., STRATH, L.J., GONZALEZ, C., RAINEY, R.L., OKUNBOR, J.I., HEATH, S.L. et al. 2018. Intersectional health-related stigma in persons living with HIV and chronic pain: implications for depressive symptoms. *AIDS Care.* 30(sup2):66-73.
GORDON, M. & LUMLEY, T. 2022. *forestplot: Advanced Forest Plot Using 'grid' Graphics*. Available: https://CRAN.R-project.org/package=forestplot [Available: https://CRAN.R-project.org/package=forestplot].
GOTTRUP, H., BACH, F.W. & JENSEN, T.S. 2004. Differential effects of peripheral ketamine and lidocaine on skin flux and hyperalgesia induced by intradermal capsaicin in humans. *Clinical physiology and functional imaging.* 24(2):103-108.
GOTTRUP, H., BACH, F.W., ARENDT-NIELSEN, L. & JENSEN, T.S. 2000a. Peripheral lidocaine but not ketamine inhibits capsaicin-induced hyperalgesia in humans. *Br J Anaesth.* 85(4):520-528. DOI:10.1093/bja/85.4.520.
GOTTRUP, H., HANSEN, P.O., ARENDT-NIELSEN, L. & JENSEN, T.S. 2000b. Differential effects of systemically administered ketamine and lidocaine on dynamic and static hyperalgesia induced by intradermal capsaicin in humans. *Br J Anaesth.* 84(2):155-162. DOI:10.1093/oxfordjournals.bja.a013396.
GUYATT, G.H., OXMAN, A.D., KUNZ, R., VIST, G.E., FALCK-YTTER, Y. & SCHÜNEMANN, H.J. 2008. What is “quality of evidence” and why is it important to clinicians? 336(7651):995-998. DOI:10.1136/bmj.39490.551019.BE %J BMJ.
HADLANDSMYTH, K., SABIC, E., ZIMMERMAN, M.B., SLUKA, K.A., HERR, K.A., CLARK, C.R., NOISEUX, N.O., CALLAGHAN, J.J. et al. 2017. Relationships among pain intensity, pain-related distress, and psychological distress in pre-surgical total knee arthroplasty patients: a secondary analysis. Psychol Health Med. 22(5):552-563. DOI:10.1080/13548506.2016.1189581.

HANSEN, K.B., YI, F., PERSZYK, R.E., FURUKAWA, H., WOLLMUTH, L.P., GIBB, A.J. & TRAYNELIS, S.F. 2018. Structure, function, and allosteric modulation of NMDA receptors. *J Gen Physiol.* 150(8):1081-1105. DOI:10.1085/jgp.201812032.

HANSEN, N., KLEIN, T., MAGERL, W. & TREEDE, R.-D. 2007. Psychophysical Evidence for Long-Term Potentiation of C-Fiber and Aδ-Fiber Pathways in Humans by Analysis of Pain Descriptors. *Journal of Neurophysiology.* 97(3):2559-2563. DOI:10.1152/jn.01125.2006.

HARBUT, R.E. & CORRELL, G.E. 2002. Successful treatment of a nine-year case of complex regional pain syndrome type-I (reflex sympathetic dystrophy) with intravenous ketamine-infusion therapy in a warfarin-anticoagulated adult female patient. Pain Medicine. 3(2):147-155.

HARRER, M., CUIJPERS, P., FURUKAWA, T. & EBERT, D.D. 2019. *dmetar: Companion R Package For The Guide 'Doing Meta-Analysis in R'*. Available: http://dmetar.protectlab.org/ [Available: http://dmetar.protectlab.org/].

HAUSER, D.J., ELLSWORTH, P.C. & GONZALEZ, R. 2018. Are Manipulation Checks Necessary? *Frontiers in Psychology.* 9(998). DOI:10.3389/fpsyg.2018.00998.

HEINZEN, E., SINNWELL, J., ATKINSON, E., GUNDERSON, T. & DOUGHERTY, G. 2021. *arsenal: An Arsenal of 'R' Functions for Large-Scale Statistical Summaries*. Available: https://CRAN.R-project.org/package=arsenal [Available:].

HENWOOD, P. & ELLIS, J.A. 2004. Chronic neuropathic pain in spinal cord injury: the patient's perspective. Pain Res Manag. 9(1):39-45. DOI:10.1155/2004/863062.

HIGGINS, J.P., THOMAS, J., CHANDLER, J., CUMPSTON, M., LI, T., PAGE, M.J. & WELCH, V.A. 2019. *Cochrane handbook for systematic reviews of interventions.* John Wiley & Sons.

HIGGINS, J.P., ALTMAN, D.G., GØTZSCHE, P.C., JÜNI, P., MOHER, D., OXMAN, A.D., SAVOVIĆ, J., SCHULZ, K.F. et al. 2011. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. 343.

HUGE, V., LAUCHART, M., MAGERL, W., SCHELLING, G., BEYER, A., THIEME, D. & AZAD, S.C. 2010. Effects of low-dose intranasal (S)-ketamine in patients with neuropathic pain. European Journal of Pain. 14(4):387-394.

HWANG, S.J., BURETTE, A., RUSTIONI, A. & VALTSCHANOFF, J.G. 2004. Vanilloid receptor VR1-positive primary afferents are glutamatergic and contact spinal neurons that co-express neurokinin receptor NK1 and glutamate receptors. *J Neurocytol.* 33(3):321-329. DOI:10.1023/B:NEUR.0000044193.31523.a1.

IASP. 2021a. *IASP Terminology*. Available: https://www.iasp-pain.org/Education/Content.aspx?ItemNumber=1698#Pain [2021, 17 June].

IASP. 2021b. *IASP Terminology*. Available: https://www.iasp-pain.org/resources/terminology/?ItemNumber=1698#hyperalgesia [2021, 29 September].

IASP. 2021c. *IASP Terminology*. Available: https://www.iasp-pain.org/Education/Content.aspx?ItemNumber=1698&navItemNumber=576#Allodynia [2020, 05 February].

IASP. 2021d. *Terminology*. Available: https://www.iasp-pain.org/resources/terminology/?ItemNumber=1698#central-sensitization [2021, 16 November].

IASP. 2021e. *IASP Terminology*. Available: https://www.iasp-pain.org/Education/Content.aspx?ItemNumber=1698&navItemNumber=576#Sensitization [2021, 19 March].

IASP. 2021f. *IASP Terminology* Available: https://www.iasp-pain.org/resources/terminology/#nociplasticpain [2023, 03 February].

IASP. 2021g. *IASP Terminology* Available: https://www.iasp-pain.org/resources/terminology/#centralsensitization [2023, 04 February].

IASP. 2021h. *IASP Terminology*. Available: https://www.iasp-pain.org/resources/terminology/#nocicpetivepain [2022, 01 February].

IASP. 2021i. *IASP Terminology*. Available: https://www.iasp-pain.org/resources/terminology/#neuropathicpain [2023, 03 February].
IASP. 2021j. *IASP Terminology*. Available: https://www.iasp-pain.org/resources/terminology/ [2022, December].
IFTINCA, M., DEFAYE, M. & ALTIER, C. 2021. TRPV1-Targeted Drugs in Development for Human Pain Conditions. *Drugs.* 81(1):7-27. DOI:10.1007/s40265-020-01429-2.
IKEDA, H., ASAI, T. & MURASE, K. 2000. Robust changes of afferent-induced excitation in the rat spinal dorsal horn after conditioning high-frequency stimulation. Journal of Neurophysiology. 83(4):2412-2420.
ILKJAER, S., PETERSEN, K.L., BRENNUM, J., WERNBERG, M. & DAHL, J.B. 1996. Effect of systemic N-methyl-D-aspartate receptor antagonist (ketamine) on primary and secondary hyperalgesia in humans. *BJA: British Journal of Anaesthesia.* 76(6):829-834. DOI:10.1093/bja/76.6.829.
ILKJAER, S., DIRKS, J., BRENNUM, J., WERNBERG, M. & DAHL, J.B. 1997. Effect of systemic N-methyl-D-aspartate receptor antagonist (dextromethorphan) on primary and secondary hyperalgesia in humans. *BJA: British Journal of Anaesthesia.* 79(5):600-605. DOI:10.1093/bja/79.5.600.
INTHOUT, J., IOANNIDIS, J.P.A. & BORM, G.F. 2014. The Hartung-Knapp-Sidik-Jonkman method for random effects meta-analysis is straightforward and considerably outperforms the standard DerSimonian-Laird method. *BMC Medical Research Methodology.* 14(1):25. DOI:10.1186/1471-2288-14-25.
JACKSON, D.L., GRAFF, C.B., RICHARDSON, J.D. & HARGREAVES, K.M. 1995. Glutamate participates in the peripheral modulation of thermal hyperalgesia in rats. *European Journal of Pharmacology.* 284(3):321-325. DOI:https://doi.org/10.1016/0014-2999(95)00449-U.
JARA-OSEGUERA, A., SIMON, S.A. & ROSENBAUM, T. 2008. TRPV1: on the road to pain relief. *Curr Mol Pharmacol.* 1(3):255-269. DOI:10.2174/1874467210801030255.
JENSEN, T.S. & BARON, R. 2003. Translation of symptoms and signs into mechanisms in neuropathic pain. *Pain.* 102(1-2):1-8. DOI:10.1016/s0304-3959(03)00006-x.
JI, G., ZHOU, S., KOCHUKOV, M.Y., WESTLUND, K.N. & CARLTON, S.M. 2007. Plasticity in intact A delta- and C-fibers contributes to cold hypersensitivity in neuropathic rats. *Neuroscience.* 150(1):182-193. DOI:10.1016/j.neuroscience.2007.09.002.
JI, R.-R., KOHNO, T., MOORE, K.A. & WOOLF, C.J. 2003. Central sensitization and LTP: do pain and memory share similar mechanisms? *Trends in Neurosciences.* 26(12):696-705. DOI:https://doi.org/10.1016/j.tins.2003.09.017.
JI, R.-R. & WOOLF, C.J. 2001. Neuronal Plasticity and Signal Transduction in Nociceptive Neurons: Implications for the Initiation and Maintenance of Pathological Pain. Neurobiology of Disease. 8(1):1-10. DOI:https://doi.org/10.1006/nbdi.2000.0360.
KAGEE, A. & MARTIN, L. 2010. Symptoms of depression and anxiety among a sample of South African patients living with HIV. *AIDS Care.* 22(2):159-165.
KAGEE, A., SAAL, W. & BANTJES, J. 2017. Distress, depression and anxiety among persons seeking HIV testing. *AIDS Care.* 29(3):280-284. DOI:10.1080/09540121.2016.1259453.
KAGEE, A.J.E. & HEALTH. 2005. Symptoms of distress and posttraumatic stress among South African former political detainees. 10(2):169-179.
KAHARUZA, F.M., BUNNELL, R., MOSS, S., PURCELL, D.W., BIKAAKO-KAJURA, W., WAMAI, N., DOWNING, R., SOLBERG, P. et al. 2006. Depression and CD4 cell count among persons with HIV infection in Uganda. *AIDS and Behavior.* 10(1):105-111.
KAMERMAN, P.R., BRADSHAW, D., LAUBSCHER, R., PILLAY-VAN WYK, V., GRAY, G.E., MITCHELL, D. & CHETTY, S. 2020. Almost 1 in 5 South African adults have chronic pain: a prevalence study conducted in a large nationally representative sample. Pain. 161(7):1629-1635. DOI:10.1097/j.pain.0000000000001844.
KIEFER, R.-T., ROHR, P., PLOPPA, A., DIETERICH, H.-J., GROTHUSEN, J., KOFFLER, S., ALTEMEYER, K.-H., UNERTL, K. et al. 2008. Efficacy of Ketamine in Anesthetic Dosage for the Treatment of

Refractory Complex Regional Pain Syndrome: An Open-Label Phase II Study. Pain Medicine. 9(8):1173-1201. DOI:10.1111/j.1526-4637.2007.00402.x.

KLEIN, T., MAGERL, W. & TREEDE, R.-D. 2006. Perceptual Correlate of Nociceptive Long-Term Potentiation (LTP) in Humans Shares the Time Course of Early-LTP. *Journal of Neurophysiology.* 96(6):3551-3555. DOI:10.1152/jn.00755.2006.

KLEIN, T., MAGERL, W., ROLKE, R. & TREEDE, R.-D. 2005. Human surrogate models of neuropathic pain. *Pain.* 115(3). Available: https://journals.lww.com/pain/Fulltext/2005/06000/Human_surrogate_models_of_neuropathic_pain.2.aspx.

KLEIN, T., STAHN, S., MAGERL, W. & TREEDE, R.D. 2008a. The role of heterosynaptic facilitation in long-term potentiation (LTP) of human pain sensation. *Pain.* 139(3):507-519. DOI:10.1016/j.pain.2008.06.001.

KLEIN, T., MAGERL, W., HOPF, H.-C., SANDKÜHLER, J. & TREEDE, R.-D. 2004. Perceptual correlates of nociceptive long-term potentiation and long-term depression in humans. *The Journal of neuroscience : the official journal of the Society for Neuroscience.* 24(4):964-971. DOI:10.1523/JNEUROSCI.1222-03.2004.

KLEIN, T., MAGERL, W., HANSCHMANN, A., ALTHAUS, M. & TREEDE, R.D. 2008b. Antihyperalgesic and analgesic properties of the N-methyl-D-aspartate (NMDA) receptor antagonist neramexane in a human surrogate model of neurogenic hyperalgesia. *Eur J Pain.* 12(1):17-29. DOI:10.1016/j.ejpain.2007.02.002.

KLEIN, T., MAGERL, W., NICKEL, U., HOPF, H.C., SANDKÜHLER, J. & TREEDE, R.D. 2007. Effects of the NMDA-receptor antagonist ketamine on perceptual correlates of long-term potentiation within the nociceptive system. *Neuropharmacology.* 52(2):655-661. DOI:10.1016/j.neuropharm.2006.09.008.

KNAPP, G. & HARTUNG, J. 2003. Improved tests for a random effects meta-regression with a single covariate. *Stat Med.* 22(17):2693-2710. DOI:10.1002/sim.1482.

KOPPERT, W., ZECK, S., BLUNK, J.A., SCHMELZ, M., LIKAR, R. & SITTL, R. 1999. The effects of intradermal fentanyl and ketamine on capsaicin-induced secondary hyperalgesia and flare reaction. *Anesth Analg.* 89(6):1521-1527. DOI:10.1097/00000539-199912000-00039.

KOPPERT, W., DERN, S.K., SITTL, R., ALBRECHT, S., SCHÜTTLER, J. & SCHMELZ, M. 2001. A new model of electrically evoked pain and hyperalgesia in human skin: the effects of intravenous alfentanil, S(+)-ketamine, and lidocaine. *Anesthesiology.* 95(2):395-402. DOI:10.1097/00000542-200108000-00022.

KOPPERT, W., SITTL, R., SCHEUBER, K., ALSHEIMER, M., SCHMELZ, M. & SCHÜTTLER, J. 2003. Differential modulation of remifentanil-induced analgesia and postinfusion hyperalgesia by S-ketamine and clonidine in humans. *Anesthesiology.* 99(1):152-159. DOI:10.1097/00000542-200307000-00025.

KWON, M., ALTIN, M., DUENAS, H. & ALEV, L. 2014. The role of descending inhibitory pathways on chronic pain modulation and clinical implications. *Pain Practice.* 14(7):656-667.

LAGANA, L., CHEN, X.-H., KOOPMAN, C., CLASSEN, C., KIMERLING, R. & SPIEGAL, D. 2002. Depressive symptomatology in relation to emotional control and chronic pain in persons who are HIV positive. *Rehabilitation Psychology.* 47(4):402.

LATREMOLIERE, A. & WOOLF, C.J. 2009. Central sensitization: a generator of pain hypersensitivity by central neural plasticity. *The journal of pain.* 10(9):895-926. DOI:10.1016/j.jpain.2009.06.012.

LAWLER, K., MOSEPELE, M., SELOILWE, E., RATCLIFFE, S., STEELE, K., NTHOBATSANG, R. & STEENHOFF, A. 2011. Depression among HIV-positive individuals in Botswana: a behavioral surveillance. *AIDS and Behavior.* 15(1):204-208.

LE BARS, D., DICKENSON, A.H. & BESSON, J.M. 1979. Diffuse noxious inhibitory controls (DNIC). I. Effects on dorsal horn convergent neurones in the rat. *Pain.* 6(3):283-304. DOI:10.1016/0304-3959(79)90049-6.

LE BARS, D., GOZARIU, M. & CADDEN, S.W. 2001. Animal Models of Nociception. *Pharmacological Reviews.* 53(4):597-652. Available: https://pharmrev.aspetjournals.org/content/pharmrev/53/4/597.full.pdf.

LEUNG, A., WALLACE, M.S., RIDGEWAY, B. & YAKSH, T. 2001. Concentration–effect relationship of intravenous alfentanil and ketamine on peripheral neurosensory thresholds, allodynia and hyperalgesia of neuropathic pain. Pain. 91(1-2):177-187.

LIBERATI, A., ALTMAN, D.G., TETZLAFF, J., MULROW, C., GØTZSCHE, P.C., IOANNIDIS, J.P.A., CLARKE, M., DEVEREAUX, P.J. et al. 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ.* 339:b2700. DOI:10.1136/bmj.b2700.

LIN, T.-C., WONG, C.-S., CHEN, F.-C., LIN, S.-Y. & HO, S.-T. 1998. Long-term epidural ketamine, morphine and bupivacaine attenuate reflex sympathetic

LIU, X.G. & ZHOU, L.J. 2015. Long-term potentiation at spinal C-fiber synapses: a target for pathological pain. *Curr Pharm Des.* 21(7):895-905. DOI:10.2174/1381612820666141027115949.

LÜDECKE, D., BEN-SHACHAR, M.S., PATIL, I., WAGGONER, P. & MAKOWSKI, D. 2021. performance: An R package for assessment, comparison and testing of statistical models. *Journal of open source software.* 6(60).

LUND, C., SCHNEIDER, M., GARMAN, E.C., DAVIES, T., MUNODAWAFA, M., HONIKMAN, S., BHANA, A., BASS, J. et al. 2020. Task-sharing of psychological treatment for antenatal depression in Khayelitsha, South Africa: Effects on antenatal and postnatal outcomes in an individual randomised controlled trial. *Behaviour Research and Therapy.* 130:103466. DOI:https://doi.org/10.1016/j.brat.2019.103466.

LYNCH, M.E., CLARK, A.J., SAWYNOK, J. & SULLIVAN, M.J. 2005. Topical 2% amitriptyline and 1% ketamine in neuropathic pain syndromes: a randomized, double-blind, placebo-controlled trial. Anesthesiology. 103(1):140-146. DOI:10.1097/00000542-200507000-00021.

MADDEN, V.J., BEDWELL, G.J., CHIKEZIE, P.C., RICE, A.S.C. & KAMERMAN, P.R. 2019. A systematic review of experimental methods to manipulate secondary hyperalgesia in humans: protocol. *Systematic Reviews.* 8(1):208. DOI:10.1186/s13643-019-1120-7.

MADDEN, V.J., MSOLO, N., MQADI, L., LESOSKY, M., BEDWELL, G.J., HUTCHINSON, M.R., PETER, J.G., PARKER, R. et al. 2022. Study protocol: an observational study of distress, immune function and persistent pain in HIV. *BMJ Open.* 12(6):e059723. DOI:10.1136/bmjopen-2021-059723.

MAKOWSKI, D., BEN-SHACHAR, M.S. & LÜDECKE, D. 2019. bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. *Journal of open source software.* 4(40):1541.

MALLOY, K.M. & MILLING, L.S. 2010. The effectiveness of virtual reality distraction for pain reduction: a systematic review. Clinical psychology review. 30(8):1011-1018.

MANGIAFICO, S. 2022. *rcompanion: Functions to Support Extension Education Program Evaluation*. Available: https://CRAN.R-project.org/package=rcompanion [2022].

MARCUS, K.S., KERNS, R.D., ROSENFELD, B. & BREITBART, W. 2000. HIV/AIDS-related Pain as a Chronic Pain Condition: Implications of a Biopsychosocial Model for Comprehensive Assessment and Effective Management. *Pain Medicine.* 1(3):260-273. DOI:10.1046/j.1526-4637.2000.00033.x.

MARTIN, E., NARJOZ, C., DECLEVES, X., LABAT, L., LAMBERT, C., LORIOT, M.A., DUCHEIX, G., DUALÉ, C. et al. 2019. Dextromethorphan Analgesia in a Human Experimental Model of Hyperalgesia. *Anesthesiology.* 131(2):356-368. DOI:10.1097/aln.0000000000002736.

MATHIESEN, O., IMBIMBO, B.P., HILSTED, K.L., FABBRI, L. & DAHL, J.B. 2006. CHF3381, a N-methyl-D-aspartate receptor antagonist and monoamine oxidase-A inhibitor, attenuates secondary hyperalgesia in a human pain model. *J Pain.* 7(8):565-574. DOI:10.1016/j.jpain.2006.02.004.

MCCRACKEN, L.M. 1998. Learning to live with the pain: acceptance of pain predicts adjustment in persons with chronic pain. Pain. 74(1):21-27.

MCGRATH, S., ZHAO, X., STEELE, R. & BENEDETTI, A. 2022. *estmeansd: Estimating the Sample Mean and Standard Deviation from Commonly Reported Quantiles in Meta-Analysis*. Available: https://CRAN.R-project.org/package=estmeansd [Available: https://CRAN.R-project.org/package=estmeansd].

MCGUINNESS, L.A. 2019. *robvis: An R package and web application for visualising risk-of-bias assessments*. Available: https://github.com/mcguinlu/robvis [Available: https://github.com/mcguinlu/robvis].

MCGUINNESS, S.K., WASIAK, J., CLELAND, H., SYMONS, J., HOGAN, L., HUCKER, T. & MAHAR, P.D. 2011. A Systematic Review of Ketamine as an Analgesic Agent in Adult Burn Injuries. *Pain Medicine.* 12(10):1551-1558. DOI:10.1111/j.1526-4637.2011.01220.x.

MEESTERS, A.M.L., TEN DUIS, K., BANIERINK, H., STIRLER, V.M.A., WOUTERS, P.C.R., KRAEIMA, J., DE VRIES, J.P.M., WITJES, M.J.H. et al. 2020. What Are the Interobserver and Intraobserver Variability of Gap and Stepoff Measurements in Acetabular Fractures? *Clin Orthop Relat Res.* 478(12):2801-2808. DOI:10.1097/corr.0000000000001398.

MELCHIOR, M., POISBEAU, P., GAUMOND, I. & MARCHAND, S. 2016. Insights into the mechanisms and the emergence of sex-differences in pain. *Neuroscience.* 338:63-80. DOI:10.1016/j.neuroscience.2016.05.007.

MELZACK, R. & KATZ, J.J.T.o.p. 2006. Pain assessment in adult patients. 5:291-304.

MERLIN, J.S. 2015. Chronic Pain in Patients With HIV Infection: What Clinicians Need To Know. *Topics in antiviral medicine.* 23(3):120-124. Available: https://pubmed.ncbi.nlm.nih.gov/26518396 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6148933/.

MEYER, R.A., RINGKAMP, M., CAMPBELL, J.N. & RAJA, S.N. 2005. Neural mechanisms of hyperalgesia after tissue injury. *Johns Hopkins APL technical digest.* 26(1):56-66.

MEZA, R.C., ANCATÉN-GONZÁLEZ, C., CHIU, C.Q. & CHÁVEZ, A.E. 2022. Transient Receptor Potential Vanilloid 1 Function at Central Synapses in Health and Disease. *Frontiers in Cellular Neuroscience.* 16. DOI:10.3389/fncel.2022.864828.

MIASKOWSKI, C., PENKO, J.M., GUZMAN, D., MATTSON, J.E., BANGSBERG, D.R. & KUSHEL, M.B. 2011. Occurrence and characteristics of chronic pain in a community-based cohort of indigent adults living with HIV infection. *The journal of pain.* 12(9):1004-1016. DOI:10.1016/j.jpain.2011.04.002.

MIKKELSEN, S., ILKJAER, S., BRENNUM, J., BORGBJERG, F.M. & DAHL, J.B. 1999. The Effect of Naloxone on Ketamine-induced Effects on Hyperalgesia and Ketamine-induced Side Effects in Humans. *Anesthesiology.* 90(6):1539-1545. DOI:10.1097/00000542-199906000-00007.

MIKKELSEN, S., JØRGENSEN, H., LARSEN, P.S., BRENNUM, J. & DAHL, J.B. 2000. Effect of oral ketamine on secondary hyperalgesia, thermal and mechanical pain thresholds, and sedation in humans. *Regional anesthesia and pain medicine.* 25(5):452-458.

MIKKELSEN, S., DIRKS, J., FABRICIUS, P., PETERSEN, K.L., ROWBOTHAM, M.C. & DAHL, J.B. 2001. Effect of intravenous magnesium on pain and secondary hyperalgesia associated with the heat/capsaicin sensitization model in healthy volunteers. *Br J Anaesth.* 86(6):871-873. DOI:10.1093/bja/86.6.871.

MILLER, T.R., HALKITIS, P.N. & DURVASULA, R. 2019. A Biopsychosocial Approach to Managing HIV-Related Pain and Associated Substance Abuse in Older Adults: a Review. *Ageing International.* 44(1):74-116. DOI:10.1007/s12126-018-9333-y.

MOINI, J., AVGEROPOULOS, N.G. & SAMSAM, M. 2021. Chapter 2 - Cytology of the nervous system. In *Epidemiology of Brain and Spinal Tumors.* J. Moini, N.G. Avgeropoulos and M. Samsam, Eds.: Academic Press. 41-63. DOI:https://doi.org/10.1016/B978-0-12-821736-8.00012-1.

MONAHAN, P.O., SHACHAM, E., REECE, M., KROENKE, K., ONG'OR, W.O., OMOLLO, O., YEBEI, V.N. & OJWANG, C. 2009. Validity/reliability of PHQ-9 and PHQ-2 depression scales among adults living with HIV/AIDS in western Kenya. *Journal of general internal medicine.* 24(2):189-197.

MORI, H. & MISHINA, M. 1995. Structure and function of the NMDA receptor channel. *Neuropharmacology.* 34(10):1219-1237. DOI:https://doi.org/10.1016/0028-3908(95)00109-J.

MORIKI, K., TUSHIMA, E., OGIHARA, H., ENDO, R., SATO, T. & IKEMOTO, Y. 2022. Combined effects of lifestyle and psychosocial factors on central sensitization in patients with chronic low back pain: A cross-sectional study. Journal of Orthopaedic Science. 27(6):1185-1189. DOI:https://doi.org/10.1016/j.jos.2021.07.012.

MÜLLER, K. 2017. *Here: A simpler way to find your files.*

MÜLLER, K. & WICKHAM, H. 2022. *tibble: Simple Data Frames*. Available: https://CRAN.R-project.org/package=tibble [Available: https://CRAN.R-project.org/package=tibble].

MYER, L., SMIT, J., ROUX, L.L., PARKER, S., STEIN, D.J. & SEEDAT, S. 2008. Common mental disorders among HIV-infected individuals in South Africa: prevalence, predictors, and validation of brief psychiatric rating scales. *AIDS patient care and STDs.* 22(2):147-158.

NEL, A. & KAGEE, A. 2013. The relationship between depression, anxiety and medication adherence among patients receiving antiretroviral treatment in South Africa. *AIDS Care.* 25(8):948-955. DOI:10.1080/09540121.2012.748867.

NELSON, K.A., PARK, K.M., ROBINOVITZ, E., TSIGOS, C. & MAX, M.B. 1997. High-dose oral dextromethorphan versus placebo in painful diabetic neuropathy and postherpetic neuralgia. Neurology. 48(5):1212-1218

NIJS, J., PAUL VAN WILGEN, C., VAN OOSTERWIJCK, J., VAN ITTERSUM, M. & MEEUS, M. 2011. How to explain central sensitization to patients with 'unexplained' chronic musculoskeletal pain: Practice guidelines. *Manual Therapy.* 16(5):413-418. DOI:https://doi.org/10.1016/j.math.2011.04.005.

OKIFUJI, A., TURK, D.C. & CURRAN, S.L. 1999. Anger in chronic pain: investigations of anger targets and intensity. J Psychosom Res. 47(1):1-12. DOI:10.1016/s0022-3999(99)00006-9.

OOMS, J. 2021. *magick: Advanced Graphics and Image-Processing in R*. Available: https://CRAN.R-project.org/package=magick [Available: https://CRAN.R-project.org/package=magick].

OSSIPOV, M.H., DUSSOR, G.O. & PORRECA, F. 2010. Central modulation of pain. *J Clin Invest.* 120(11):3779-3787. DOI:10.1172/jci43766.

PARK, K.M., MAX, M.B., ROBINOVITZ, E., GRACELY, R.H. & BENNETT, G.J. 1995. Effects of intravenous ketamine, alfentanil, or placebo on pain, pinprick hyperalgesia, and allodynia produced by intradermal capsaicin in human subjects. *PAIN®.* 63(2):163-172.

PARKER, R., JELSMA, J. & STEIN, D.J. 2017. Pain in amaXhosa women living with HIV/AIDS: a cross-sectional study of ambulant outpatients. *BMC Women's Health.* 17(1):31. DOI:10.1186/s12905-017-0388-9.

PATIL, I. 2021. Visualizations with statistical details: The'ggstatsplot'approach. *Journal of open source software.* 6(61):3167. Available: https://doi.org/10.21105/joss.03167.

PEDERSEN, J.L., GALLE, T.S. & KEHLET, H. 1998. Peripheral analgesic effects of ketamine in acute inflammatory pain. *The Journal of the American Society of Anesthesiologists.* 89(1):58-66.

PEDERSEN, T.L. 2022. *patchwork: The Composer of Plots*. Available: https://CRAN.R-project.org/package=patchwork [Available: https://CRAN.R-project.org/package=patchwork].

PEFURA-YONE, E.W., SOH, E., KENGNE, A.P., BALKISSOU, A.D. & KUABAN, C. 2013. Non-adherence to antiretroviral therapy in Yaounde: prevalence, determinants and the concordance of two screening criteria. *Journal of infection and public health.* 6(4):307-315.

PETERS, J.L., SUTTON, A.J., JONES, D.R., ABRAMS, K.R. & RUSHTON, L. 2008. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *J Clin Epidemiol.* 61(10):991-996. DOI:10.1016/j.jclinepi.2007.11.010.

PETERSEN, K.K., ARENDT-NIELSEN, L., SIMONSEN, O., WILDER-SMITH, O. & LAURSEN, M.B. 2015. Presurgical assessment of temporal summation of pain predicts the development of chronic postoperative pain 12 months after total knee replacement. *Pain.* 156(1):55-61. DOI:10.1016/j.pain.0000000000000022.

PETRENKO, A.B., YAMAKURA, T., BABA, H. & SHIMOJI, K. 2003. The role of N-methyl-D-aspartate (NMDA) receptors in pain: a review. *Anesth Analg.* 97(4):1108-1116. DOI:10.1213/01.Ane.0000081061.12235.55.

PFAU, D.B., KLEIN, T., PUTZER, D., POGATZKI-ZAHN, E.M., TREEDE, R.D. & MAGERL, W. 2011. Analysis of hyperalgesia time courses in humans after painful electrical high-frequency stimulation identifies a possible transition from early to late LTP-like pain plasticity. *Pain.* 152(7):1532-1539. DOI:10.1016/j.pain.2011.02.037.

PILAT, D. & FUKASAKU, Y. 2007. OECD principles and guidelines for access to research data from public funding. *Data Science Journal.* 6:OD4-OD11.

POGATZKI-ZAHN, E.M., SEGELCKE, D. & SCHUG, S.A. 2017. Postoperative pain—from mechanisms to treatment. *Pain reports.* 2(2):e588. DOI:10.1097/pr9.0000000000000588.

PÖYHIÄ, R. & VAINIO, A. 2006. Topically administered ketamine reduces capsaicin-evoked mechanical hyperalgesia. *Clin J Pain.* 22(1):32-36. DOI:10.1097/01.ajp.0000149800.39240.95.

PUD, D., EISENBERG, E., SPITZER, A., ADLER, R., FRIED, G. & YARNITSKY, D. 1998. The NMDA receptor antagonist amantadine reduces surgical neuropathic pain in cancer patients: a double blind, randomized, placebo controlled trial. Pain. 75(2-3):349-354.

QUESADA, C., KOSTENKO, A., HO, I., LEONE, C., NOCHI, Z., STOUFFS, A., WITTAYER, M., CASPANI, O. et al. 2021. Human surrogate models of central sensitization: A critical review and practical guide. *Eur J Pain.* 25(7):1389-1428. DOI:10.1002/ejp.1768.

RAJA, S.N., CARR, D.B., COHEN, M., FINNERUP, N.B., FLOR, H., GIBSON, S., KEEFE, F.J., MOGIL, J.S. et al. 2020. The revised International Association for the Study of Pain definition of pain: concepts, challenges, and compromises. *Pain.* 161(9):1976-1982. DOI:10.1097/j.pain.0000000000001939.

RANDIC, M., JIANG, M. & CERNE, R. 1993. Long-term potentiation and long-term depression of primary afferent neurotransmission in the rat spinal cord. Journal of Neuroscience. 13(12):5228-5241.

RASMUSSEN, V., BORGEN, A., JANSEN, E., ROTBØLL NIELSEN, P. & WERNER, M. 2015. Hyperbaric oxygen therapy attenuates central sensitization induced by a thermal injury in humans. *Acta Anaesthesiologica Scandinavica.* 59(6):749-762.

RICHARDSON, J.L., HEIKES, B., KARIM, R., WEBER, K., ANASTOS, K. & YOUNG, M. 2009. Experience of pain among women with advanced HIV disease. *AIDS Patient Care STDS.* 23(7):503-511. DOI:10.1089/apc.2008.0128.

RODITI, D., ROBINSON, M.E. & LITWINS, N. 2009. Effects of coping statements on experimental pain in chronic pain patients. Journal of pain research.109-116.

ROSENFELD, B., BREITBART, W., MCDONALD, M.V., PASSIK, S.D., THALER, H. & PORTENOY, R.K. 1996. Pain in ambulatory AIDS patients. II: Impact of pain on psychological functioning and quality of life. *Pain.* 68(2-3):323-328. DOI:10.1016/s0304-3959(96)03220-4.

ROTHERAM-BORUS, M.J. 2000. Variations in perceived pain associated with emotional distress and social identity in AIDS. *AIDS Patient Care STDS.* 14(12):659-665. DOI:10.1089/10872910050206586.

RUPINSKI, M.T. & DUNLAP, W.P. 1996. Approximating Pearson product-moment correlations from Kendall's tau and Spearman's rho. *Educational and psychological measurement.* 56(3):419-429.

SABIN, C.A., HARDING, R., BAGKERIS, E., NKHOMA, K., POST, F.A., SACHIKONYE, M., BOFFITO, M., ANDERSON, J. et al. 2018. Pain in people living with HIV and its association with healthcare resource use, well being and functional status. *Aids.* 32(18). Available: .

SADAT-SHIRAZI, M.S., VOUSOOGHI, N., ALIZADEH, B., MAKKI, S.M., ZAREI, S.Z., NAZARI, S. & ZARRINDAST, M.R. 2018. Expression of NMDA receptor subunits in human blood lymphocytes: A peripheral biomarker in online computer game addiction. J Behav Addict. 7(2):260-268. DOI:10.1556/2006.7.2018.35.

SCHULTE, H., SOLLEVI, A. & SEGERDAHL, M. 2004. The synergistic effect of combined treatment with systemic ketamine and morphine on experimentally induced windup-like pain in humans. *Anesthesia & Analgesia.* 98(6):1574-1580.

SCHWARTZ, M. & MERCALDO, N. 2022. *BI: Blinding Assessment Indexes for Randomized, Controlled, Clinical Trials*. Available: https://CRAN.R-project.org/package=BI [2022].
SCHWARTZMAN, R.J., ALEXANDER, G.M., GROTHUSEN, J.R., PAYLOR, T., REICHENBERGER, E. & PERREAULT, M. 2009. Outpatient intravenous ketamine for the treatment of complex regional pain syndrome: a double-blind placebo controlled study. PAIN®. 147(1-3):107-115.
SETHNA, N.F., LIU, M., GRACELY, R., BENNETT, G.J. & MAX, M.B. 1998. Analgesic and cognitive effects of intravenous ketamine-alfentanil combinations versus either drug alone after intradermal capsaicin in normal subjects. *Anesthesia & Analgesia.* 86(6):1250-1256.
SHY, M.E., FROHMAN, E.M., SO, Y.T., AREZZO, J.C., CORNBLATH, D.R., GIULIANI, M.J., KINCAID, J.C., OCHOA, J.L. et al. 2003. Quantitative sensory testing. *Report of the Therapeutics and Technology Assessment Subcommittee of the American Academy of Neurology.* 60(6):898-904. DOI:10.1212/01.Wnl.0000058546.16985.11.
SIGNORELL, A. *Area Under the Curve*. Available: https://search.r-project.org/CRAN/refmans/DescTools/html/AUC.html [Available: https://search.r-project.org/CRAN/refmans/DescTools/html/AUC.html].
SIGNORELL, A. 2022. *DescTools: Tools for Descriptive Statistics*. Available: https://cran.r-project.org/package=DescTools [2022].
SLEPIAN, P.M., ANKAWI, B., HIMAWAN, L.K. & FRANCE, C.R. 2016. Development and initial validation of the pain resilience scale. The journal of pain. 17(4):462-472.
SLOWIKOWSKI , K. 2021. *ggrepel: Automatically Position Non-Overlapping Text Labels with 'ggplot2'*. Available: https://CRAN.R-project.org/package=ggrepel [Available:].
SMITH, B.W., TOOLEY, E.M., MONTAGUE, E.Q., ROBINSON, A.E., COSPER, C.J. & MULLINS, P.G. 2008. Habituation and sensitization to heat and cold pain in women with fibromyalgia and healthy controls. Pain. 140(3):420-428.
SODERBERG, C.K. 2018. Using OSF to share data: A step-by-step guide. Advances in Methods and Practices in Psychological Science. 1(1):115-120.
SPRUYT, A., CLARYSSE, J., VANSTEENWEGEN, D., BAEYENS, F. & HERMANS, D. 2010. Affect 4.0: a free software package for implementing psychological and psychophysiological experiments. *Exp Psychol.* 57(1):36-45. DOI:10.1027/1618-3169/a000005.
STATISTICS, S.J.C.T., SA: Strategic Development Information & DEPARTMENT, G. 2011. City of Cape Town–2011 census suburb overview.1-7.
STURGEON, J.A., DIXON, E.A., DARNALL, B.D. & MACKEY, S.C. 2015. Contributions of physical function and satisfaction with social roles to emotional distress in chronic pain: a Collaborative Health Outcomes Information Registry (CHOIR) study. Pain. 156(12):2627-2633. DOI:10.1097/j.pain.0000000000000313.
STURGEON, J.A. & ZAUTRA, A.J. 2010. Resilience: a new paradigm for adaptation to chronic pain. Current Pain and Headache Reports. 14:105-112.
SWEATT, J.D. 2008. 4.20 - The NMDA Receptor. In *Learning and Memory: A Comprehensive Reference.* J.H. Byrne, Ed. Oxford: Academic Press. 409-426. DOI:https://doi.org/10.1016/B978-012370509-9.00020-6.
TAKAHASHI, H., MIYAZAKI, M., NANBU, T., YANAGIDA, H. & MORITA, S. 1998. The NMDA-receptor antagonist ketamine abolishes neuropathic pain after epidural administration in a clinical case. Pain. 75(2-3):391-394.
TEZEL, Ö., TİRYAKİ, B., ÖZKUL, E. & KESEMEN, O. 2021. A New Goodness-of-Fit Test: Free Chi-Square (FCS). *GAZI UNIVERSITY JOURNAL OF SCIENCE.* 34:879-897. DOI:10.35378/gujs.743444.
TREEDE, R.-D. 2016. Neurogenic hyperalgesia: illuminating its mechanisms with an infrared laser. *The Journal of physiology.* 594(22):6441-6442. DOI:10.1113/JP273072.
TREEDE, R.-D. & MAGERL, W. 2000. Multiple mechanisms of secondary hyperalgesia. *Progress in brain research.* 129:331-341.
TURK, D.C., FILLINGIM, R.B., OHRBACH, R. & PATEL, K.V. 2016. Assessment of psychosocial and functional impact of chronic pain. *The journal of pain.* 17(9):T21-T49.

USHIDA, T., TANI, T., KANBARA, T., ZINCHUK, V.S., KAWASAKI, M. & YAMAMOTO, H. 2002. Analgesic effects of ketamine ointment in patients with complex regional pain syndrome type 1. Regional Anesthesia & Pain Medicine. 27(5):524-528.

VAN DEN BROEKE, E., GEENE, N., VAN RIJN, C., WILDER-SMITH, O. & OOSTERMAN, J. 2014a. Negative expectations facilitate mechanical hyperalgesia after high-frequency electrical stimulation of human skin. *European Journal of Pain.* 18(1):86-91.

VAN DEN BROEKE, E.N. & MOURAUX, A.J.J.o.n. 2014b. High-frequency electrical stimulation of the human skin induces heterotopical mechanical hyperalgesia, heat hyperalgesia, and enhanced responses to nonnociceptive vibrotactile input. 111(8):1564-1573.

VAN DEN BROEKE, E.N., GOUSSET, S., BOUVY, J., STOUFFS, A., LEBRUN, L., VAN NEERVEN, S.G.A. & MOURAUX, A. 2019. Heterosynaptic facilitation of mechanical nociceptive input is dependent on the frequency of conditioning stimulation. *Journal of Neurophysiology.* 122(3):994-1001. DOI:10.1152/jn.00274.2019.

VAN MIDDENDORP, H., LUMLEY, M.A., JACOBS, J.W., BIJLSMA, J.W. & GEENEN, R. 2010. The effects of anger and sadness on clinical pain reports and experimentally-induced pain thresholds in women with and without fibromyalgia. Arthritis Care Res (Hoboken). 62(10):1370-1376. DOI:10.1002/acr.20230.

VENTEVOGEL, P., DE VRIES, G., SCHOLTE, W.F., SHINWARI, N.R., FAIZ, H., NASSERY, R., VAN DEN BRINK, W. & OLFF, M. 2007. Properties of the Hopkins Symptom Checklist-25 (HSCL-25) and the Self-Reporting Questionnaire (SRQ-20) as screening instruments used in primary care in Afghanistan. *Soc Psychiatry Psychiatr Epidemiol.* 42(4):328-335. DOI:10.1007/s00127-007-0161-8.

VERHOEVEN, K., CHRISTOPHER, E., DIMITRI, M.V.R., VALÉRY, L. & GEERT, C. 2011. Distraction from pain and executive functioning: an experimental investigation of the role of inhibition, task switching and working memory. European Journal of Pain. 15(8):866-873.

VERONIKI, A.A., JACKSON, D., VIECHTBAUER, W., BENDER, R., BOWDEN, J., KNAPP, G., KUSS, O., HIGGINS, J.P. et al. 2016. Methods to estimate the between-study variance and its uncertainty in meta-analysis. *Res Synth Methods.* 7(1):55-79. DOI:10.1002/jrsm.1164.

VIECHTBAUER, W. 2005. Bias and Efficiency of Meta-Analytic Variance Estimators in the Random-Effects Model. *Journal of Educational and Behavioral Statistics.* 30(3):261-293. DOI:10.3102/10769986030003261.

VO, L. & DRUMMOND, P.D. 2013. High frequency electrical stimulation concurrently induces central sensitization and ipsilateral inhibitory pain modulation. *European Journal of Pain.* 17(3):357-368.

WALLACE, M.S., BRAUN, J. & SCHULTEIS, G. 2002. Postdelivery of Alfentanil and Ketamine Has No Effect on Intradermal Capsaicin–Induced Pain and Hyperalgesia. *The Clinical journal of pain.* 18(6):373-379.

WALLACE, M.S., RIDGEWAY III, B., LEUNG, A., SCHULTEIS, G. & YAKSH, T.L. 2002. Concentration-effect relationships for intravenous alfentanil and ketamine infusions in human volunteers: effects on acute thresholds and capsaicin-evoked hyperpathia. *The Journal of Clinical Pharmacology.* 42(1):70-80.

WARNCKE, T., STUBHAUG, A. & JØRUM, E. 1997. Ketamine, an NMDA receptor antagonist, suppresses spatial and temporal properties of burn-induced secondary hyperalgesia in man: a double-blind, cross-over comparison with morphine and placebo. *Pain.* 72(1-2):99-106.

WARNCKE, T., JØRUM, E. & STUBHAUG, A. 1997. Local treatment with the N-methyl-D-aspartate receptor antagonist ketamine, inhibit development of secondary hyperalgesia in man by a peripheral action. *Neuroscience letters.* 227(1):1-4.

WARNCKE, T., STUBHAUG, A. & JØRUM, E. 2000. Preinjury treatment with morphine or ketamine inhibits the development of experimentally induced secondary hyperalgesia in man. *Pain.* 86(3):293-303.

WICKHAM, H. 2022. *stringr: Simple, Consistent Wrappers for Common String Operations*. Available: https://CRAN.R-project.org/package=stringr [Available: https://CRAN.R-project.org/package=stringr].
WICKHAM, H. & GIRLICH, M. 2022. *tidyr: Tidy Messy Data*. Available: https://CRAN.R-project.org/package=tidyr [Available: https://CRAN.R-project.org/package=tidyr].
WICKHAM, H., HESTER, J. & BRYAN, J. 2022. *Readr: Read rectangular text data.*
WICKHAM, H., FRANÇOIS, R., HENRY , L. & MÜLLER, K. 2022. *dplyr: A Grammar of Data Manipulation*. Available: https://CRAN.R-project.org/package=dplyr [Available: https://CRAN.R-project.org/package=dplyr].
WICKHAM, H., BRYAN, J., KALICINSKI, M., VALERY, K., LEITIENNE, C., COLBERT, B., HOERL, D., MILLER, E. et al. 2019a. Package 'readxl'. *Version, 1.3.* 1.
WICKHAM, H., AVERICK, M., BRYAN, J., CHANG, W., MCGOWAN, L.D.A., FRANÇOIS, R., GROLEMUND, G., HAYES, A. et al. 2019b. Welcome to the Tidyverse. *Journal of open source software.* 4(43):1686.
WILLIS, W.D. 2009. The role of TRPV1 receptors in pain evoked by noxious thermal and chemical stimuli. *Experimental Brain Research.* 196(1):5-11. DOI:10.1007/s00221-009-1760-2.
WOO, A.K. 2010. Depression and Anxiety in Pain. *Reviews in pain.* 4(1):8-12. DOI:10.1177/204946371000400103.
WOOLF, C.J. 2014. What to call the amplification of nociceptive signals in the central nervous system that contribute to widespread pain? *Pain.* 155(10):1911-1912. DOI:10.1016/j.pain.2014.07.021.
WOOLF, C.J., BENNETT, G.J., DOHERTY, M., DUBNER, R., KIDD, B., KOLTZENBURG, M., LIPTON, R., LOESER, J.D. et al. 1998. Towards a mechanism-based classification of pain? *Pain.* 77(3):227-229. DOI:10.1016/s0304-3959(98)00099-2.
WORLDMEDICAL, A. 2019. *Handbook of WMA Policies*. Available: https://www.wma.net/policy/hb-e-version-2019-2/ [Available: https://www.wma.net/policy/hb-e-version-2019-2/].
XIE, Y. 2022. *formatR: Format R Code Automatically*. Available: https://CRAN.R-project.org/package=formatR [2022].
YASUMOTO, A. 2022. *ftExtra: Extensions for 'Flextable'*. Available: https://CRAN.R-project.org/package=ftExtra [Available: https://CRAN.R-project.org/package=ftExtra].
YOU, D.S., CREECH, S.K. & MEAGHER, M.W. 2016. Enhanced Area of Secondary Hyperalgesia in Women with Multiple Stressful Life Events: A Pilot Study. *Pain Medicine.* 17(10):1859-1864. DOI:10.1093/pm/pnw049.
YOU, D.S., CREECH, S.K., VICHAYA, E.G., YOUNG, E.E., SMITH, J.S. & MEAGHER, M.W. 2014. Effect of written emotional disclosure on secondary hyperalgesia in women with trauma history. *Psychosom Med.* 76(5):337-346. DOI:10.1097/psy.0000000000000064.
YUNUSA, M.A., NJOKU, C.H. & OBEMBE, A. 2014. Self-reported adherence to treatment: a study of socioeconomic factors and psychiatric morbidity among male and female patients with HIV infection in Sokoto, Nigeria. *Nigerian Journal of Medicine.* 23(1):33-39.
ZENG, C., GUO, Y., HONG, Y.A., GENTZ, S., ZHANG, J., ZHANG, H., QIAO, J., XU, Z. et al. 2019. Differential effects of unemployment on depression in people living with HIV/AIDS: a quantile regression approach. *AIDS Care.* 31(11):1412-1419. DOI:10.1080/09540121.2019.1587366.
ZHANG, C., LI, X., LIU, Y., ZHOU, Y., SHEN, Z. & CHEN, Y. 2018. Impacts of HIV Stigma on Psychosocial Well-Being and Substance Use Behaviors Among People Living With HIV/AIDS In China: Across the Life Span. *AIDS education and prevention : official publication of the International Society for AIDS Education.* 30(2):108-119. DOI:10.1521/aeap.2018.30.2.108.
ZHOU, H.-Y., CHEN, S.-R., BYUN, H.-S., CHEN, H., LI, L., HAN, H.-D., LOPEZ-BERESTEIN, G., SOOD, A.K. et al. 2012. N-methyl-D-aspartate receptor-and calpain-mediated proteolytic cleavage of K+-Cl− cotransporter-2 impairs spinal chloride homeostasis in neuropathic pain. Journal of Biological Chemistry. 287(40):33853-33864.

ZHU, H. 2019. KableExtra: Construct complex table with'kable'and pipe syntax. *R package version.* 1(0). Available: https://CRAN.R-project.org/package=kableExtra.
ZHUO, M. 2017. Ionotropic glutamate receptors contribute to pain transmission and chronic pain. Neuropharmacology. 112(Pt A):228-234. DOI:10.1016/j.neuropharm.2016.08.014.
ZIMMER, Z., FRASER, K., GROL-PROKOPCZYK, H. & ZAJACOVA, A. 2022. A global study of pain prevalence across 52 countries: examining the role of country-level contextual factors. Pain. 163(9):1740-1750. DOI:10.1097/j.pain.0000000000002557.

SUPPLEMENTARY INFORMATION

Chapter 3: model assumptions for primary outcome

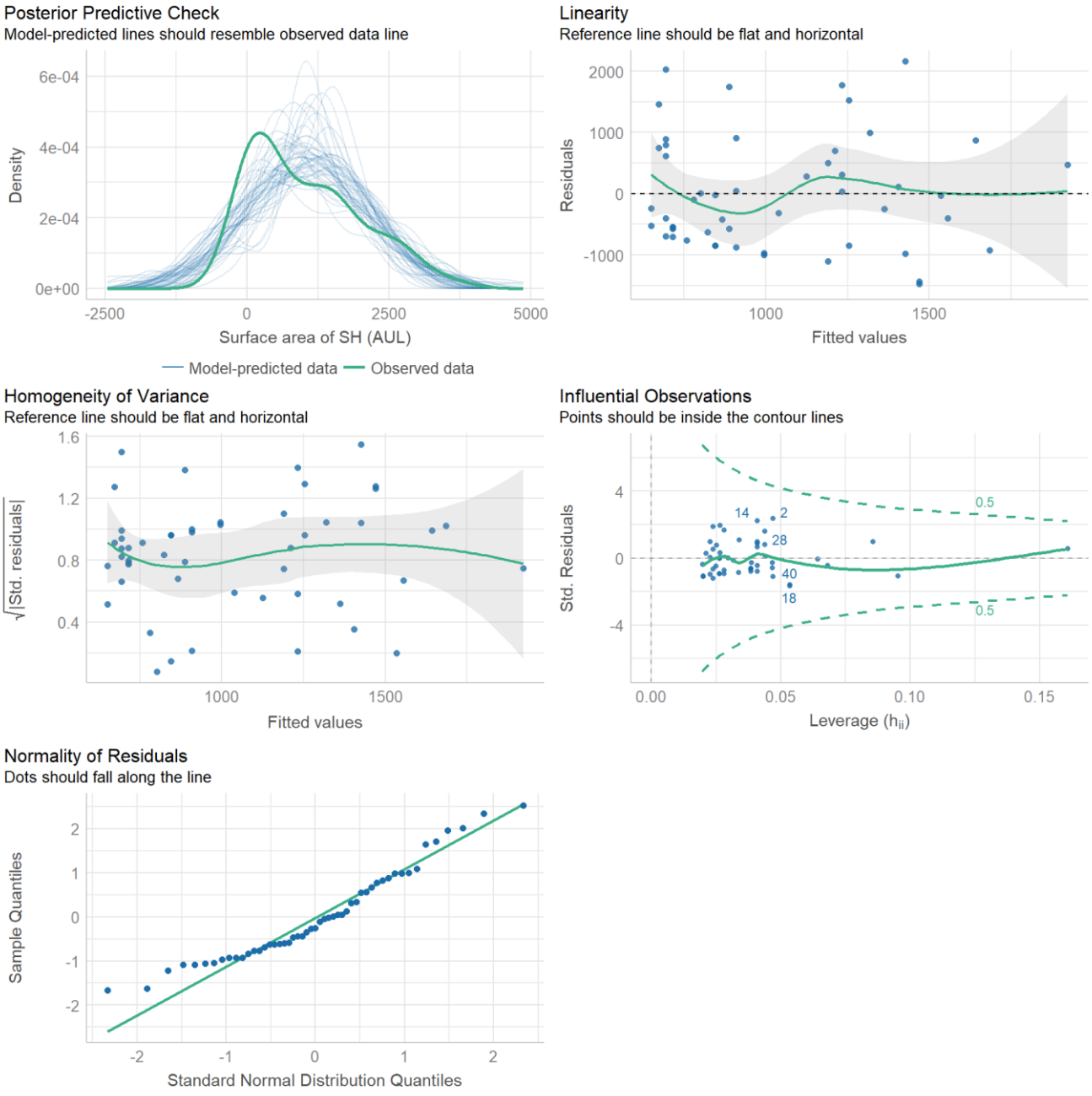


Figure S1*: Model 1 assumptions for the primary analysis. The following model assumptions were met; posterior predictive check, linearity, homogeneity of variance, and influential observations; however, the assumption of normality of residuals was not met.*

Chapter 3: model assumptions for secondary outcome

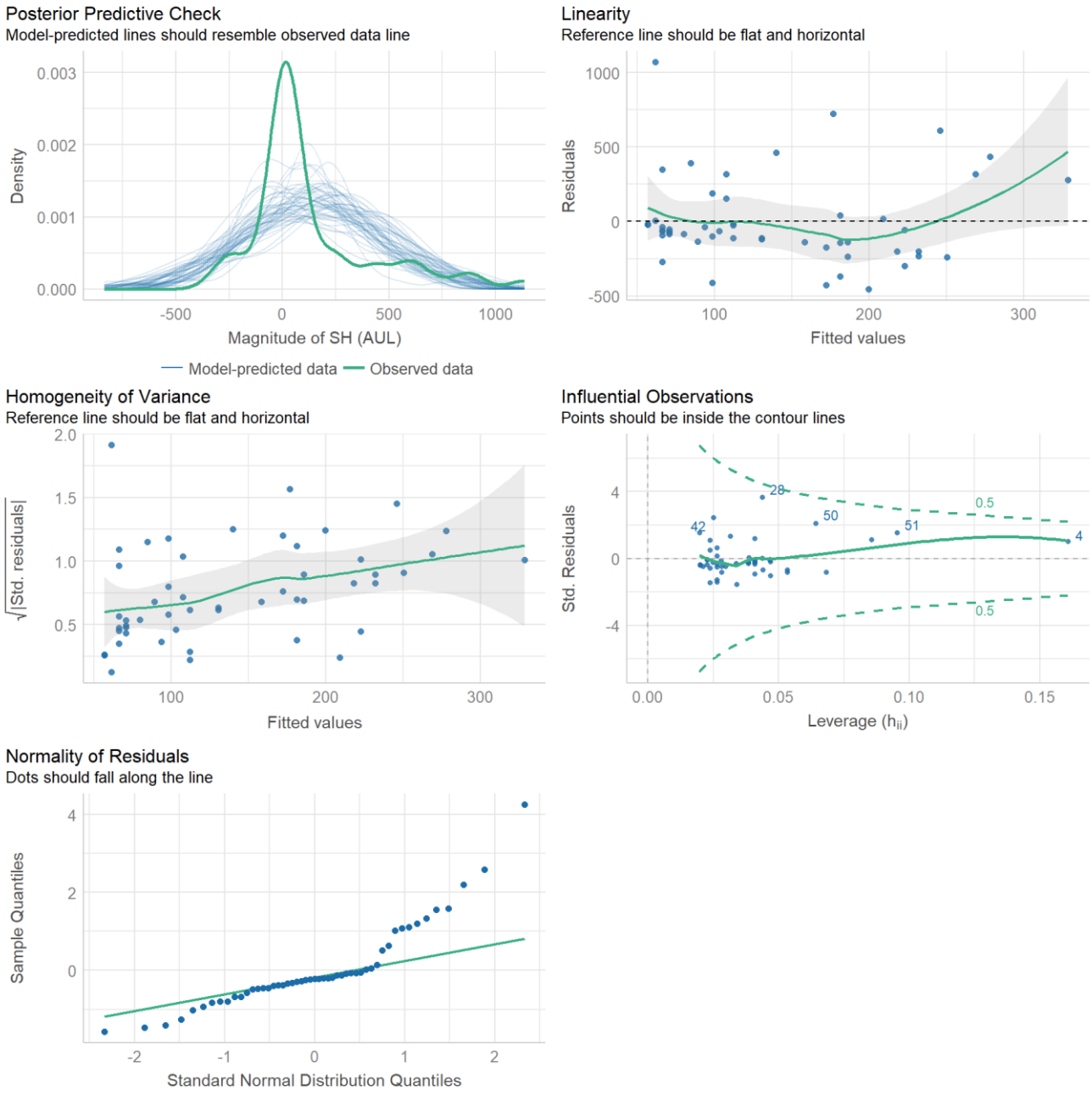


Figure S2*: Model 1 assumptions for the secondary analysis. Four (posterior predictive check, linearity, homogeneity of variance, and normality of residuals) of five assumptions were violated by Model 1. The assumption of influential observations was deemed to have been met.*

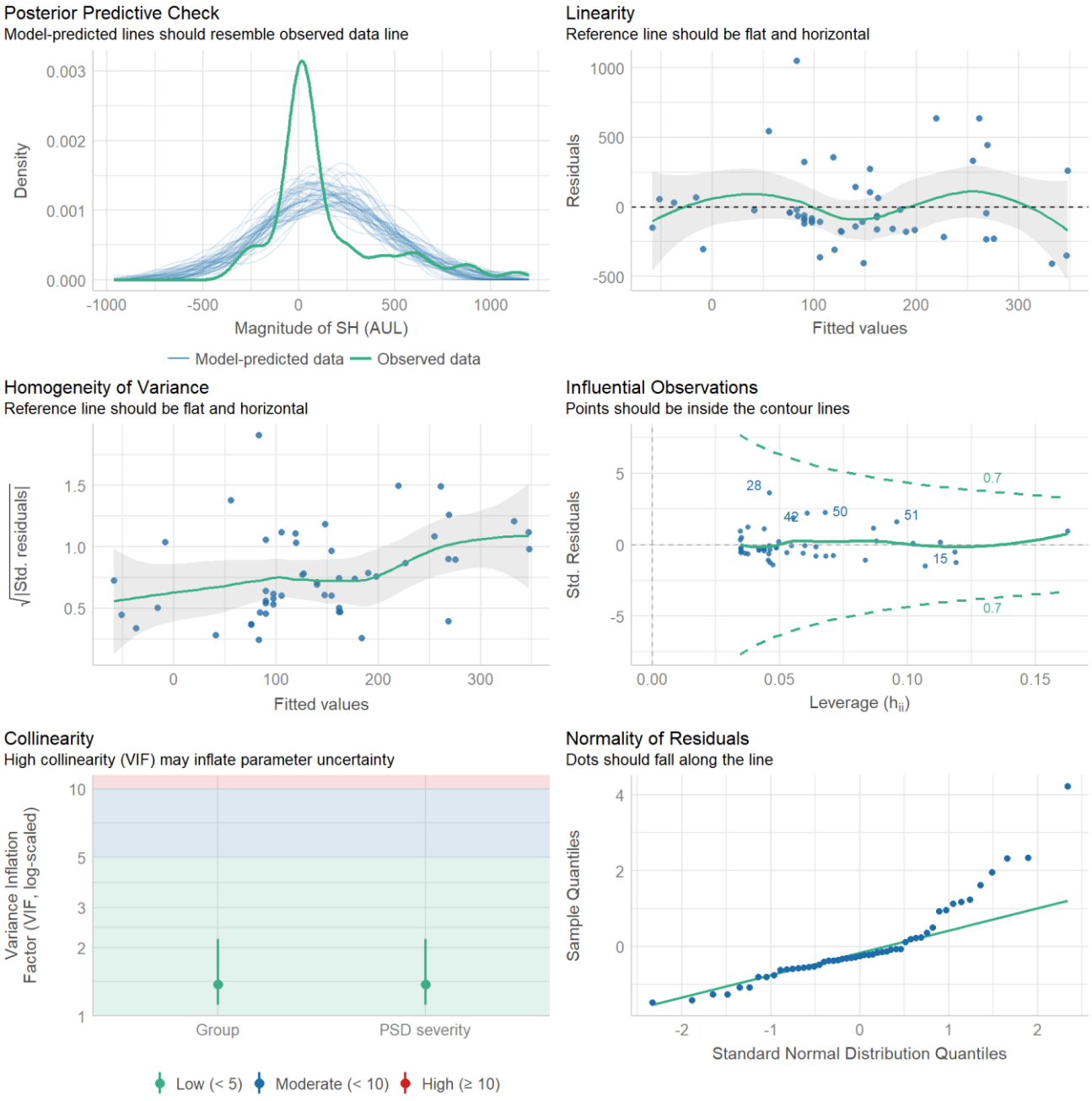


***Figure S3**: Model 2 assumptions for the secondary analysis. Three (posterior predictive check, homogeneity of variance, and normality of residuals) of six assumptions were violated by Model 2. The assumptions of linearity, collinearity and influential observations were deemed to have been met.*

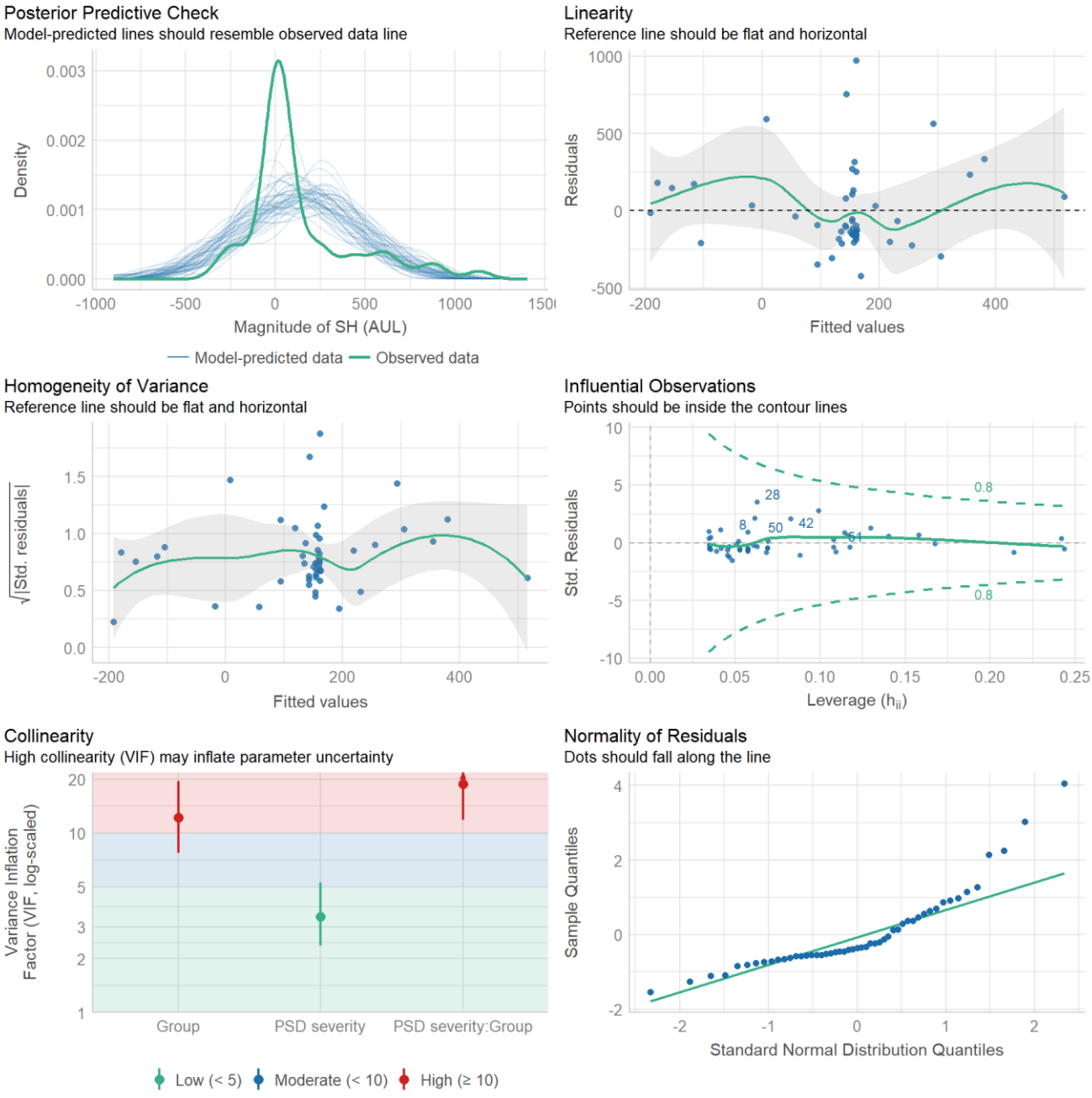


Figure S4*: Model 3 assumptions for the secondary analysis. Five (posterior predictive check, homogeneity of variance, linearity, collinearity and normality of residuals) of six assumptions were violated by model 2, and the assumption of influential observations was deemed to have been met.*

APPENDICES

Appendix A: eligibility criteria and grouping table

	Inclusion	Exclusion	
Participants	Only pain-free, healthy humans included	Included animals OR people with pain	
Study design	Used an experimental procedure with the aim of inducing AND manipulating secondary hyperalgesia (identifiable goal AND site AND induction procedure AND manipulation procedure)	Review (set aside for cross-checking) OR Not an experimental procedure OR No identifiable manipulation procedure	
Outcomes	Pain or sensitivity to provocation assessed subsequent to induction AND manipulation Acceptable: pain yes or no, self-report of intensity, quality, pain threshold.	Subjective ratings not provided Unacceptable: facial expression, physical behaviour measurement, or psychophysiology in absence of self-report	
Include?	Tick in *every* box above: include ☐	Tick in *any* box above: exclude ☐	Review ☐

Appendix B: risk of bias assessment tool

Article ID:	Reviewer:	
Selection bias		
	Decision	Justification
Was the sampling/recruitment strategy appropriate to minimise bias?	☐ Yes ☐ No ☐ Unclear	
Was it clearly and appropriately determined that participants were pain-free?	☐ Yes ☐ No ☐ Unclear	
[B-G only] Similar baseline demographics among participants (age/sex/medical/psychological state)?	☐ Yes ☐ No ☐ Unclear	
[Psych manip] Neutral psych status?	☐ Yes ☐ No ☐ Unclear	
[B-G only] Random allocation [B-site] Random allocation	☐ Yes ☐ No ☐ Unclear	
Risk of selection bias summary	☐ High (failure to include any of the above probably influenced results FOR THE QUESTION OF THIS REVIEW) ☐ Low (results unlikely to have been influenced) ☐ Unclear (not enough information)	

Performance bias		
Blinding	Decision	Justification
Were participants blinded to the research question and paradigm and [if relevant] group allocation?	☐ Yes ☐ No ☐ Unclear	
Risk of performance bias summary	☐ High ☐ Low ☐ Unclear	
Detection bias		
Were outcome assessors blinded to the research question and paradigm?	☐ Yes ☐ No ☐ Unclear	
Were analysing researchers blinded to the group allocation of participants and/or to site allocation?	☐ Yes ☐ No ☐ Unclear	
Risk of detection bias summary	☐ High ☐ Low ☐ Unclear	

Comments:

Manipulation veracity		
[Psych] Did a manipulation check confirm the effectiveness of the manipulation?	□ Yes □ No □ Unclear	
Risk of manipulation veracity problem	□ High □ Low □ Unclear	
Attrition bias		
Incomplete outcome data	Decision	Justification
Have attrition/exclusions/ withdrawals been reported and appropriately dealt with in analysis?	□ Yes □ No □ Unclear	
Risk of attrition bias summary	□ High □ Low □ Unclear	
Measurement bias		
	Decision	Justification
Were valid and reliable outcome measurements used to assess severity & SA of secondary hyperalgesia?	2H: □ Yes □ No □ Unclear SA: □ Yes □ No □ Unclear	
Were identical equipment items used for measurements between groups/sites/time points?	2H: □ Yes □ No □ Unclear SA: □ Yes □ No □ Unclear	
Did the same assessor conduct assessments between groups/sites/time points?	2H: □ Yes □ No □ Unclear SA: □ Yes □ No □ Unclear	
Risk of measurement bias summary	2H: □ Yes □ No □ Unclear SA: □ Yes □ No □ Unclear	
Reporting bias		
Selective reporting	Decision	Justification

Were all outcomes for experimental and control groups reported on?	□ Yes □ No	
Were conflicts of interest and funding sources declared?	□ Yes □ No	
Risk of reporting bias summary	□ High □ Low □ Unclear	

Risk of bias summary

Risk of bias	Description	Study bias outcome
High risk of bias	Plausible bias that seriously weakens confidence in the results.	
Low risk of bias	Plausible bias unlikely to seriously alter or diminish trust in the results.	
Unclear risk of bias	Insufficient information available to make a judgement.	

Article ID:	Reviewer:	
Selection bias		
	Decision	Justification
Was the sampling/recruitment strategy appropriate to minimise bias?	□ Yes □ No □ Unclear	Yes: general population or subgroup. Convenience sampling is acceptable as long as eligibility criteria do not restrict to a certain group that could plausibly respond differently to the induction. No: group selected on basis of particular feature (e.g. high catastrophising positive affect / athletes in training)
Was it clearly and appropriately determined that participants were pain-free?	□ Yes □ No □ Unclear	Yes: participant self-report of no pain at time of testing AND no history of chronic pain (pain on most days for > 3 mo) in preceding 2 years. No: reports failure to ask BOTH questions. Unclear: does not report asking both questions.
[B-G only] Similar baseline demographics among participants (age/sex/medical/psychological state)?	□ Yes □ No □ Unclear	Yes: Psych (trauma Hx, stress status, general affect, sex, age, medication variables accounted for and similar) No: Psychiatric diagnoses or medication use (esp analgesics/anti-inflammatories/SNRI, etc) amongst participants. Unclear: not reported *Consider design features, e.g. within-subject control or pre-post design
[Psych manip] Neutral psych status?	□ Yes □ No □ Unclear	Yes: Psych variables accounted for and normal No: selected for responses on psych assessment
[B-G only] Random allocation [B-site] Random allocation	□ Yes □ No □ Unclear	Yes: random sequence generation / roll of die / other truly random procedure named No: counterbalancing of group size (i.e. pseudo-randomisation)[but consider ROB in context] / sequential allocation Unclear: not reported in enough detail to allow decision
Risk of selection bias summary	□ High □ Low □ Unclear	(failure to include any of the above probably influenced results FOR THE QUESTION OF THIS REVIEW) (results unlikely to have been influenced) (not enough information)

Guide to decision-making for risk of bias assessment

Performance bias		
Blinding	Decision	Justification
Were participants blinded to the research question and paradigm and [if relevant] group allocation?	☐ Yes ☐ No ☐ Unclear	Yes: evidence provided - blinding strategy AND blinding check AND results reported AND analysis done accordingly No: Blinding reported broken Unclear: not enough information / failure to report)
Risk of performance bias summary	☐ High ☐ Low ☐ Unclear	High: Plausible doubt that participant blinding was applied and maintained throughout Low: Confident that participant blinding was applied and maintained throughout Unclear: not enough information to make informed judgement (e.g. blinding strategy AND blinding check AND results mentioned BUT not fully reported)
Detection bias		
Were outcome assessors blinded to the research question and paradigm?	☐ Yes ☐ No ☐ Unclear	Yes: evidence provided - blinding strategy AND blinding check AND results reported AND analysis done accordingly No: Blinding reported broken Unclear: not enough information / failure to report)
Were analysing researchers blinded to the group allocation of participants and/or to site allocation?	☐ Yes ☐ No ☐ Unclear	Yes: evidence provided - blinding strategy AND blinding check AND results reported AND analysis done accordingly No: Blinding reported broken Unclear: not enough information / failure to report)
Risk of detection bias summary	☐ High ☐ Low ☐ Unclear	High: Plausible doubt that participant blinding was applied and maintained throughout Low: Confident that participant blinding was applied and maintained throughout Unclear: not enough information to make informed judgement (e.g. blinding strategy AND blinding check AND results mentioned BUT not fully reported)

Risk of manipulation veracity problem		
[Psych] Did a manipulation check confirm the effectiveness of the manipulation?	☐ Yes ☐ No ☐ Unclear	Yes: manip check done and results reported and confirmed effectiveness No: no manipulation check done OR manip check done but results not reported. Unclear: manip check done and results confirmed ineffectiveness or were inconclusive

Risk of manipulation veracity problem	□ High □ Low □ Unclear	
Attrition bias		
Incomplete outcome data	Decision	Justification
Have attrition/exclusions/ withdrawals been reported and appropriately dealt with in analysis?	□ Yes □ No □ Unclear	Yes: no attrition/withdrawals OR stats handled withdrawals appropriately AND relevant adverse events reported
Risk of attrition bias summary	□ High □ Low □ Unclear	
Measurement bias		
	Decision	Justification
Were valid and reliable outcome measurements used to assess severity & SA of secondary hyperalgesia?	2H: □ Yes □ No □ Unclear SA: □ Yes □ No □ Unclear	Yes: Self-report: VAS / NRS / validated scale Surface area: independently duplicated measurements or validated approach Consider test-retest reliability if relevant No: single measurement of distance/SA; un-validated self-report scale
Were identical equipment items used for measurements between groups/sites/time points?	2H: □ Yes □ No □ Unclear SA: □ Yes □ No □ Unclear	
Did the same assessor conduct assessments between groups/sites/time points?	2H: □ Yes □ No □ Unclear SA: □ Yes	

	□ No □ Unclear	
Risk of measurement bias summary	2H: □ Yes □ No □ Unclear SA: □ Yes □ No □ Unclear	

Reporting bias		
Selective reporting	Decision	Justification
Were all outcomes for experimental and control groups reported on?	□ Yes □ No	Check each outcome (compare methods vs results)
Were conflicts of interest and funding sources declared?	□ Yes □ No	Consider relevant conflicts
Risk of reporting bias summary	□ High □ Low □ Unclear	

Risk of bias summary		
Risk of bias	Description	Study bias outcome
High risk of bias	Plausible bias that seriously weakens confidence in the results.	
Low risk of bias	Plausible bias unlikely to seriously alter or diminish trust in the results.	
Unclear risk of bias	Insufficient information available to make a judgement.	

Appendix C: meta-analysis script

Luyanduthando Mqadi, Tory Madden and Peter Kamerman

12 February, 2023

Contents


```
## knitr setup

knitr::opts_chunk$set(
  tidy.opts = list(
    tidy = TRUE,
    warning = FALSE,
    message = FALSE,
    fig.align = 'center',
    fig.path = 'figures',
    fig.retina = 1
  )
)

## Load packages
```

```
## Warning: package 'dplyr' was built under R version 4.2.1

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##       filter, lag

## The following objects are masked from 'package:base':
##
##       intersect, setdiff, setequal, union
```

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 4.2.1

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.2 ──

## ✔ ggplot2 3.3.6     ✔ purrr   0.3.5
## ✔ tibble  3.1.8     ✔ stringr 1.4.1
## ✔ tidyr   1.2.1     ✔ forcats 0.5.2
## ✔ readr   2.1.3

## Warning: package 'tibble' was built under R version 4.2.1

## Warning: package 'tidyr' was built under R version 4.2.1

## Warning: package 'readr' was built under R version 4.2.1

## Warning: package 'purrr' was built under R version 4.2.1

## Warning: package 'stringr' was built under R version 4.2.1

## Warning: package 'forcats' was built under R version 4.2.1

## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
```

```
library(readxl)
```

```
## Warning: package 'readxl' was built under R version 4.2.1
```

```
library(gridExtra)
```

```
##
## Attaching package: 'gridExtra'##
## The following object is masked from 'package:dplyr':##
##       combine
```

```
library(readr)
library(here)
```

here() starts at C:/Users/user/Desktop/systematic-review

```
library(kableExtra)
```

Warning in !is.null(rmarkdown::metadata$output) && rmarkdown::metadata$output## %in% :
'length(x) = 3 > 1' in coercion to 'logical(1)'

##
Attaching package: 'kableExtra'##
The following object is masked from 'package:dplyr':
##
group_rows

```
library(ggstatsplot)
```

You can cite this package as:
Patil, I. (2021). Visualizations with statistical details: The 'ggstatsplot' approach.##Journal of Open Source Software, 6(61), 3167, doi:10.21105/joss.03167

```
library(formatR)
library(forestplot)
```

Warning: package 'forestplot' was built under R version 4.2.1

Loading required package: grid
Loading required package: checkmate##
Loading required package: abind

```
library(estmeansd)
```

Warning: package 'estmeansd' was built under R version 4.2.1

```
library(meta)
```

Warning: package 'meta' was built under R version 4.2.1## Loading

'meta' package (version 6.0-0).
Type 'help(meta)' for a brief overview.
Readers of 'Meta-Analysis with R (Use R!)' should install
older version of 'meta' package: https://tinyurl.com/dt4y5drs

```
library(dmetar)
```

Extensive documentation for the dmetar package can be found at:##
www.bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/

```
library(robvis)
```

Warning: package 'robvis' was built under R version 4.2.1

```
library(ftExtra)
```

Warning: package 'ftExtra' was built under R version 4.2.1

```
library(tidyr)
library(stringr)
library(tibble)
library(flextable)
```

Warning: package 'flextable' was built under R version 4.2.1

```
##
## Attaching package: 'flextable'## 
## The following object is masked from 'package:ftExtra':
##
##     separate_header## 
## The following objects are masked from 'package:kableExtra':
##
##     as_image, footnote## 
## The following object is masked from 'package:purrr':
##
##     compose
```

```
library(magrittr)
```

```
##
## Attaching package: 'magrittr'## 
## The following object is masked from 'package:purrr':
##
##     set_names
##
## The following object is masked from 'package:tidyr':
##
##     extract
```

```
library(patchwork)
```

Warning: package 'patchwork' was built under R version 4.2.1

```
library(magick )
```

Warning: package 'magick' was built under R version 4.2.2## Linking

to ImageMagick 6.9.12.3
Enabled features: cairo, freetype, fftw, ghostscript, heic, lcms, pango, raw, rsvg, webp## Disabled features: fontconfig, x11

Forest plots

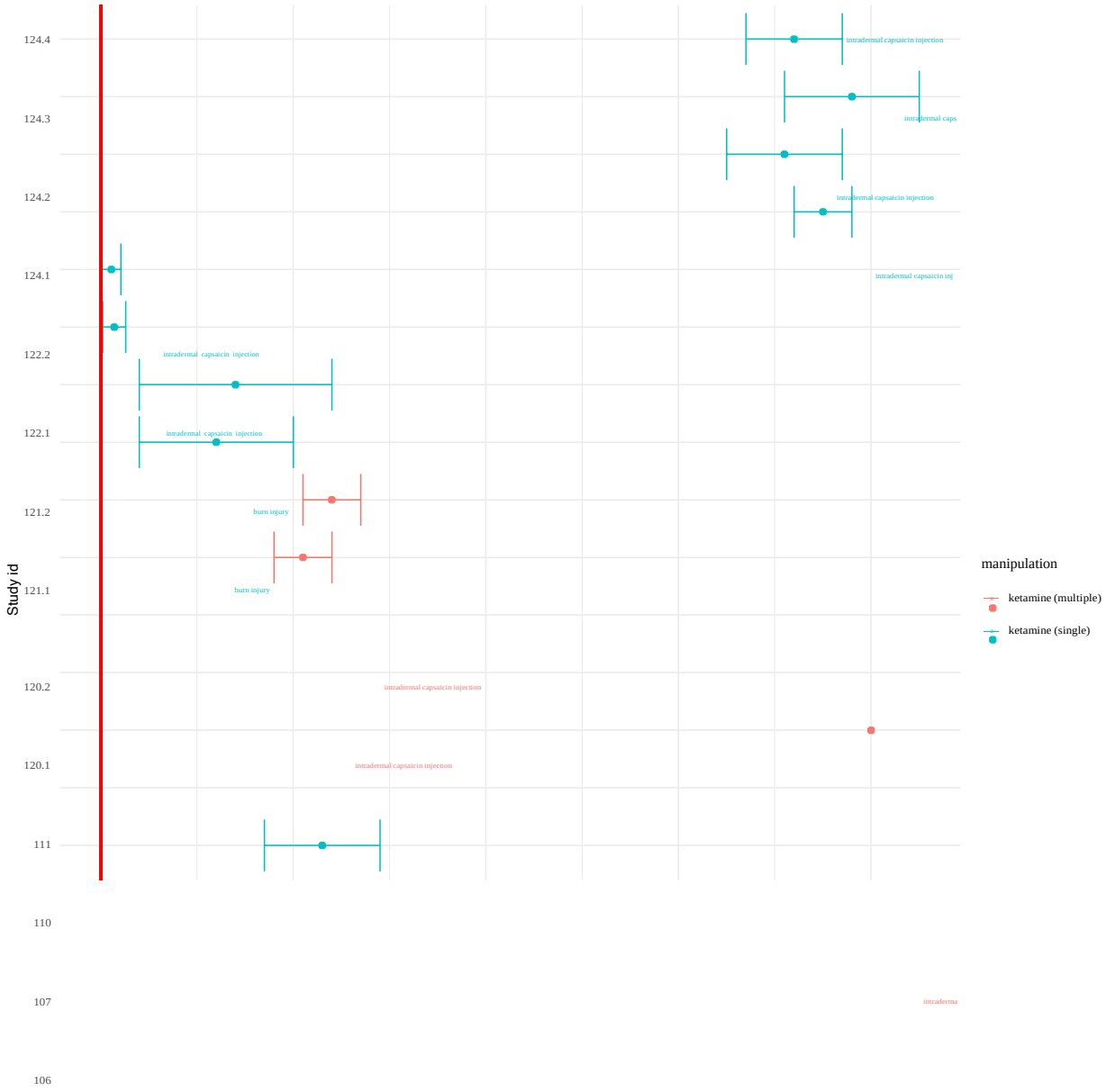
Magnitude of SH

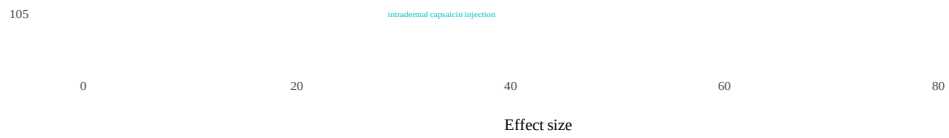
```
## Import mag data
mag <- read_excel(path =   here("sr data/mag_area_bl_fu.xlsx"),sheet =
                            "mag") %>%
   filter(!manipulation == "neramaxane") %>% ## 1 study excluded from meta-analysis

   mutate(m.baseline.CTM = as.numeric(m.baseline.CTM)) %>% mutate(c.baseline.CTM
   = as.numeric(c.baseline.CTM)) %>% mutate(m.baseline.upper.bound =
   as.numeric(m.baseline.upper.bound))   %>%mutate(c.baseline.upper.bound =
   as.numeric(c.baseline.upper.bound)) %>%mutate(m.baseline.lower.bound =
   as.numeric(m.baseline.lower.bound)) %>%mutate(c.baseline.lower.bound =
   as.numeric(c.baseline.lower.bound))   %>%mutate(m.CTM = as.numeric(m.CTM)) %>%
   mutate(c.CTM = as.numeric(c.CTM)) %>%
   mutate(m.upper.bound = as.numeric(m.upper.bound)) %>%
   mutate(c.upper.bound = as.numeric(c.upper.bound)) %>%
   mutate(m.lower.bound = as.numeric(m.lower.bound)) %>%
   mutate(c.lower.bound = as.numeric(c.lower.bound))

## create plot (manipulation)
ggplot(data = mag,
   aes(y = study_id,
         x = m.CTM,
         group = manipulation,fill =
         manipulation, colour =
         manipulation,
         scale_x_discrete(limits = c(-20, 100)))) + geom_point(size =
   2, aes(group = manipulation)) +geom_errorbar(aes(
      y = study_id,
      height = .1,
      xmax = m.upper.bound, xmin
      = m.lower.bound)) +
   geom_text(aes(
      label = induction,
      x = as.numeric(m.CTM)),
      hjust = -0.5,
      size = 2) +
   labs(title = 'Magnitude of SH of all studies that gave an induction and administered a pharmacologica
      subtitle = 'Data in each study is expressed by different measures of central tendency. The text in
```

```
    x = 'Effect size',
    y = 'Study id') +
  geom_vline(
    xintercept = 0,
    color = 'red',
    linetype = 'dashed',
    alpha = .5,
    size = 1
  ) +
  theme_minimal()
```

Magnitude of SH of all studies that gave an induction and administered a pharmacological manipulation to manipulate SH

Data in each study is expressed by different measures of central tendency. The text in different colours shows the induction which is coloured coded by the type ofmanipuation received. Dotted vertical line is the point of no effect





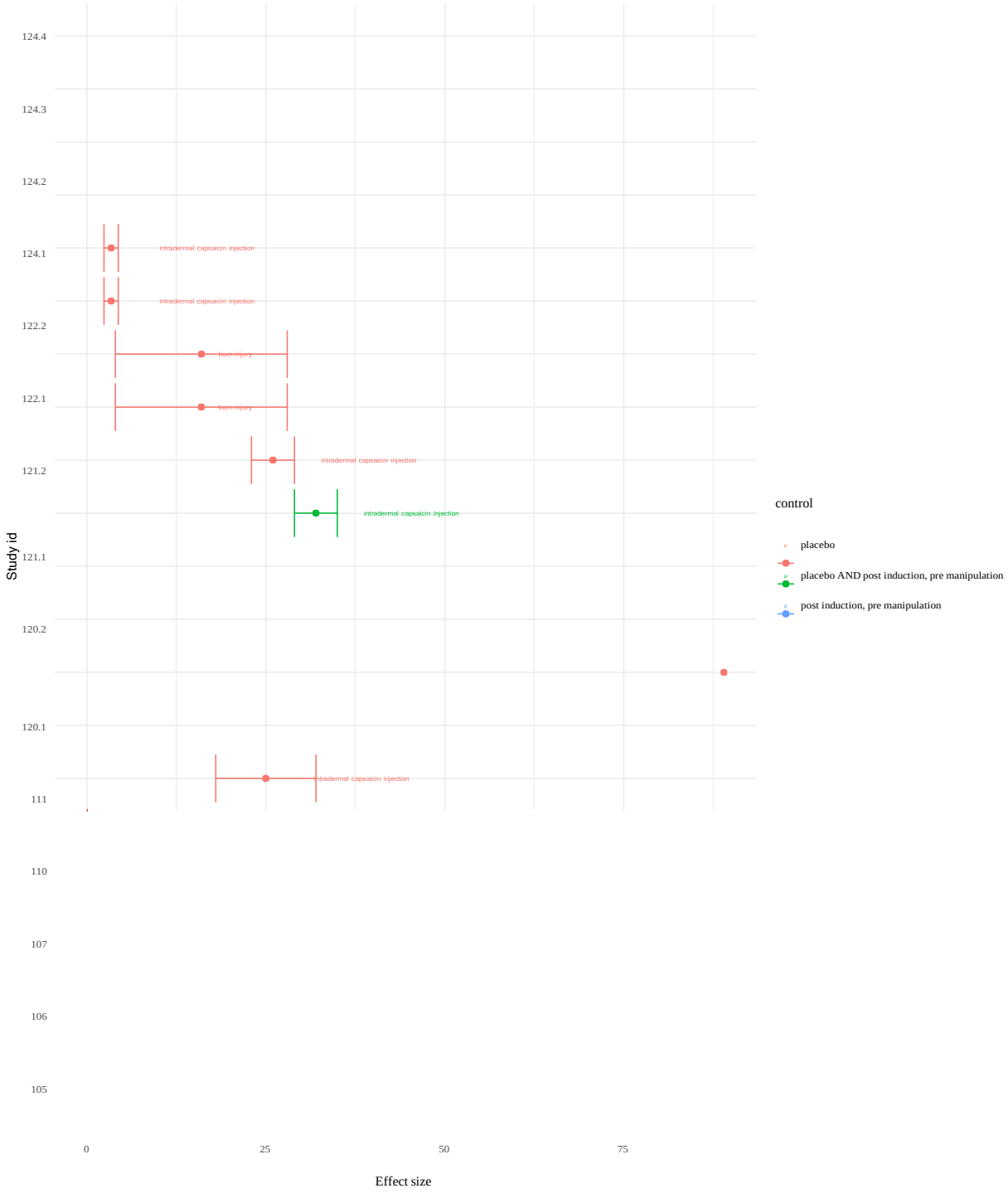

```
## create forest plot (control)

ggplot(data = mag,
  aes(y = study_id,
      x = c.CTM,
      group = control,
      fill = control,
      colour = control,


## create forest plot (control)
```

```
        scale_x_continuous(limits = c(-10, 80)))) +
  geom_point(size = 2, aes(group = control)) +
  geom_errorbar(aes(
    y = study_id,
    height = .1,
    xmax = c.upper.bound, xmin
    = c.lower.bound)) +
  geom_text(aes(
    label = induction,
    x = as.numeric(c.CTM)),
    hjust = -0.5,
    size = 2) +
  labs(title = 'Magnitude of SH of all studies that gave an induction and administered a control for th
    subtitle = 'Data in each study is expressed by different measures of central tendency.x = 'Effect size',
    y = 'Study id') +
  geom_vline(
    xintercept = 0, color =
    'red', linetype = 'dashed',
    alpha = .5,
    size = 1
  ) +
  theme_minimal()
```

The text inm

Magnitude of SH of all studies that gave an induction and administered a control for the manipulation

Data in each study is expressed by different measures of central tendency. The text inmdifferent colours shows the induction which is coloured coded by the type ofcontrol. Dotted vertical line is the point of no effect



Surface area of SH

```r
## import surface area data

area <- read_excel(path = here("sr data/mag_area_bl_fu.xlsx"),
  sheet = "area") %>%
  mutate(m.baseline.CTM = as.numeric(m.baseline.CTM)) %>%
  mutate(c.baseline.CTM = as.numeric(c.baseline.CTM)) %>%
  mutate(m.baseline.upper.bound = as.numeric(m.baseline.upper.bound)) %>%
  mutate(c.baseline.upper.bound = as.numeric(c.baseline.upper.bound)) %>%
  mutate(m.baseline.lower.bound = as.numeric(m.baseline.lower.bound)) %>%
  mutate(c.baseline.lower.bound = as.numeric(c.baseline.lower.bound)) %>%
  mutate(m.CTM = as.numeric(m.CTM)) %>%
  mutate(c.CTM = as.numeric(c.CTM)) %>%
  mutate(m.upper.bound = as.numeric(m.upper.bound)) %>%
  mutate(c.upper.bound = as.numeric(c.upper.bound)) %>%
  mutate(m.lower.bound = as.numeric(m.lower.bound)) %>%
```

```
## create plot (manipulation)
ggplot(data = area, aes(y = study_id, x = m.CTM, group = manipulation,
    fill = manipulation, colour = manipulation, scale_x_continuous(limits = c(-800,1900)))) +
       geom_point(size = 2, aes(group = manipulation)) +
    geom_errorbar(aes(y = study_id, height = 0.1, xmax = m.upper.bound,xmin =
       m.lower.bound)) + geom_text(aes(label = induction,
    x = as.numeric(m.CTM)), hjust = -1, size = 2) + labs(title = "Surface of SH of all studies that gavesubtitle = "Data in each
    study is expressed by different measures of central tendency. The text in dix = "Effect size", y = "Study id") +
    geom_vline(xintercept = 0,
    color = "red", linetype = "dashed", alpha = 0.5, size = 1) +
    theme_minimal()
```

```
## Warning: Ignoring unknown aesthetics: height
```

```
## Warning: Removed 1 rows containing missing values (geom_point).##
```

```
Warning: Removed 1 rows containing missing values (geom_text).
```

Surface of SH of all studies that gave an induction and administered a pharmacological manipulation to manipulate SH

Data in each study is expressed by different measures of central tendency. The text in different colours shows the induction which is coloured coded by the type of manipuation received. Dotted vertical line is the point of no effect


```
## create forest plot (control)
ggplot(data = area, aes(y = study_id, x = c.CTM, group = control,
    fill = control, colour = control, scale_x_continuous(limits = c(-800,1900)))) +
       geom_point(size = 2, aes(group = control)) +
    geom_errorbar(aes(y = study_id, height = 0.1, xmax = c.upper.bound,xmin =
       c.lower.bound)) + geom_text(aes(label = induction,
```

```
  x = as.numeric(c.CTM)), hjust = -1, size = 2) + labs(title = "Surface of SH of all studies that gavesubtitle = "Data in each
  study is expressed by different measures of central tendency. The text in\ndx = "Effect size", y = "Study id") +
  geom_vline(xintercept = 0,
  color = "red", linetype = "dashed", alpha = 0.5, size = 1) +
  theme_minimal()
```

```
## Warning: Ignoring unknown aesthetics: height
```

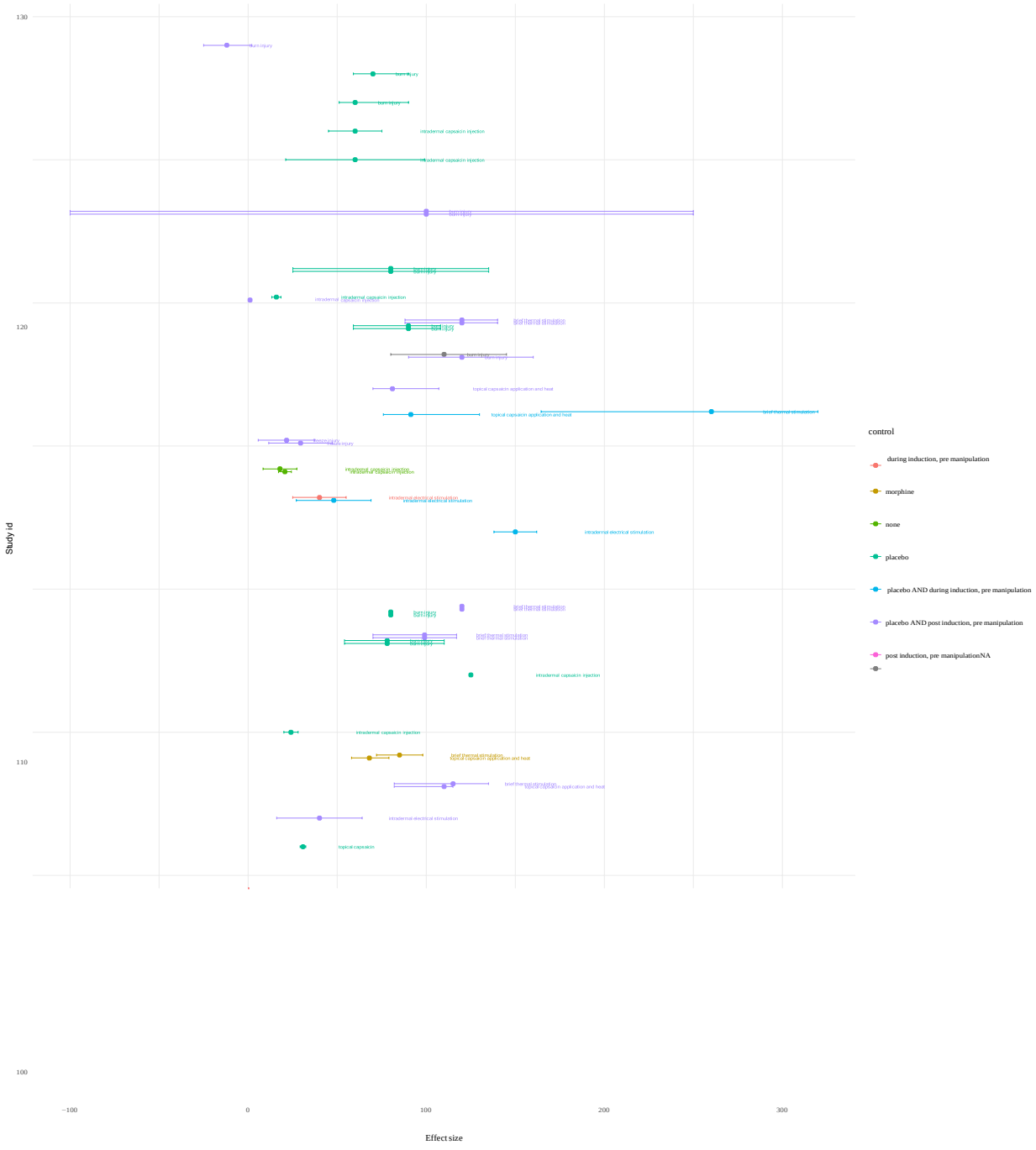
```
## Warning: Removed 5 rows containing missing values (geom_point).##
```

```
Warning: Removed 5 rows containing missing values (geom_text).
```

Surface of SH of all studies that gave an induction and administered a control for the manipulation

Data in each study is expressed by different measures of central tendency. The text in

different colours shows the induction which is coloured coded by the type of control. Dotted vertical line is thepoint of no effect



Calculations and conversions

Magnitude of SH

```
# studies excluded from meta-analysis

mag %<>% mutate(
    `reason for exclusion from meta-analysis` =
      case_when(
        study_id == "106" ~ "missing data for meta-analysis",
        study_id == "107" ~ "missing data for meta-analysis",
        study_id == "124.1" ~ "missing data for meta-analysis",
        study_id == "124.2" ~ "missing data for meta-analysis",
        study_id == "124.3" ~ "missing data for meta-analysis",
        study_id == "124.4" ~ "missing data for meta-analysis",


# studies excluded from meta-analysis

mag %<>% mutate(
    `reason for exclusion from meta-analysis` =
      case_when(
        study_id == "106" ~ "missing data for meta-analysis",
        study_id == "107" ~ "missing data for meta-analysis",
        study_id == "124.1" ~ "missing data for meta-analysis",
        study_id == "124.2" ~ "missing data for meta-analysis",
        study_id == "124.3" ~ "missing data for meta-analysis",
        study_id == "124.4" ~ "missing data for meta-analysis",
```

Table 1: Studies assessing the magnitude of SH but excluded from the meta-analysis

study_id	first_author	year	induction	manipulation	reason for exclusion from meta-analysis
106	106-Gottrup	2004	intradermal capsaicin injection	ketamine (single)	missing data for meta-analysis
107	107-Gottrup	2000	intradermal capsaicin injection	ketamine (multiple)	missing data for meta-analysis
110	110-Klein	2008	intradermal capsaicin injection	ketamine (single)	only one study is available with the following manipulation
111	111-Klein	2007	HFS	ketamine (single)	missing data for meta-analysis
124.1	124.1-Sethna	1998	intradermal capsaicin injection	ketamine (single)	missing data for meta-analysis
124.2	124.2-Sethna	1998	intradermal capsaicin injection	ketamine (single)	missing data for meta-analysis
124.3	124.3-Sethna	1998	intradermal capsaicin injection	ketamine (single)	missing data for meta-analysis
124.4	124.4-Sethna	1998	intradermal capsaicin injection	ketamine (single)	missing data for meta-analysis

```
            study_id == "111" ~ "missing data for meta-analysis",
            study_id == "110" ~ "only one study is available with the following manipulation"))

mag %>%
    filter(
            study_id == "106"|
            study_id == "107" |
            study_id == "124.1" |
            study_id == "124.2" |
            study_id == "124.3" |
            study_id == "124.4"|
            study_id == "110"|
            study_id == "111") %>%
  select("study_id", "first_author", "year", "induction", "manipulation", "reason for exclusion from mekbl(caption = "Studies
  assessing the magnitude of SH but excluded from the meta-analysis", booktabskable_styling(bootstrap_options = "striped",
  "scale_down", full_width = F) %>%
    kable_material(lightable_options = "hover")# %>%
```

```
  #save_kable('Tables/Studies excluded from meta-analysis_mag.png')


mag %<>%filter(
    !study_id == "106",!study_id == "107" ,!study_id == "124.1" ,!study_id == "124.2" ,!study_id == "12
  )

## Now we need to convert different central measures of tendency and spread
## indicators to the same central measure and spread indicator (mean (SD)) prior## to conducting the meta-
analysis.

# extract all data expressed as mean (sd)

mag_raw_ratings_mean_sd <- mag %>%
  filter(`central_tendency_measure` == "mean",
           spread_indicator == "SD") %>%
  rename(m.mean = m.CTM) %>%
  rename(c.mean = c.CTM) %>%
  mutate(m.sd = (m.mean - m.lower.bound)) %>%
  mutate(c.sd = (c.mean - c.lower.bound)) %>% # calculate sd

  select( study_id,
    first_author,
    induction,
```

```
      manipulation,control,
      rob,
      n.m,
      n.c, c.mean,
      c.sd,
      m.mean,
      m.sd,
      scale_anchors
   )

# convert all data expressed as mean (sem) to mean (sd)

mag_raw_ratings_mean_sem <- mag %>%
   filter(`central_tendency_measure` == "mean",
             spread_indicator == "SEM") %>%
   rename(m.mean = m.CTM) %>%
   rename(c.mean = c.CTM) %>%
   mutate(m.sem = (m.mean - m.lower.bound)) %>%
   mutate(c.sem = (c.mean - c.lower.bound))

# manipulation: estimate mean (sd) from mean (sem)

m_mag_sd <-
   function(m.sem,

                n.m) {
      m_mag_sd = round((m.sem * sqrt(n.m)), 1)
      list("m.sd" = m_mag_sd)
   }

mag_raw_ratings_mean_sem
   %<>%rowwise() %>%
   mutate(m = list(m_mag_sd(m.sem,
                                                n.m))) %>%
   unnest_wider(m)


# control: estimate mean (sd) from mean (sem)

c_mag_sd <-
   function(c.sem,

                n.c) {
      c_mag_sd = round((c.sem * sqrt(n.c)), 1)
      list("c.sd" = c_mag_sd)
   }

mag_raw_ratings_mean_sem
   %<>%rowwise() %>%
   mutate(c = list(c_mag_sd(c.sem,
                                                n.c))) %>%
   unnest_wider(c)

mag_raw_ratings_mean_sem

%<>%
   select( study_id,
      first_author,
```

```
    induction,
    manipulation,
    control,
    rob,
    n.m,
    n.c,
    c.mean,
    c.sd,
    m.mean,
    m.sd,
    scale_anchors
  )

# bind all data frames

mag_raw_ratings_conditions <-
  rbind(mag_raw_ratings_mean_sd, mag_raw_ratings_mean_sem)

rm(mag_raw_ratings_mean_sd, mag_raw_ratings_mean_sem)
```

Surface area of SH

```
area
%<>%
mutate(
      `reason for exclusion from meta-analysis` =
        case_when(
          study_id == "106" ~ "missing data for meta-analysis", study_id ==
          "107" ~ "missing data for meta-analysis", study_id == "124.1" ~
          "missing data for meta-analysis",study_id == "124.2" ~ "missing
          data for meta-analysis",study_id == "124.3" ~ "missing data for
          meta-analysis",study_id == "124.4" ~ "missing data for meta-
          analysis",study_id == "109.1" ~ "missing data for meta-analysis",
          study_id == "109.2" ~ "missing data for meta-analysis",study_id ==
          "109.3" ~ "missing data for meta-analysis",study_id == "109.4" ~
          "missing data for meta-analysis",
          study_id == "117" ~ "only one study is available with the following manipulation",
          study_id == "123.1" ~ "failed to estimate the mean and standard deviation from bc.mean.sd formustudy_id == "123.2"
          ~ "failed to estimate the mean and standard deviation from bc.mean.sd formustudy_id == "128" ~ "failed to estimate the
          mean and standard deviation from bc.mean.sd formulastudy_id == "129" ~ "failed to estimate the mean and standard
          deviation from bc.mean.sd formula
        ))

area %>%
      filter(study_id == "106"|
            study_id == "107" |
            study_id == "124.1" |
            study_id == "124.2" |
            study_id == "124.3" |
            study_id == "124.4" |
            study_id == "109.1" |
            study_id == "109.2" |
```

Table 2: Studies assessing the surface area of SH but excluded from the meta-analysis

study_id	first_author	year	induction	manipulation	reason for exclusion from meta-analysis
106.0	106-Gottrup	2004	intradermal capsaicin injection	ketamine (single)	missing data for meta-analysis
107.0	107-Gottrup	2000	intradermal capsaicin injection	ketamine (multiple)	missing data for meta-analysis
109.1	109,1-Ilkjaer	1996	burn injury	ketamine (multiple)	missing data for meta-analysis
109.2	109,2-Ilkjaer	1996	burn injury	ketamine (multiple)	missing data for meta-analysis
109.3	109,3-Ilkjaer	1996	brief thermal stimulation	ketamine (multiple)	missing data for meta-analysis
109.4	109,4-Ilkjaer	1996	brief thermal stimulation	ketamine (multiple)	missing data for meta-analysis
117.0	117-Mikkelsen	2001	topical capsaicin application and heat	magnesium sulphate (multiple)	only one study is available with the following manipulation
123.1	123,1-Schulte	2004	burn injury	ketamine (single)	failed to estimate the mean and standard deviation from bc.mean.sd formula due to the nature of the data.
123.2	123,2-Schulte	2004	burn injury	ketamine and morphine (single)	failed to estimate the mean and standard deviation from bc.mean.sd formula due to the nature of the data.
124.1	124,1-Sethna	1998	intradermal capsaicin injection	ketamine (single)	missing data for meta-analysis
124.2	124,2-Sethna	1998	intradermal capsaicin injection	ketamine (single)	missing data for meta-analysis
124.3	124,3-Sethna	1998	intradermal capsaicin injection	ketamine and alfentanil (single)	missing data for meta-analysis
124.4	124,4-Sethna	1998	intradermal capsaicin injection	ketamine and alfentanil (single)	missing data for meta-analysis
128.0	128-Warncke	2000	burn injury	ketamine (multiple)	failed to estimate the mean and standard deviation from bc.mean.sd formula due to the nature of the data.
129.0	129-Warncke	1997	burn injury	ketamine (single)	failed to estimate the mean and standard deviation from bc.mean.sd formula due to the nature of the data.

```
            study_id == "109.3" |
            study_id == "109.4" |
            study_id == "117" |
            study_id == "123.1" |
            study_id == "123.2" |
            study_id == "128" |
            study_id == "129"

)
%>%

  select("study_id", "first_author", "year", "induction", "manipulation", "reason for exclusion from mekbl(caption = "Studies
  assessing the surface area of SH but excluded from the meta-analysis",   booktakable_styling(bootstrap_options = "striped",
  "scale_down", full_width = F) %>%
    kable_material(lightable_options = "hover") #%>%
```

```
  #save_kable('Tables/Studies excluded from meta-analysis_area.png')

area %<>%
filter(
    !study_id == "106",
    !study_id == "107" ,
    !study_id == "124.1" ,
    !study_id == "124.2",
    !study_id == "124.3" ,
    !study_id == "124.4",
    !study_id == "109.1" ,
    !study_id == "109.2",
    !study_id == "109.3",
    !study_id == "109.4",
    !study_id == "117",
    !study_id == "123.1",
    !study_id == "123.2",
    !study_id == "128",
    !study_id == "129"
  )

# all surface area expressed as raw ratings

sa_raw_ratings <- area %>%
  filter(`outcome data expressed as` == "raw ratings")

## convert all data to mean (sd)

# extract all data expressed as mean (sd)#save_kable('Tables/Studies excluded from meta-
analysis_area.png')


area %<>%
  filter(
    !study_id == "106",
    !study_id == "107" ,
    !study_id == "124.1" ,
    !study_id == "124.2",
    !study_id == "124.3" ,
    !study_id == "124.4",
    !study_id == "109.1" ,
    !study_id == "109.2",
    !study_id == "109.3",
    !study_id == "109.4",
    !study_id == "117",
    !study_id == "123.1",
    !study_id == "123.2",
    !study_id == "128",
    !study_id == "129"
  )

# all surface area expressed as raw ratings

sa_raw_ratings <- area %>%
  filter(`outcome data expressed as` == "raw ratings")
```

```
sa_raw_ratings_mean_sd <- sa_raw_ratings %>%
  filter(`central_tendency_measure (ctm)` == "mean",
           spread_indicator == "SD") %>%
  rename(m.mean = m.CTM) %>%
  rename(c.mean = c.CTM) %>%
  mutate(m.sd = (m.mean - m.lower.bound)) %>%
  mutate(c.sd = (c.mean - c.lower.bound)) %>% # calculate  sd

  select( study_id,
    first_author,
    induction,
    manipulation,
    control,
    rob,
    n.m,
    n.c, c.mean,
    c.sd,
    m.mean,
    m.sd,
    area_unit
  )

# convert all data from median (iqr) to mean (sd)

sa_raw_ratings_median_iqr <- sa_raw_ratings %>%
  filter(`central_tendency_measure (ctm)` == "median"|
           spread_indicator == "IQR")

# manipulation: estimate mean SD from median iqr #
step 1: convert median (iqr) to mean (sd), then
m_area_sd <-

  function(m.lower.bound,
             m.CTM,

             m.upper.bound,
             n.m) {
    m_area_sd = bc.mean.sd(
      q1.val = m.lower.bound,
      med.val = m.CTM,
      q3.val   = m.upper.bound,n
      = n.m
    )
    list("m.mean" = m_area_sd$est.mean,
           "m.sd" = m_area_sd$est.sd)
  }

sa_raw_ratings_median_iqr %<>%
  rowwise() %>%
  mutate(m = list(m_area_sd(m.lower.bound,
                            m.CTM,
                            m.upper.bound,n.m)))
                            %>%
  unnest_wider(m)
```

Warning in stats::optimize(lambda.to.err, c(-2, 3)): NA/Inf replaced by maximum## positive value

```
# control: estimate mean sd from median iqr

c_area_sd <-
  function(c.lower.bound,
                c.CTM,

                c.upper.bound,
                n.c) {
    c_area_sd = bc.mean.sd(
      q1.val = c.lower.bound,
      med.val = c.CTM,
      q3.val  = c.upper.bound,n  =
      n.c
    )
    list("c.mean" = c_area_sd$est.mean,"c.sd"
         = c_area_sd$est.sd)
  }

sa_raw_ratings_median_iqr %<>%
  rowwise() %>%
  mutate(c = list(c_area_sd(c.lower.bound,
                                      c.CTM,
                                      c.upper.bound,n.c)))
                                      %>%
  unnest_wider(c)
sa_raw_ratings_median_iqr %<>%
  select( study_id,
    first_author,
    induction,
    manipulation,
    control,
    rob,
    n.m,
    n.c, c.mean,
    c.sd,
    m.mean,
    m.sd,
    area_unit
  )

## convert all data expressed as mean (sem) to mean (sd)
sa_raw_ratings_mean_sem <- sa_raw_ratings %>%
  filter(`central_tendency_measure (ctm)` == "mean",
           spread_indicator == "SEM") %>%
  rename(m.mean = m.CTM) %>%
  rename(c.mean = c.CTM) %>%
  mutate(m.sem = (m.mean - m.lower.bound)) %>% mutate(c.sem =
  (c.mean - c.lower.bound)) # calculate sem

# manipulation: estimate mean (sd) from mean (sem)

m_area_sd_1 <-
```

```
  function(m.sem,

                  n.m) {
    m_area_sd_1 = round((m.sem * sqrt(n.m)), 1)
    list("m.sd" = m_area_sd_1)
  }

sa_raw_ratings_mean_sem %<>%
  rowwise() %>%
  mutate(m = list(m_area_sd_1(m.sem,
                                        n.m))) %>%
  unnest_wider(m)

# control: estimate mean (sd) from mean (sem)

c_area_sd_1 <-
  function(c.sem,

                  n.c) {
    c_area_sd_1 = round((c.sem * sqrt(n.c)), 1)list("c.sd"
    = c_area_sd_1)
  }

sa_raw_ratings_mean_sem %<>%
  rowwise() %>%
  mutate(c = list(c_area_sd_1(c.sem,
                                        n.c))) %>%
  unnest_wider(c)

sa_raw_ratings_mean_sem %<>%
  select(
    study_id, first_author,
    induction,
    manipulation,control,
    rob,
    n.m,
    n.c, c.mean,
    c.sd,
    m.mean,
    m.sd,
    area_unit
  )

sa_raw_data <-
  rbind(sa_raw_ratings_mean_sd,
        sa_raw_ratings_mean_sem, sa_raw_ratings_median_iqr) #
        bind all data frames


#####

# all surface area expressed as change from baseline

sa_change_from_bl <- area %>%
  filter(`outcome data expressed as` == "% baseline"|
```

```
            `outcome data expressed as` == "difference from baseline")

# convert all data from median (iqr) to mean (sd)

sa_change_from_bl_median_iqr <- sa_change_from_bl %>%
  filter(`central_tendency_measure (ctm)` == "median"|
           spread_indicator == "IQR")

# manipulation: estimate mean SD from median iqr#
step 1: convert median (iqr) to mean (sd) m_area_sd_2
<-

  function(m.lower.bound,
           m.CTM,

           m.upper.bound,
           n.m) {
    m_area_sd_2  =  bc.mean.sd(
      q1.val   =   m.lower.bound,
      med.val = m.CTM,
      q3.val = m.upper.bound,n   =
      n.m
    )
    list("m.mean" = m_area_sd_2$est.mean,
         "m.sd" = m_area_sd_2$est.sd)
  }

sa_change_from_bl_median_iqr %<>%
  rowwise() %>%
  mutate(m = list(m_area_sd_2(m.lower.bound,
                              m.CTM,
                              m.upper.bound,n.m)))
                              %>%
  unnest_wider(m)

# control: estimate mean sd from median iqr

c_area_sd_2 <-
  function(c.lower.bound,
           c.CTM,

           c.upper.bound,
           n.c) {
    c_area_sd_2   =   bc.mean.sd(
      q1.val   =   c.lower.bound,
      med.val = c.CTM,
      q3.val = c.upper.bound,n   =
      n.c
    )
    list("c.mean" = c_area_sd_2$est.mean,"c.sd"
         = c_area_sd_2$est.sd)
  }

sa_change_from_bl_median_iqr %<>%
  rowwise() %>%
  mutate(c = list(c_area_sd_2(c.lower.bound,
                              c.CTM,
                              c.upper.bound,
```

```
                                        n.c))) %>%
  unnest_wider(c)

sa_change_from_bl_median_iqr %<>%
  select(
    study_id, first_author,
    induction,
    manipulation,control,
    rob,
    n.m,
    n.c, c.mean,
    c.sd,
    m.mean,
    m.sd,
    area_unit
  )

## convert all data expressed as mean (sem) to mean (sd)
sa_change_from_bl_mean_sem <- sa_change_from_bl %>%
  filter(`central_tendency_measure (ctm)` == "mean",
           spread_indicator == "SEM") %>%
  rename(m.mean = m.CTM) %>%
  rename(c.mean = c.CTM) %>%
  mutate(m.sem = (m.mean - m.lower.bound)) %>% mutate(c.sem =
  (c.mean - c.lower.bound)) # calculate sem

# manipulation: estimate mean (sd) from mean (sem)

m_area_sd_3 <-
  function(m.sem,

           n.m) {
    m_area_sd_3 = round((m.sem * sqrt(n.m)), 1)
    list("m.sd" = m_area_sd_3)
  }

sa_change_from_bl_mean_sem
  %<>%rowwise() %>%
  mutate(m = list(m_area_sd_3(m.sem,
                              n.m))) %>%
  unnest_wider(m)

# control: estimate mean (sd) from mean (sem)

c_area_sd_3 <-
  function(c.sem,

           n.c) {
    c_area_sd_3 = round((c.sem * sqrt(n.c)), 1)list("c.sd"
    = c_area_sd_3)
  }

sa_change_from_bl_mean_sem
  %<>%rowwise() %>%
  mutate(c = list(c_area_sd_3(c.sem,
```

```
                                        n.c))) %>%
  unnest_wider(c)

sa_change_from_bl_mean_sem
  %<>%select(
    study_id, first_author,
    induction,
    manipulation,control,
    rob,
    n.m,
    n.c, c.mean,
    c.sd,
    m.mean,
    m.sd,
    area_unit
  )

## convert all data expressed as mean (95% CI) to mean (sd)
sa_change_from_bl_mean_ci <-
  sa_change_from_bl %>%
  filter(`central_tendency_measure (ctm)` == "mean") %>%
  filter(spread_indicator == "95% confidence intervals")

## To convert the spread indicator from 95% ci to sd we will use 2 steps.
## Step 1: convert to SEM using the source here (https://handbook-5-1.cochrane.org/chapter_7/7_7_7_2_ob## Step 2, convert the SEM to SD.

## step 1 formula; SE = (upper limit – lower limit) / 3.92.
sa_change_from_bl_mean_ci  %<>%
  rename(c.mean = c.CTM) %>%
  rename(m.mean = m.CTM ) %>%
  mutate(c.sem = (c.upper.bound - c.lower.bound) / 3.92) %>%mutate(m.sem =
  (m.upper.bound - m.lower.bound) / 3.92)

# manipulation:  estimate  sd  from  sem

m_area_sd_4 <-
  function(m.sem,

           n.m) {
    m_area_sd_4 = round((m.sem * sqrt(n.m)), 1)
    list("m.sd" = m_area_sd_4)
  }

sa_change_from_bl_mean_ci %<>%
  rowwise() %>%
  mutate(m = list(m_area_sd_4(m.sem,
                                        n.m))) %>%
  unnest_wider(m)

# control:  estimate  sd  from  sem

c_area_sd_4 <-
  function(c.sem,
```

```
                n.c) {
    c_area_sd_4 = round((c.sem * sqrt(n.c)), 1)list("c.sd"
    = c_area_sd_4)
  }

sa_change_from_bl_mean_ci %<>%
  rowwise() %>%
  mutate(c = list(c_area_sd_4(c.sem,
                                        n.c))) %>%
  unnest_wider(c)

sa_change_from_bl_mean_ci %<>%
  select(
    study_id, first_author,
    induction,
    manipulation,control,
    rob,
    n.m,
    n.c, c.mean,
    c.sd,
    m.mean,
    m.sd,
    area_unit
  )

# bind all data frames

sa_change_from_bl_data <-
  rbind(
    sa_change_from_bl_mean_ci,
    sa_change_from_bl_mean_sem,
    sa_change_from_bl_median_iqr
  )

rm(
  sa_change_from_bl_mean_ci,
  sa_change_from_bl_mean_sem,
  sa_change_from_bl_median_iqr,
  sa_change_from_bl, sa_raw_ratings,
  sa_raw_ratings_mean_sd,
  sa_raw_ratings_mean_sem,
  sa_raw_ratings_median_iqr
)

rm(c_area_sd, c_area_sd_1,
   c_area_sd_2,
   c_area_sd_3,
   c_area_sd_4,c_mag_sd,
```

```
    m_area_sd,
    m_area_sd_1,
    m_area_sd_2,
    m_area_sd_3,
    m_area_sd_4,
    m_mag_sd)

all_sa_data_converted <- rbind(sa_raw_data ,sa_change_from_bl_data) # use later for publication bias
```

Meta-analysis

Magnitude

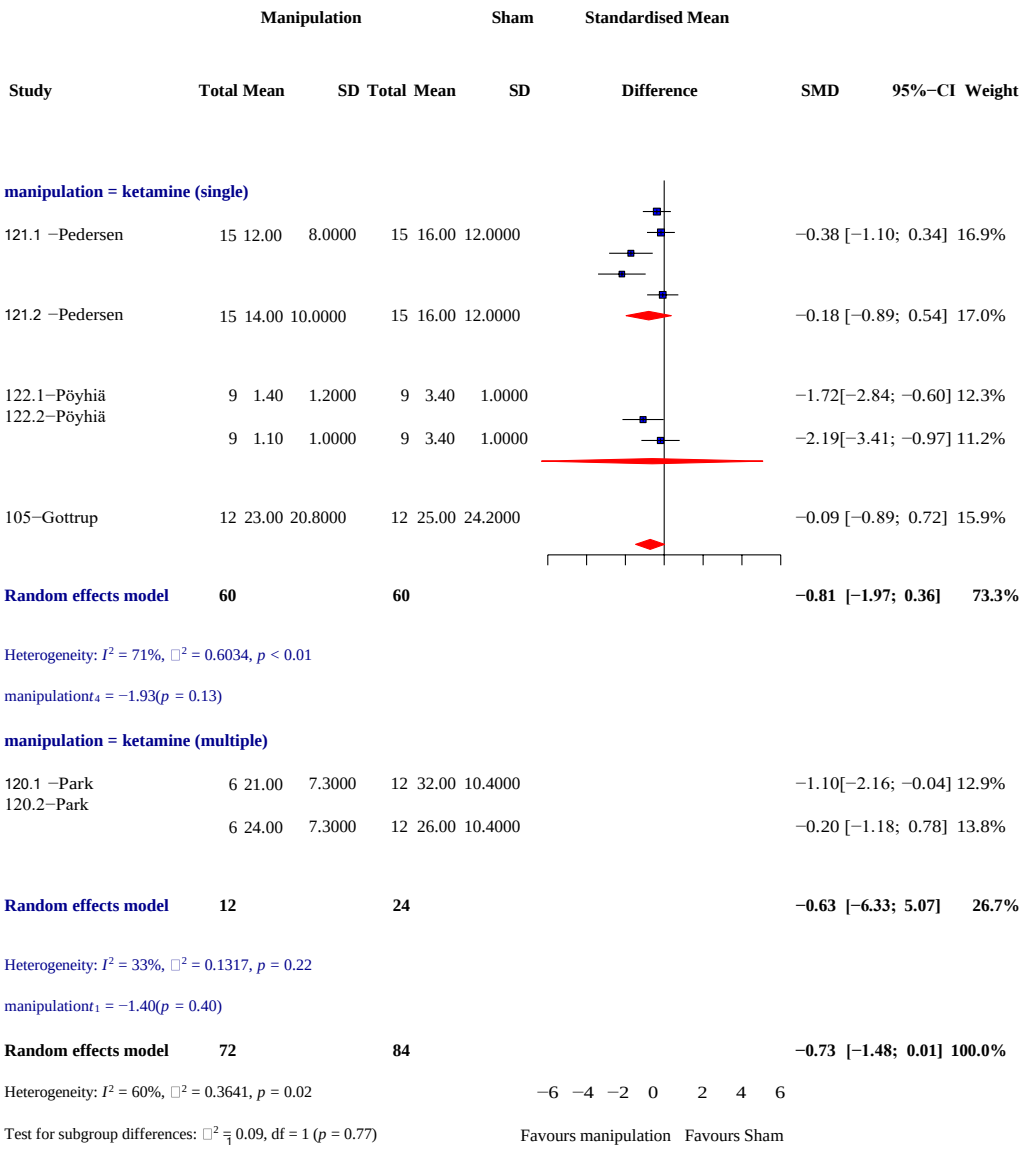
```
## create forest plot for all studies that manipulated the
## magnitude of SH.Subgroup plot by manipulation

meta_mag <- meta::metacont(
  n.e = n.m,
  mean.e = m.mean,
  sd.e = m.sd,
  n.c = n.c,
  mean.c = c.mean,
  sd.c = c.sd,
  studlab = first_author,
  data = mag_raw_ratings_conditions,
  sm = "SMD",
  # standardised mean difference

  method.smd = "Hedges",
  # most preferred because it can better handle small sample sizes

  fixed = FALSE,
  random = TRUE,
  method.tau = "REML",
  method.random.ci = "HK",
  subgroup = manipulation,
  label.e = "Manipulation",
  label.c = "Sham",
  layout = "meta")
```

```
##                      SMD             95%-CI %W(random)           manipulation
##  121.1-Pedersen   -0.3816 [-1.1047;  0.3415]       16.9     ketamine (single)
##  121.2-Pedersen   -0.1762 [-0.8934;  0.5411]       17.0     ketamine (single)
##  122.1-Pöyhiä     -1.7243 [-2.8441; -0.6044]       12.3     ketamine (single)
##  122.2-Pöyhiä     -2.1902 [-3.4148; -0.9656]       11.2     ketamine (single)
##  105-Gottrup      -0.0856 [-0.8862;  0.7150]       15.9     ketamine (single)
##  120.1-Park       -1.0980 [-2.1576; -0.0384]       12.9   ketamine (multiple)
##  120.2-Park       -0.1996 [-1.1823;  0.7831]       13.8   ketamine (multiple)
##
## Number of studies combined: k = 7##
Number of observations: o = 156
```

```
##
##                                      SMD                  95%-CI         t p-value
##  Random effects model -0.7344 [-1.4789; 0.0101] -2.41   0.0523
##
##  Quantifying heterogeneity:
##   tau^2 = 0.3641 [0.0125; 3.1808]; tau = 0.6034 [0.1118; 1.7835]
##   I^2 = 60.3% [8.9%; 82.7%]; H = 1.59 [1.05; 2.40]
##
##  Test of heterogeneity:
##         Q d.f. p-value
##   15.11        6   0.0194
##
##  Results for subgroups (random effects model):
##                                                k         SMD                  95%-CI   tau^2         tau
##  manipulation = ketamine (single)              5 -0.8088 [-1.9741; 0.3565] 0.6034 0.7768
##  manipulation = ketamine (multiple)            2 -0.6261 [-6.3260; 5.0739] 0.1317 0.3629
##                                                    Q     I^2
##  manipulation = ketamine (single)              13.62 70.6%
##  manipulation = ketamine (multiple)   1.48 32.6%
##
##  Test for subgroup differences (random effects model):
##                                Q d.f. p-value
##  Between groups        0.09        1   0.7661
##
##  Details on meta-analytical method:
##  - Inverse variance method
##  - Restricted maximum-likelihood estimator for tau^2
##  - Q-Profile method for confidence interval of tau^2 and tau
##  - Hartung-Knapp (HK) adjustment for random effects model (df = 6)
##  - Hedges' g (bias corrected standardised mean difference; using exact formulae)
```

```
forest(
   meta_mag,
   col.square = "blue", col.square.lines =
   "black", col.diamond.random = "red",
   col.diamond.lines.random = "red",
   allstudies = TRUE,
   comb.fixed = FALSE,
   comb.random = TRUE,
   label.left = "Favours manipulation",
   col.label.left = "black", label.right =
   "Favours Sham", col.label.right = "black",
   test.effect.subgroup.random = TRUE,overall
   = TRUE,
   print.byvar = FALSE, col.by
   = "dark blue",
   addrow.subgroups = TRUE,
   addrow.overall = TRUE,
   addrow = TRUE,
   label.test.effect.subgroup.random = "manipulation",
   test.subgroup.random = TRUE)
grid.text("Effect of the manipulation and Sham on magnitude of SH, when\n sub-grouped by manipulation (
```

Effect of the manipulation and Sham on magnitude of SH, when sub−grouped by manipulation (N = 7)



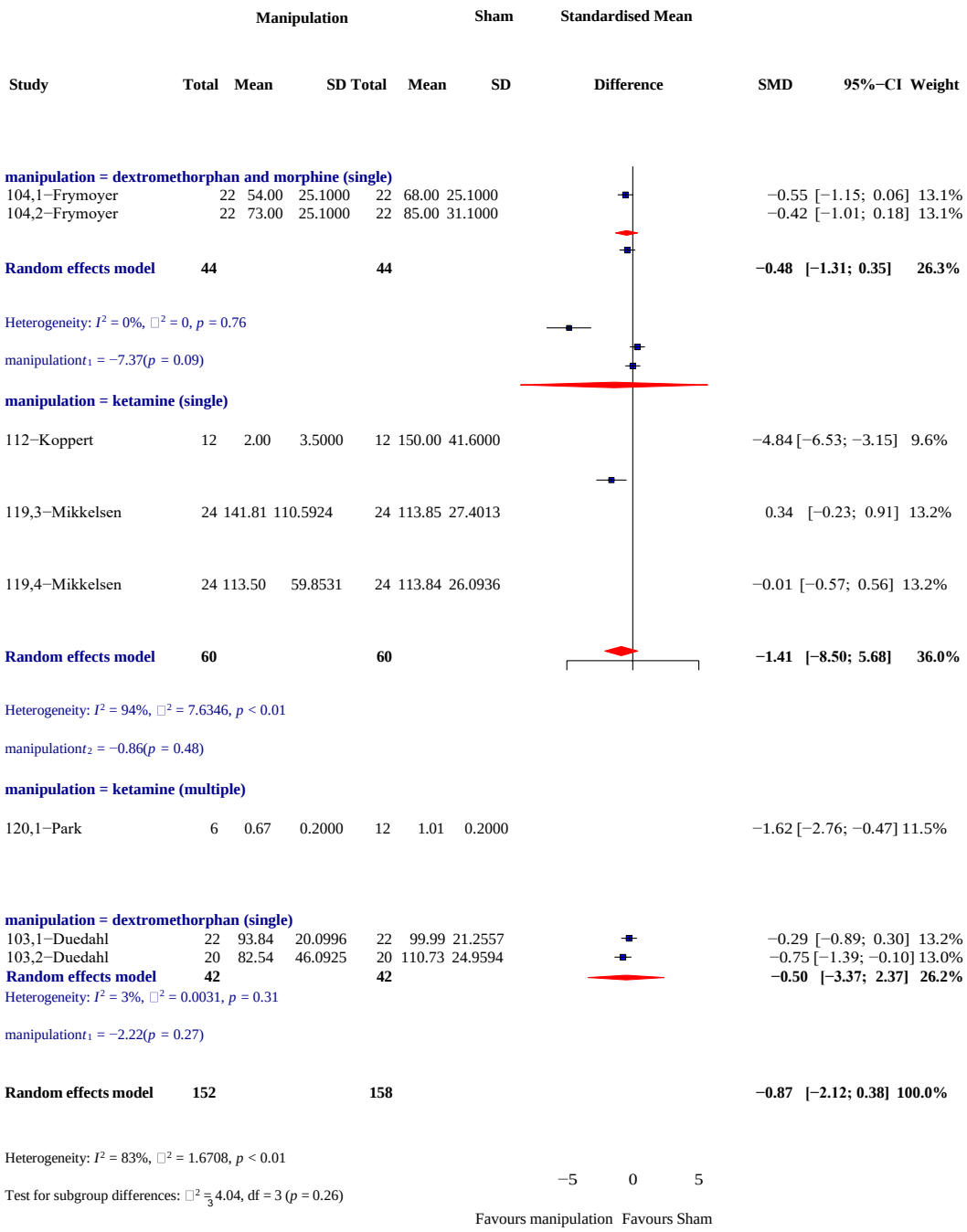
Surface area

```
## create forest plot for all studies that manipulated the

## surface area of SH and expressed their follow-up ratings as the

# % change or difference from baseline ratings.
# Subgroup analysis by manipulation
meta_sa_change_from_bl<- meta::metacont(

  n.e = n.m,
  mean.e = m.mean,
  sd.e = m.sd,
  n.c = n.c,
  mean.c = c.mean,
  sd.c = c.sd,
  studlab = first_author,
  data = sa_change_from_bl_data,
  sm = "SMD",
```

```
  # most preferred because it can better handle small sample sizes
  fixed = FALSE, random
  = TRUE, method.tau =
  "REML",
  method.random.ci = "HK",
  subgroup = manipulation,
  label.e = "Manipulation",
  label.c = "Sham",
  layout = "meta")


forest( meta_sa_change_from_bl,
  col.square = "blue",
  col.square.lines = "black",
  col.diamond.random = "red",
  col.diamond.lines.random = "red",
  allstudies = TRUE,
  comb.fixed = FALSE,
  comb.random = TRUE,
  label.left = "Favours manipulation",
  col.label.left = "black", label.right =
  "Favours Sham", col.label.right = "black",
  test.effect.subgroup.random = TRUE,overall
  = TRUE,
  print.byvar = FALSE, col.by
  = "dark blue",
  addrow.subgroups = TRUE,
  addrow.overall = TRUE,
  addrow = TRUE,
  label.test.effect.subgroup.random = "manipulation",
  test.subgroup.random = TRUE
)
grid.text(
  "Effect of the manipulation and sham on surface area of SH, when\n sub-grouped by manipulation in cro
  .5,
  .9,
  gp = gpar(cex = 2)
)
```

Effect of the manipulation and sham on surface area of SH, when sub−grouped by manipulation in crossover and within−subject study designs

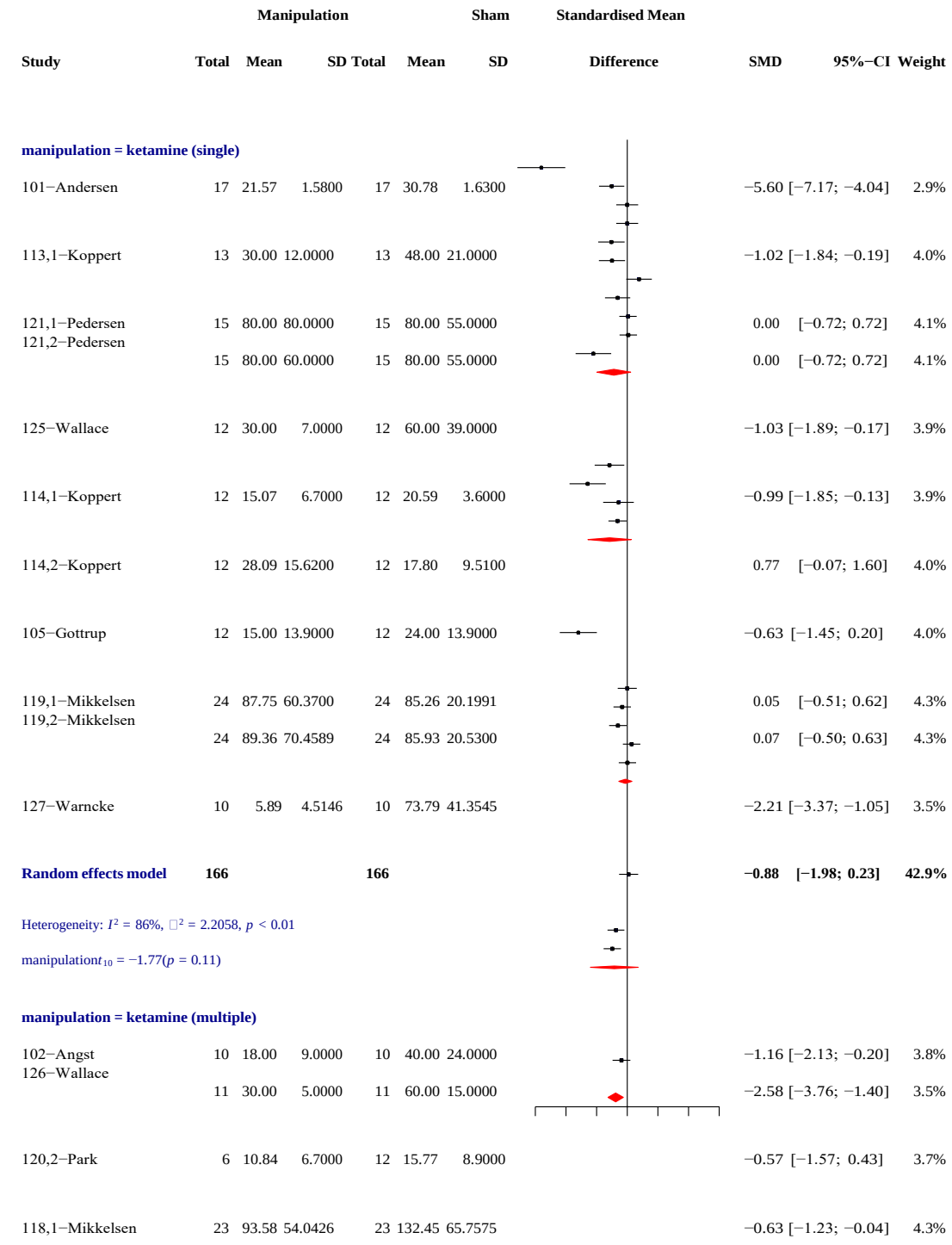
(change from baseline; N = 8)



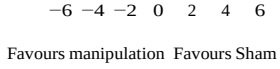
```
## create forest plot for all studies that manipulated the
## surface area of SH and reported raw ratings.

# Subgroup analysis by manipulation

meta_sa_raw_data <- meta::metacont(
  n.e = n.m,
  mean.e = m.mean,
  sd.e = m.sd,
  n.c = n.c,
  mean.c = c.mean,
  sd.c = c.sd,
  studlab = first_author,
  data = sa_raw_data,
  sm = "SMD",
  # standardised mean difference

  method.smd = "Hedges",
  # most preferred because it can better handle small sample sizes

  fixed = FALSE,
  random = TRUE,
  method.tau = "REML",
  method.random.ci = "HK",


## create forest plot for all studies that manipulated the
## surface area of SH and reported raw ratings.

# Subgroup analysis by manipulation

meta_sa_raw_data <- meta::metacont(
  n.e = n.m,
  mean.e = m.mean,
  sd.e = m.sd,
  n.c = n.c,
  mean.c = c.mean,
  sd.c = c.sd,
  studlab = first_author,
  data = sa_raw_data,
  sm = "SMD",
  # standardised mean difference

  method.smd = "Hedges",
  # most preferred because it can better handle small sample sizes

  fixed = FALSE,
  random = TRUE,
  method.tau = "REML",
  method.random.ci = "HK",
```

```
    subgroup = manipulation,
    label.e = "Manipulation",
    label.c = "Sham",
    layout = "meta")

forest( meta_sa_raw_data,
    col.square = "blue",
    col.square.lines = "black",
    col.diamond.random = "red",
    col.diamond.lines.random = "red",
    allstudies = TRUE,
    comb.fixed = FALSE,
    comb.random = TRUE,
    label.left = "Favours manipulation",
    col.label.left = "black", label.right =
    "Favours Sham", col.label.right = "black",
    test.effect.subgroup.random = TRUE,overall
    = TRUE,
    print.byvar = FALSE, col.by
    = "dark blue",
    addrow.subgroups = TRUE,
    addrow.overall = TRUE,
    addrow = TRUE,
    label.test.effect.subgroup.random = "manipulation",
    test.subgroup.random = TRUE
)
grid.text("Effect of the manipulation and sham on surface area of SH, when\n sub-grouped by manipulatio
```

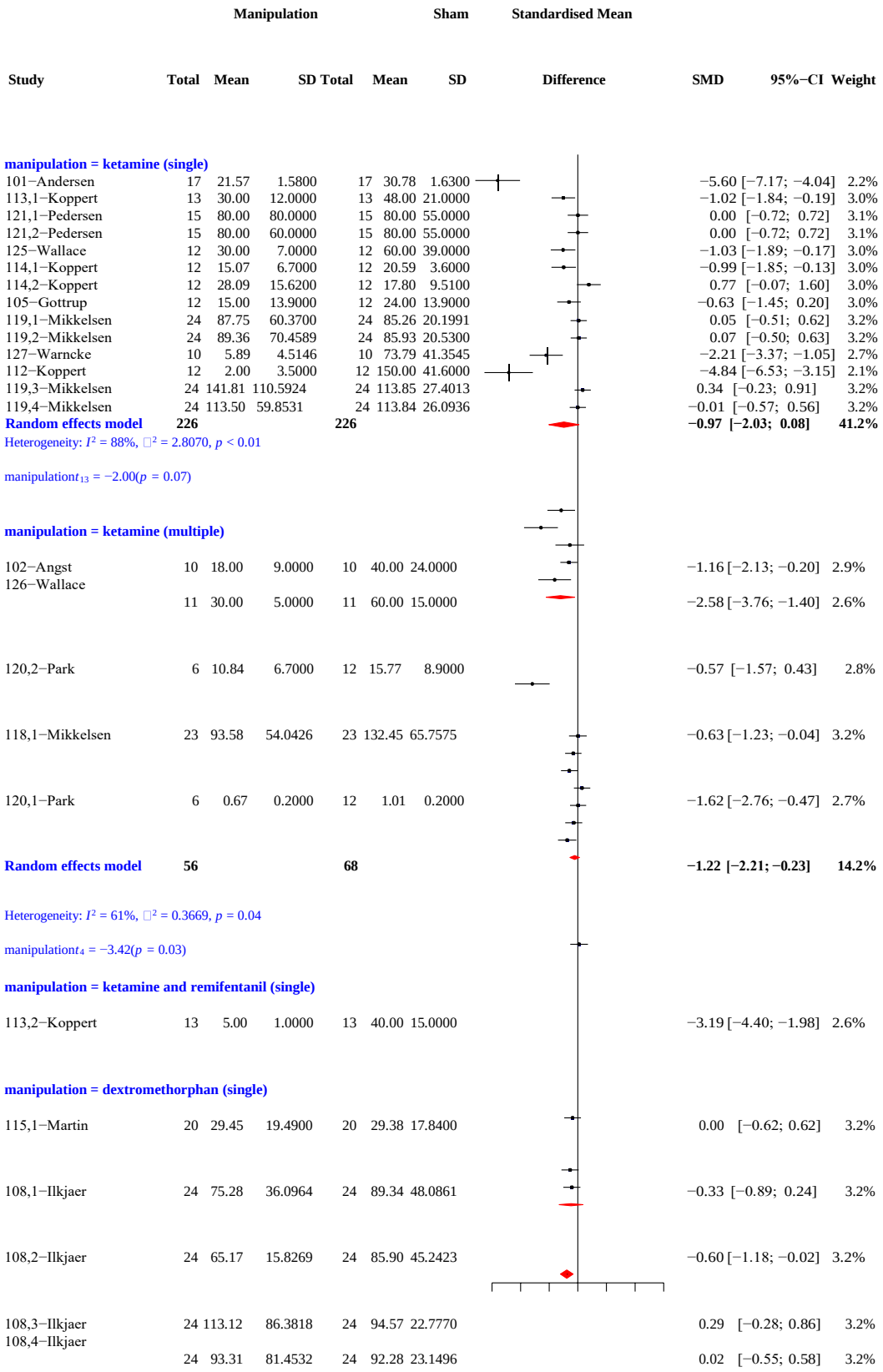
Effect of the manipulation and sham on surface area of SH, when b−grouped by manipulation in crossover and within−suject study desi(raw data; N = 25)



Random effects model	**50**			**56**			**−1.16 [−2.57; 0.26]**	**15.2%**
Heterogeneity: I^2 = 67%, τ^2 = 0.4979, p = 0.03								
manipulationt_3 = −2.60(p = 0.08)								
manipulation = ketamine and remifentanil (single)								
113,2−Koppert	13	5.00	1.0000	13	40.00	15.0000	−3.19 [−4.40; −1.98]	3.4%
manipulation = dextromethorphan (single)								
115,1−Martin	20	29.45	19.4900	20	29.38	17.8400	0.00 [−0.62; 0.62]	4.2%
108,1−Ilkjaer	24	75.28	36.0964	24	89.34	48.0861	−0.33 [−0.89; 0.24]	4.3%
108,2−Ilkjaer	24	65.17	15.8269	24	85.90	45.2423	−0.60 [−1.18; −0.02]	4.3%
108,3−Ilkjaer	24	113.12	86.3818	24	94.57	22.7770	0.29 [−0.28; 0.86]	4.3%
108,4−Ilkjaer	24	93.31	81.4532	24	92.28	23.1496	0.02 [−0.55; 0.58]	4.3%
Random effects model	**116**			**116**			**−0.12 [−0.55; 0.31]**	**21.4%**
Heterogeneity: I^2 = 27%, τ^2 = 0.0341, p = 0.24								
manipulationt_4 = −0.80(p = 0.47)								
manipulation = dextromethorphan (multiple)								
115,2−Martin	20	23.12	19.3000	20	21.55	15.9200	0.09 [−0.53; 0.71]	4.2%
manipulation = CHF3381 (single)								
116,1−Mathiesen	27	77.14	30.8152	27	106.08	45.5840	−0.73 [−1.29; −0.18]	4.3%
116,2−Mathiesen	27	172.62	81.8039	27	246.79	67.0654	−0.98 [−1.54; −0.41]	4.3%
Random effects model	**54**			**54**			**−0.85 [−2.40; 0.70]**	**8.6%**
Heterogeneity: I^2 = 0%, τ^2 = 0, p = 0.55								
manipulationt_1 = −6.98(p = 0.09)								
manipulation = ketamine and naloxone (multiple)								
118,2−Mikkelsen	23	99.36	33.1171	23	115.44	49.9330	−0.37 [−0.96; 0.21]	4.3%
Random effects model	**442**			**448**			**−0.75 [−1.26; −0.25]**	**100.0%**
Heterogeneity: I^2 = 81%, τ^2 = 1.0868, p < 0.01								
Test for subgroup differences: χ^2_6 = 38.38, df = 6 (p < 0.01)								



```r
# All sa studies

meta_sa_pub_bias <- meta::metacont(
   n.e = n.m,
   mean.e = m.mean,
   sd.e = m.sd,
   n.c = n.c,
   mean.c = c.mean,
   sd.c = c.sd,
   subgroup = manipulation,
   data = all_sa_data_converted,
   studlab = first_author,
   sm = "SMD",# standardised mean difference

   method.smd = "Hedges",# most preferred because it can better handle small sample sizes

  fixed = FALSE,
   random = TRUE,
   method.tau = "REML",
   method.random.ci = "HK",
   label.e = "Manipulation",
   label.c = "Sham",
   layout = "meta")

# use for gradepro, delete later

forest(
  meta_sa_pub_bias,
  col.square = "blue",
  col.square.lines = "black",
  col.diamond.random = "red",
  col.diamond.lines.random = "red",
  allstudies = TRUE,
  comb.fixed = FALSE,
  comb.random = TRUE,
  label.left = "Favours manipulation",
  col.label.left = "black",
  label.right = "Favours Sham",
  col.label.right = "black",
  test.effect.subgroup.random = TRUE,
  overall = TRUE,
  print.byvar = FALSE,
  col.by = "blue",
  addrow.subgroups = TRUE,
  addrow.overall = TRUE,
  addrow = TRUE,
  label.test.effect.subgroup.random = "manipulation",
  test.subgroup.random = TRUE)
```


Manipulation
Sham
Standardised Mean
Study
Total Mean SD Total Mean SD
Difference
SMD 95%−CI Weight
manipulation = ketamine (single)
101−Andersen 17 21.57 1.5800 17 30.78 1.6300 −5.60 [−7.17; −4.04] 2.2%
113,1−Koppert 13 30.00 12.0000 13 48.00 21.0000 −1.02 [−1.84; −0.19] 3.0%
121,1−Pedersen 15 80.00 80.0000 15 80.00 55.0000 0.00 [−0.72; 0.72] 3.1%
121,2−Pedersen 15 80.00 60.0000 15 80.00 55.0000 0.00 [−0.72; 0.72] 3.1%
125−Wallace 12 30.00 7.0000 12 60.00 39.0000 −1.03 [−1.89; −0.17] 3.0%
114,1−Koppert 12 15.07 6.7000 12 20.59 3.6000 −0.99 [−1.85; −0.13] 3.0%
114,2−Koppert 12 28.09 15.6200 12 17.80 9.5100 0.77 [−0.07; 1.60] 3.0%
105−Gottrup 12 15.00 13.9000 12 24.00 13.9000 −0.63 [−1.45; 0.20] 3.0%
119,1−Mikkelsen 24 87.75 60.3700 24 85.26 20.1991 0.05 [−0.51; 0.62] 3.2%
119,2−Mikkelsen 24 89.36 70.4589 24 85.93 20.5300 0.07 [−0.50; 0.63] 3.2%
127−Warncke 10 5.89 4.5146 10 73.79 41.3545 −2.21 [−3.37; −1.05] 2.7%
112−Koppert 12 2.00 3.5000 12 150.00 41.6000 −4.84 [−6.53; −3.15] 2.1%
119,3−Mikkelsen 24 141.81 110.5924 24 113.85 27.4013 0.34 [−0.23; 0.91] 3.2%
119,4−Mikkelsen 24 113.50 59.8531 24 113.84 26.0936 −0.01 [−0.57; 0.56] 3.2%
Random effects model 226 226 −0.97 [−2.03; 0.08] 41.2%
Heterogeneity: I^2 = 88%, τ^2 = 2.8070, p < 0.01
manipulation t_{13} = −2.00(p = 0.07)
manipulation = ketamine (multiple)
102−Angst 10 18.00 9.0000 10 40.00 24.0000 −1.16 [−2.13; −0.20] 2.9%
126−Wallace 11 30.00 5.0000 11 60.00 15.0000 −2.58 [−3.76; −1.40] 2.6%
120,2−Park 6 10.84 6.7000 12 15.77 8.9000 −0.57 [−1.57; 0.43] 2.8%
118,1−Mikkelsen 23 93.58 54.0426 23 132.45 65.7575 −0.63 [−1.23; −0.04] 3.2%
120,1−Park 6 0.67 0.2000 12 1.01 0.2000 −1.62 [−2.76; −0.47] 2.7%
Random effects model 56 68 −1.22 [−2.21; −0.23] 14.2%
Heterogeneity: I^2 = 61%, τ^2 = 0.3669, p = 0.04
manipulation t_4 = −3.42(p = 0.03)
manipulation = ketamine and remifentanil (single)
113,2−Koppert 13 5.00 1.0000 13 40.00 15.0000 −3.19 [−4.40; −1.98] 2.6%
manipulation = dextromethorphan (single)
115,1−Martin 20 29.45 19.4900 20 29.38 17.8400 0.00 [−0.62; 0.62] 3.2%
108,1−Ilkjaer 24 75.28 36.0964 24 89.34 48.0861 −0.33 [−0.89; 0.24] 3.2%
108,2−Ilkjaer 24 65.17 15.8269 24 85.90 45.2423 −0.60 [−1.18; −0.02] 3.2%
108,3−Ilkjaer 24 113.12 86.3818 24 94.57 22.7770 0.29 [−0.28; 0.86] 3.2%
108,4−Ilkjaer 24 93.31 81.4532 24 92.28 23.1496 0.02 [−0.55; 0.58] 3.2%

103,1−Duedahl	22	93.84	20.0996	22	99.99	21.2557	−0.29 [−0.89; 0.30]	3.2%
103,2−Duedahl	20	82.54	46.0925	20	110.73	24.9594	−0.75 [−1.39; −0.10]	3.2%
Random effects model	**158**			**158**			**−0.23 [−0.56; 0.11]**	**22.6%**

Heterogeneity: I^2 = 31%, τ^2 = 0.0395, p = 0.19

manipulationt_6 = −1.65(p = 0.15)

manipulation = dextromethorphan (multiple)

115,2−Martin	20	23.12	19.3000	20	21.55	15.9200	0.09 [−0.53; 0.71]	3.2%

manipulation = CHF3381 (single)

116,1−Mathiesen	27	77.14	30.8152	27	106.08	45.5840	−0.73 [−1.29; −0.18]	3.3%
116,2−Mathiesen	27	172.62	81.8039	27	246.79	67.0654	−0.98 [−1.54; −0.41]	3.2%
Random effects model	**54**			**54**			**−0.85 [−2.40; 0.70]**	**6.5%**

Heterogeneity: I^2 = 0%, τ^2 = 0, p = 0.55

manipulationt_1 = −6.98(p = 0.09)

manipulation = ketamine and naloxone (multiple)

118,2−Mikkelsen	23	99.36	33.1171	23	115.44	49.9330	−0.37 [−0.96; 0.21]	3.2%

manipulation = dextromethorphan and morphine (single)

104,1−Frymoyer	22	54.00	25.1000	22	68.00	25.1000	−0.55 [−1.15; 0.06]	3.2%
104,2−Frymoyer	22	73.00	25.1000	22	85.00	31.1000	−0.42 [−1.01; 0.18]	3.2%
Random effects model	**44**			**44**			**−0.48 [−1.31; 0.35]**	**6.4%**

Heterogeneity: I^2 = 0%, τ^2 = 0, p = 0.76

manipulationt_1 = −7.37(p = 0.09)

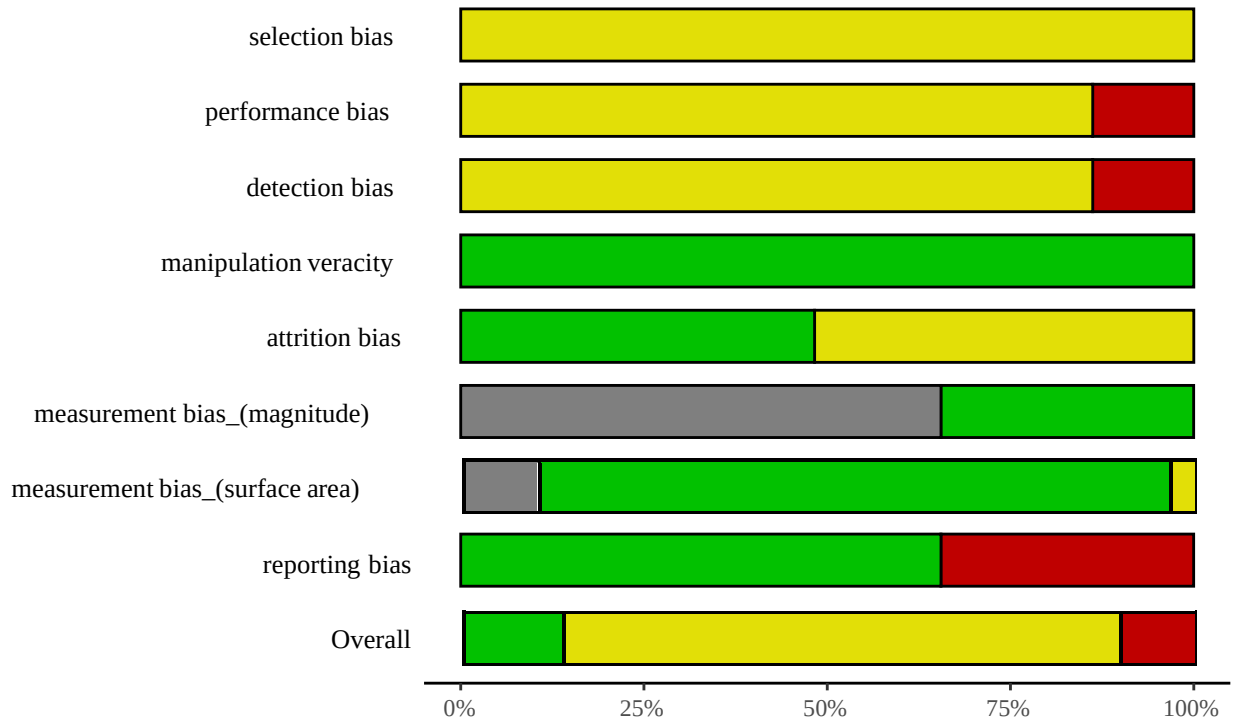
Random effects model	**594**			**606**			**−0.77 [−1.22; −0.33]**	**100.0%**

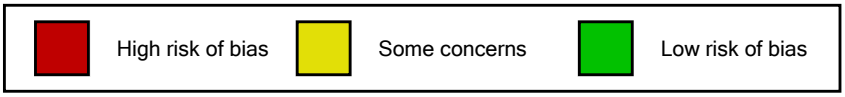
Heterogeneity: I^2 = 81%, τ^2 = 1.1329, p < 0.01

Test for subgroup differences: χ^2_7 = 39.55, df = 7 (p < 0.01)

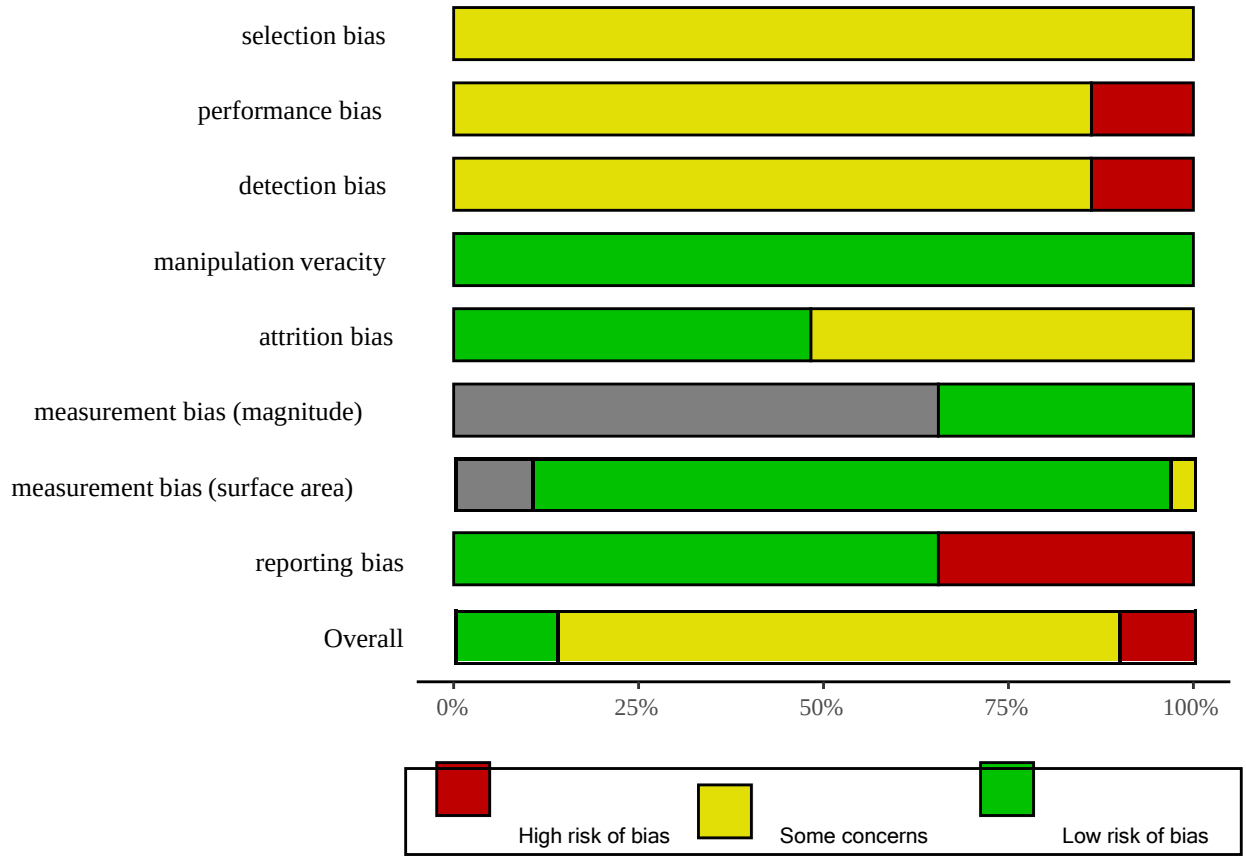
RoB analysis

```
rob <- read_excel(path =   here("sr data/mag_area_bl_fu.xlsx"),sheet = "rob")
# trial  run rob2
<- rob
names(rob2) <-

   gsub(" ", ".", names(rob2)) #  replace  gaps  in  headings

rob_trial <-
   rob2 %>% select(Study = study_id,
                           selection.bias:reporting.bias,
                           Overall = RoB_summary) #  removed  study  column  since  the  ROB1  tool  recognizes  only  spe

rob_trial$Weight <-
   rep((1:1), 29) # creating  a  weight  variable  with  a  value  of  1


rob_trial$Weight <- as.numeric(rob_trial$Weight)

rob_summary(
   data = rob_trial,tool
   = "ROB1", weighted
   = FALSE,
   # ROB1 allows for the labeling of the bias variables as is

   overall = TRUE
)
```

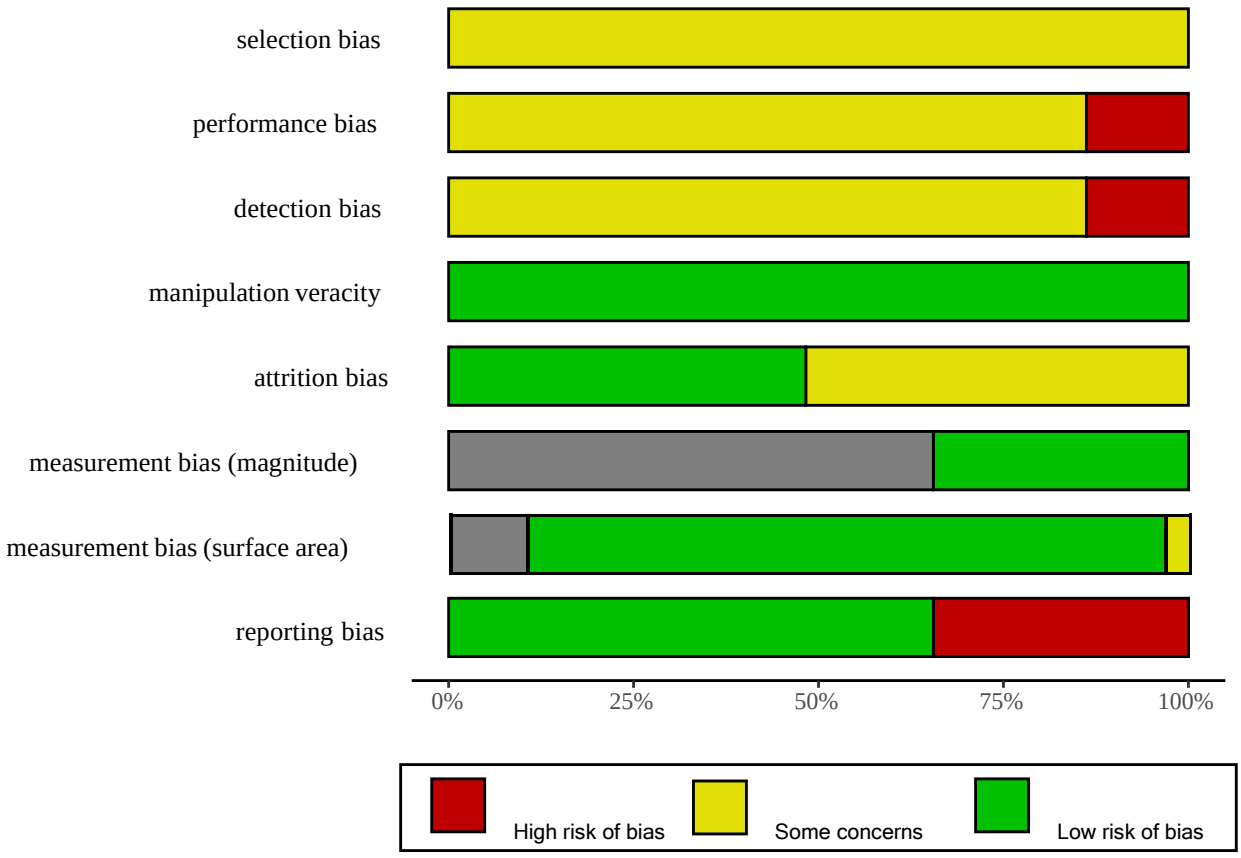



High risk of bias
Some concerns
Low risk of bias

```
#replace the '_' in bias_intensity and bias_area with a dot to have consistent labeling in the figure

rob_trial2 <- rob_trial

names(rob_trial2) <- gsub("_", ".", names(rob_trial2))
rob_summary(
  data = rob_trial2,
  tool = "ROB1",
  weighted = FALSE,
  overall = TRUE
) # grey shades represents NAs (some studies did not look at both outcomes)
```



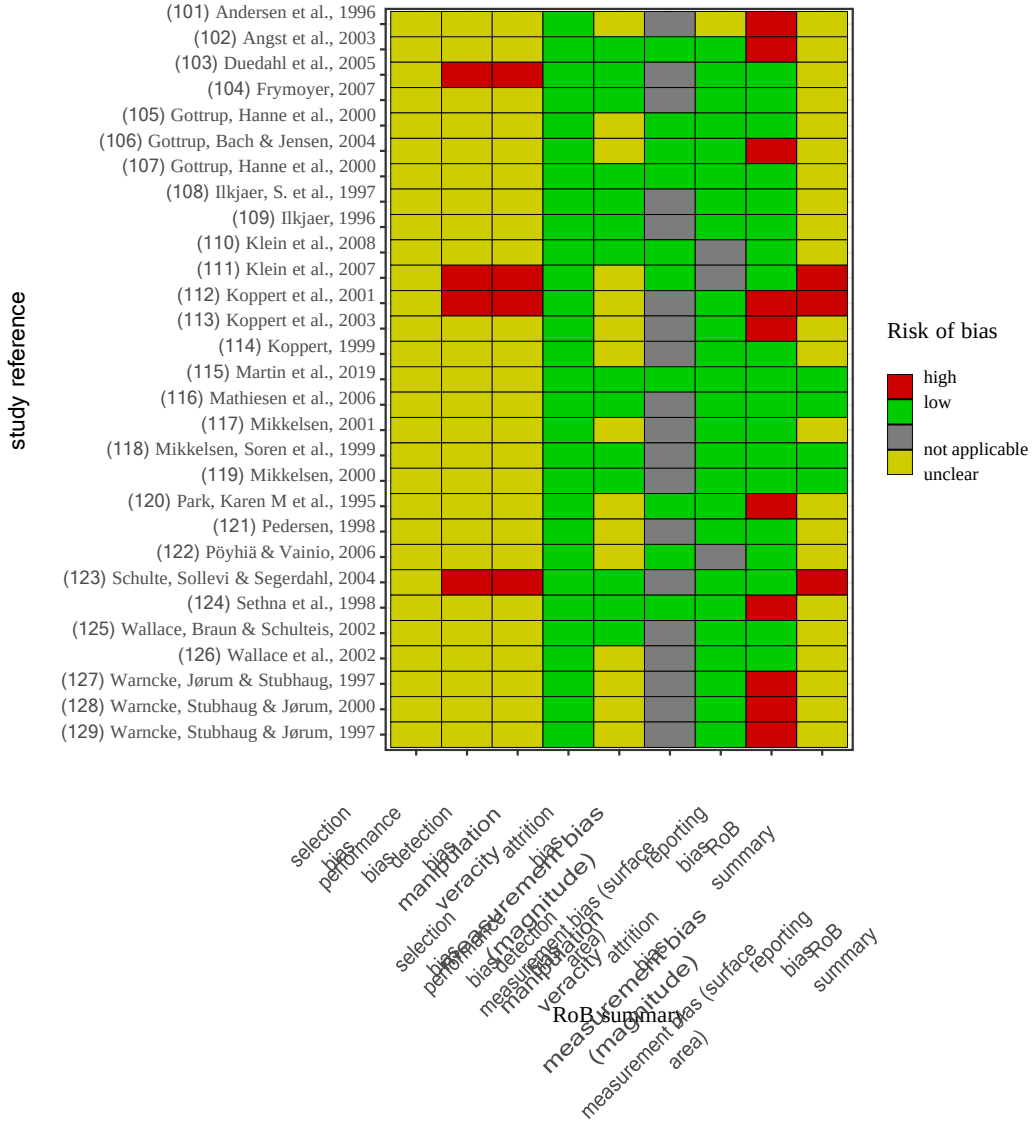

```
# remove weight category to see if it has an effect on figure

rob_trial3 <- rob_trial2
rob_trial3$Weight <- NULL

rob_summary(
  data = rob_trial3,tool
  = "ROB1", weighted
  = FALSE,# no effect
  seen overall = TRUE

)
```




```
# RoB summary table

heatmap_rob <- rob %>%
  rename(`measurement bias (magnitude)` = `measurement bias_(magnitude)`,
    `measurement bias (surface area)` = `measurement bias_(surface area)`,
    `RoB summary` = RoB_summary) %>%
  pivot_longer(cols = `selection bias`:`RoB summary`, names_to = c("RoB summary")) %>%rename(`Risk of
  bias` = value)

heatmap_rob$`RoB summary` = factor(heatmap_rob$`RoB summary`,
  levels = c("selection bias", "performance bias", "detection bias", "manipulation veracity",
    "attrition bias", "measurement bias (magnitude)","measurement bias (surface area)",
    "reporting bias",
    "RoB summary"))

heatmap <- ggplot(heatmap_rob) + aes(x = `RoB summary`, y = `study reference`,fill = `Risk of
  bias`, ) + geom_tile(colour = "black") + scale_fill_manual(drop = FALSE, values = c("red3",
  "green3",
    "#7C7C7C", "yellow3"), name = "Risk of bias") + scale_y_discrete(limits = rev) +
  theme(panel.background = element_blank(), axis.ticks = element_blank(),
    legend.title = element_blank(), legend.text = element_text(size = 8),legend.key.size = unit(0.5,
    "cm")) + coord_fixed(ratio = 0.5) +
  theme_bw(base_size = 16) + theme(axis.text.x = element_text(angle = 45,hjust = 1))

heatmap + plot_annotation(title = "Risk of bias summary: review authors' judgements about each RoB domaincluded study.
```

```
')
```

Risk of bias summary: review authors' judgements about each RoB domain for eachincluded study.

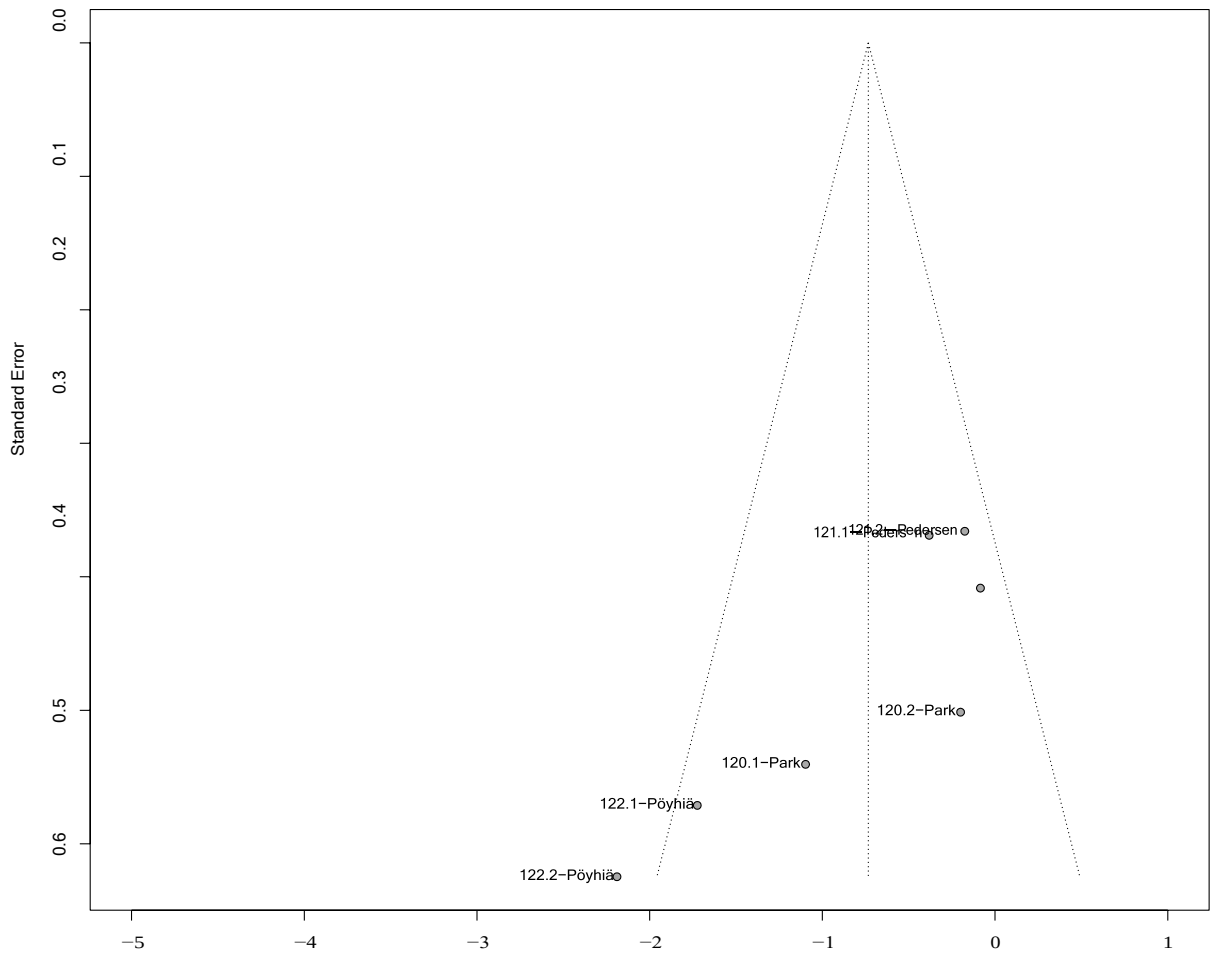


Publication bias

Magnitude of SH

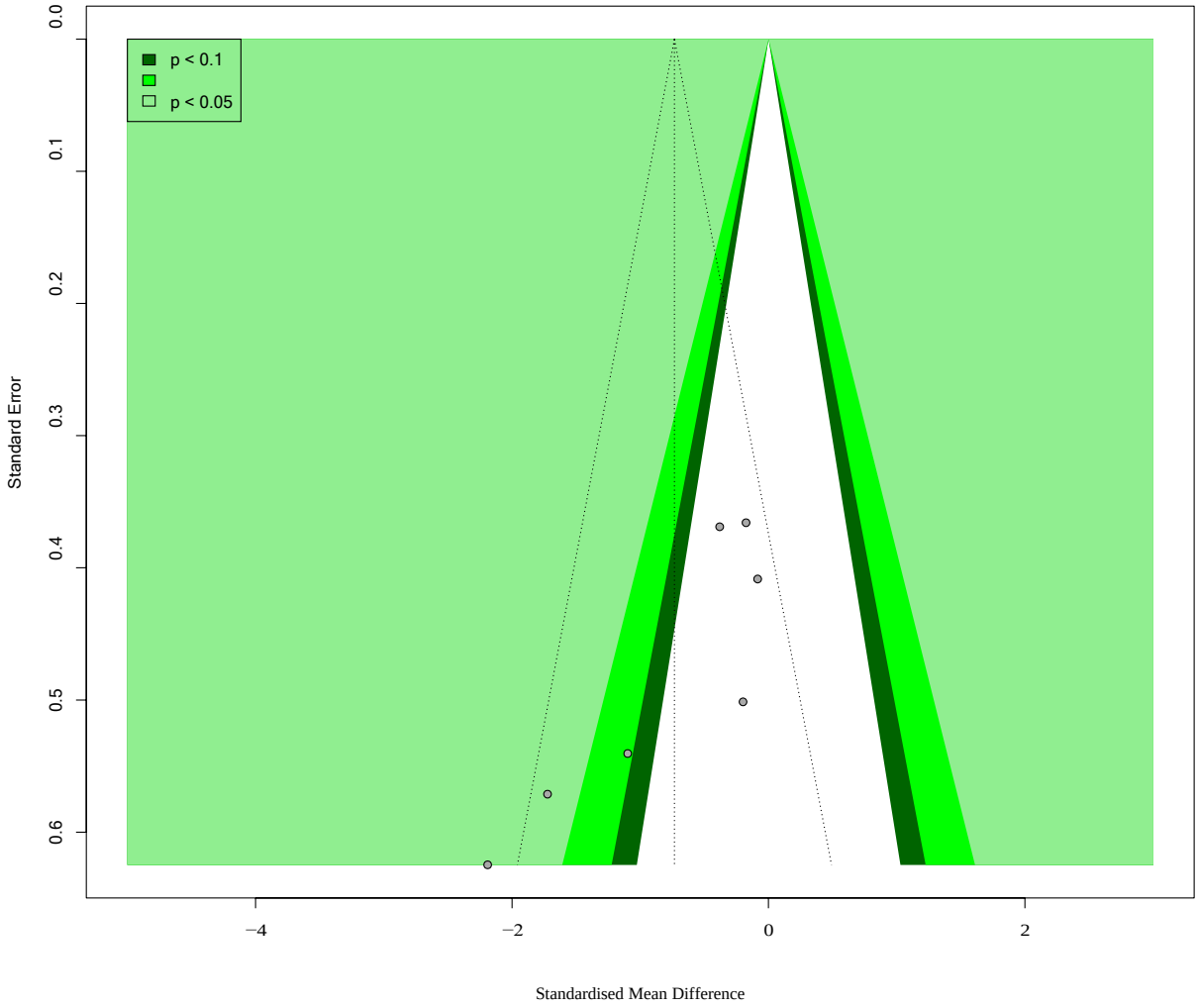
```r
## Magnitude of SH

meta_pub_bias_mag <- meta::metacont(
  n.e = n.m,
  mean.e = m.mean,
  sd.e = m.sd,
  n.c = n.c,
  mean.c = c.mean,
  sd.c = c.sd,
  studlab = first_author,


## Magnitude of SH

meta_pub_bias_mag <- meta::metacont(
  n.e = n.m,
  mean.e = m.mean,
  sd.e = m.sd,
  n.c = n.c,
  mean.c = c.mean,
  sd.c = c.sd,
  studlab = first_author,
```

```
  data = mag_raw_ratings_conditions,sm =
  "SMD",
  #  standardised  mean  difference

  method.smd = "Hedges",
  # most preferred because it can better handle small sample sizes

  fixed = FALSE, random
  = TRUE, method.tau =
  "REML",
  method.random.ci = "HK",
  label.e = "Manipulation",
  label.c = "Sham",
  layout = "meta")

# prepare  publication  bias  plots

funnel.meta(meta_pub_bias_mag,
            xlim = c(-5, 1),
            studlab = TRUE) # method 1 ; not quite clear but we can see that average effect size = -0.

 title("Funnel Plot: Publication bias of studies assessing the magnitude of SH using crossover study de
```



```r
# data set shows an asymmetrical pattern in the funnel plot that might be
# indicative of publication bias.now generate contour-enhanced funnel plot
# to see how asymmetry patterns relate to statistical significance

# Define fill colors for contour

col.contour_mag = c("darkgreen", "green", "lightgreen")

# Generate funnel plot (we do not include study labels here)

 funnel.meta(
   meta_pub_bias_mag,
   xlim = c(-5, 3),
   contour = c(0.9, 0.95, 0.99),
   col.contour = col.contour_mag)
 legend(
   x = -5,
   y = 0.0,
   legend = c("p < 0.1", "p < 0.05", "p < 0.01"),
   fill = col.contour_mag
 )# Add a legend

 title("Contour-Enhanced Funnel Plot (Publication bias of studies assessing the magnitude of SH)")


# data set shows an asymmetrical pattern in the funnel plot that might be
# indicative of publication bias.now generate contour-enhanced funnel plot
# to see how asymmetry patterns relate to statistical significance

# Define fill colors for contour

col.contour_mag = c("darkgreen", "green", "lightgreen")

# Generate funnel plot (we do not include study labels here)

 funnel.meta(
   meta_pub_bias_mag,
   xlim = c(-5, 3),
   contour = c(0.9, 0.95, 0.99),
   col.contour = col.contour_mag)
 legend(
   x = -5,
   y = 0.0,
   legend = c("p < 0.1", "p < 0.05", "p < 0.01"),
   fill = col.contour_mag
 )# Add a legend

 title("Contour-Enhanced Funnel Plot (Publication bias of studies assessing the magnitude
```

Contour−Enhanced Funnel Plot (Publication bias of studies assessing the magnitude of SH)




```
# NOTE: We are particularly interested in the p<0.05 and p<0.01 regions,
# # because effect sizes falling into this area are traditionally considered significant.

## assess funnel plots asymmetry (Egger's regression test)
```

Linear regression test of funnel plot asymmetry##
Test result: t = -3.61, df = 5, p-value = 0.0154##
Sample estimates:
bias se.bias intercept se.intercept## -
6.3261 1.7537 2.2275 0.7977
##
Details:
- multiplicative residual heterogeneity variance (tau^2 = 0.8390)## - predictor:
standard error

```
## - weight:            inverse variance
## - reference: Egger et al. (1997), BMJ
```

```
# number of studies required to run test is 10 and we have 7, so we have to adjust kmin to 7

# which method.bias is correct https://www.rdocumentation.org/packages/meta/versions/4.9-7/topics/metab

eggers.test(meta_pub_bias_mag)
```

```
##  Eggers'  test  of  the  intercept ##
============================
=##
##  intercept       95% CI      t         p##   -6.326 -9.76
- -2.89 -3.607 0.01542415##
## Eggers' test indicates the presence of funnel plot asymmetry.
```

```
# Therefore, eggers test could not be performed because it was lack statistical
# power since we have few studies (k =7)

# method 2; no effect, data in the funnel plot is indeed asymmetrical
# I^2 heterogeneity observed in this meta-analysis meta_pub_bias_mag$I2
# I^2 = 60 %
```

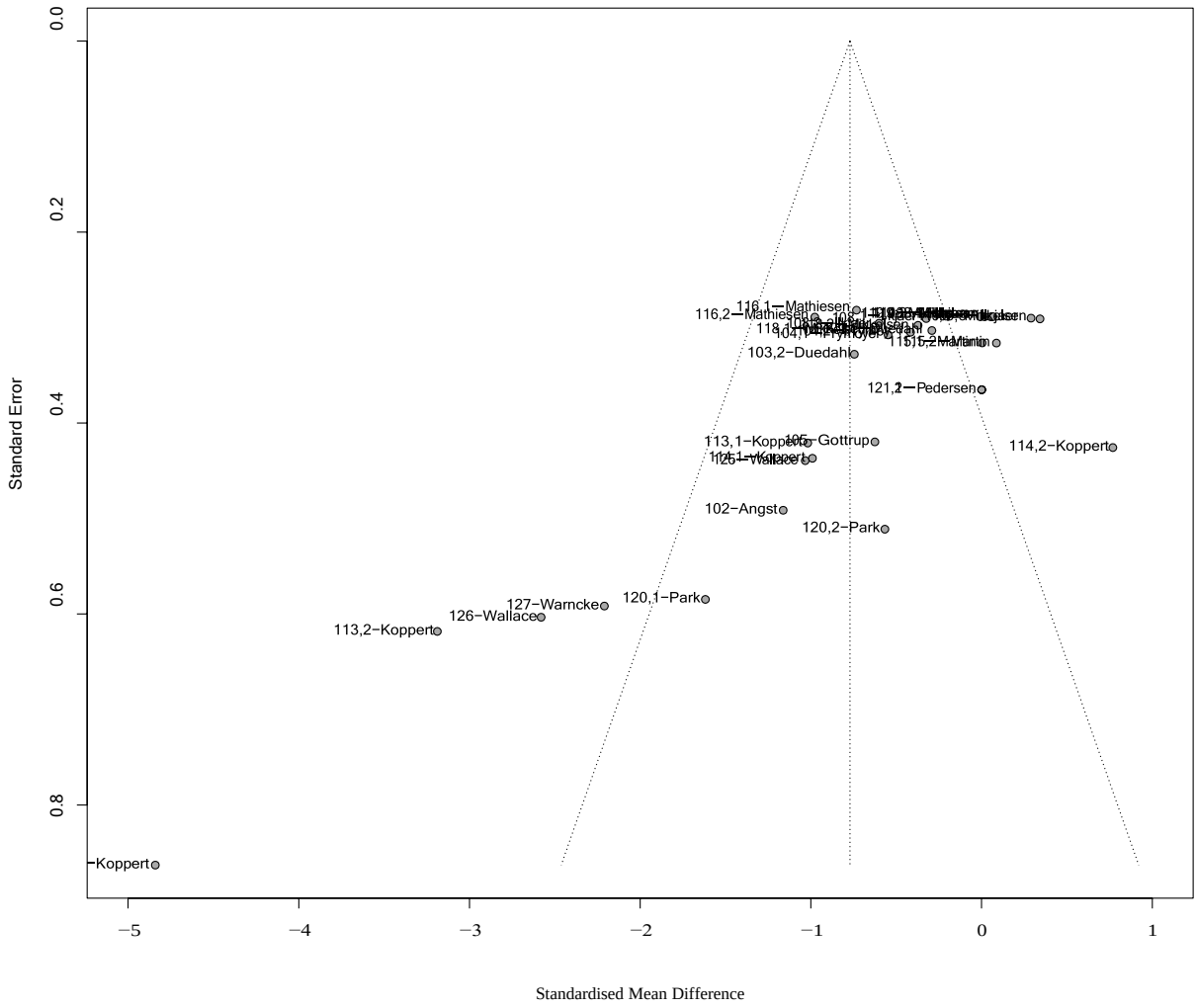
```
## [1] 0.602961
```

```
# the heterogeneity in our analysis is substantial
```

Surface area of SH

```
## Surface area of SH Produce funnel plot prepare## publication
bias plots
funnel.meta(meta_sa_pub_bias, xlim = c(-5, 1), studlab = TRUE)  # method 1 ; not quite clear but we can

title("Funnel Plot: Publication bias of studies assessing the surface area of SH using crossover and wi
```

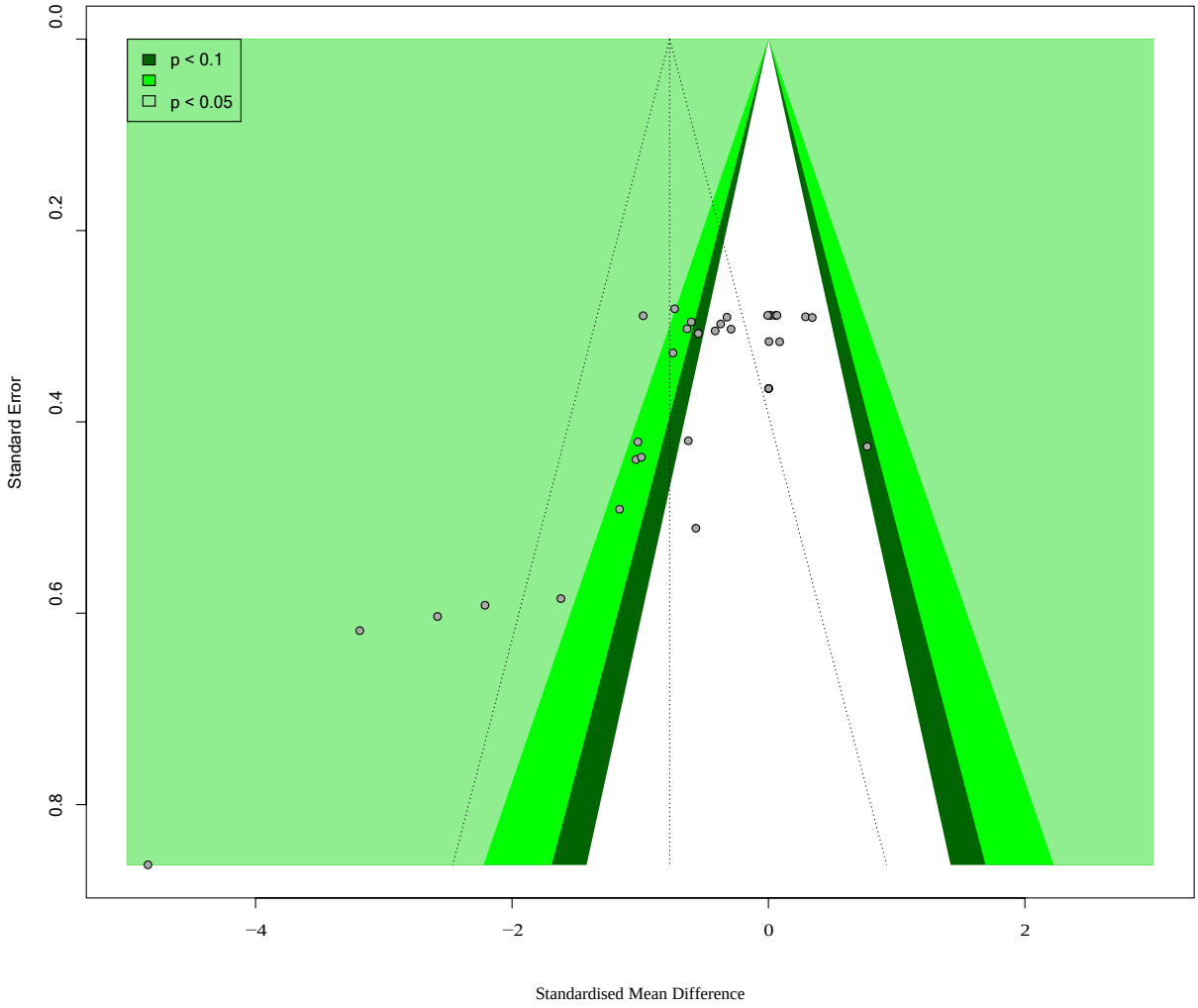
Funnel Plot: Publication bias of studies assessing the surface area of SH using crossover and within subject study designs




```
# data set shows an asymmetrical pattern in the funnel plot#
that might be indicative of publication bias (for both

# outcomes). now generate contour-enhanced funnel plot to#
see how asymmetry patterns relate to statistical

#   significance



# surface area Define fill colors for contour

col.contour_area = c("darkgreen", "green", "lightgreen")

# Generate funnel plot (we do not include study labels#
here)

funnel.meta(meta_sa_pub_bias, xlim = c(-5, 3), contour = c(0.9,0.95, 0.99),
```

```
    col.contour = col.contour_area)
legend(x = -5, y = 0, legend = c("p < 0.1", "p < 0.05", "p < 0.01"),fill =
    col.contour_area)   # Add a legend
title("Contour-Enhanced Funnel Plot (Publication bias of studies assessing the Surface area of SH)")
```

Contour−Enhanced Funnel Plot (Publication bias of studies assessing the Surface area of SH)




```
# NOTE: We are particularly interested in the p<0.05 and

# p<0.01 regions, # because effect sizes falling into this
# area are traditionally considered significant.



## assess funnel plots asymmetry (Egger's regression test)
## surface area

metabias(meta_sa_pub_bias, method.bias = "linreg")  # method 1
# which method.bias is correct

#       https://www.rdocumentation.org/packages/meta/versions/4.9-7/topics/metabias


# eggers.test(meta_sa_pub_bias) method 2; no effect, data
# in the funnel plot is indeed asymmetrical I^2
```

```
## [1] 0.8090836
# the heterogeneity in our analysis is substantial
```

Appendix D: data conversion for meta-analysis

Original reported data							
Study id	**First author**	**Manipulation group** (mean)	**Manipulation group** SD (SEM)	**Control group** (mean)	**Control group** SD (SEM)	**Area units**	**Formulae to convert to cm^2**
101	Andersen	2,62	0,71	3,13	0.72	cm	Area = πr^2
114.1	Koppert	21,9	14,6	25,6	10.7	mm	**step 1**: use *Area* = πr^2 to convert data to mm^2 **step 2**: divide answer from step 1 by 100 to convert to data to cm^2
114.2	Koppert	29,9	22,3	23,8	17.4	mm	**step 1**: use *Area* = πr^2 to convert data to mm^2 **step 2**: divide answer from step 1 by 100 to convert
120.1	Park	67	(7)	101	(7)	mm^2	divide data by 100
120.2	Park	1084	(273)	1577	(256)	mm^2	divide data by 100
Converted data							
101	Andersen	21,57	1,58	30,78	1,63	cm^2	
114.1	Koppert	15,07	6,7	20,59	3,6	cm^2	
114,2	Koppert	28,09	15,62	17,8	9,51	cm^2	
120.1	Park	0,67	(0,07)	1,01	(0,07)	cm^2	
120.2	Park	10,84	(2,73)	15,77	(2,56)	cm^2	

Appendix E: sample size calculations

Correlation sample size

Total sample size required to determine whether a correlation coefficient differs from zero.

Instructions: Enter parameters in the **green** cells. Answers will appear in the **blue** box below.

α (two-tailed) =	0.05	Threshold probability for rejecting the null hypothesis. Type I error rate.
β =	0.15	Probability of failing to reject the null hypothesis under the alternative hypothesis. Type II error rate.
r =	0.40	The expected correlation coefficient.

Calculate

The standard normal deviate for α = Z_α = 1.9600
The standard normal deviate for β = Z_β = 1.0364
C = 0.5 * ln[(1+r)/(1-r)] = 0.4236

Total sample size = N = $[(Z_\alpha+Z_\beta)/C]^2 + 3$ = **53**

Screenshot 1: sample size calculations

Correlation sample size

Total sample size required to determine whether a correlation coefficient differs from zero.

Instructions: Enter parameters in the **green** cells. Answers will appear in the **blue** box below.

α (two-tailed) =	0.05	Threshold probability for rejecting the null hypothesis. Type I error rate.
β =	0.20	Probability of failing to reject the null hypothesis under the alternative hypothesis. Type II error rate.
r =	0.40	The expected correlation coefficient.

Calculate

The standard normal deviate for α = Z_α = 1.9600
The standard normal deviate for β = Z_β = 0.8416
C = 0.5 * ln[(1+r)/(1-r)] = 0.4236

Total sample size = N = $[(Z_\alpha+Z_\beta)/C]^2 + 3$ = **47**

Screenshot 2: illustrates how many participants may be lost to follow-up before we can interpret the results meaningfully.

Appendix F: adapted brief pain inventory

Description: The English version of the BPI used in the parent study.

Brief Pain Inventory

Throughout our lives, most of us have had pain from time to time (such as minor headaches, sprains, and toothaches). Have you had pain other than these everyday kinds of pain during the last week?	○ Yes ○ No

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Please select all the zones where the participant reports having pain this last week	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Which site(s) is the most problematic site(s) this past week?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

	0 - No pain	1	2	3	4	5	6	7	8	9	10 - Pain as bad as you can imagine
Please use this scale to describe your pain at its WORST in the last week	○	○	○	○	○	○	○	○	○	○	○
Please use this scale to describe your pain at its LEAST in the last week	○	○	○	○	○	○	○	○	○	○	○
Please use this scale to describe your pain on AVERAGE in the last week	○	○	○	○	○	○	○	○	○	○	○
Please use this scale to describe how much pain you have RIGHT NOW	○	○	○	○	○	○	○	○	○	○	○

Throughout our lives, most of us have had pain from time to time (such as minor headaches, sprains, and toothaches). Have you had pain other than these everyday kinds of pain for most days (4 or more days) of the week for the last 3 months?	○ Yes ○ No

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Please select all the sites where the participant reports having had pain for most days of the week for the past 3 months.	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Which ongoing pain site(s) is the most problematic?	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

	0 - No pain	1	2	3	4	5	6	7	8	9	10 - Pain as bad as you can imagine
Please use this scale to describe your pain at its WORST in the last week	○	○	○	○	○	○	○	○	○	○	○
Please use this scale to describe your pain at its LEAST in the last week	○	○	○	○	○	○	○	○	○	○	○
Please use this scale to describe your pain on AVERAGE in the last week	○	○	○	○	○	○	○	○	○	○	○
Please use this scale to describe how much pain you have RIGHT NOW	○	○	○	○	○	○	○	○	○	○	○

	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
In the last week, how much relief have pain treatments or medications provided? Please circle the one percentage that most shows how much relief you have received	○	○	○	○	○	○	○	○	○	○	○

	0 - Does not interfere	1	2	3	4	5	6	7	8	9	10 - Completely interferes
Choose the one number that describes how much, during the past week, pain has interfered with your: GENERAL ACTIVITY	○	○	○	○	○	○	○	○	○	○	○
Choose the one number that describes how much, during the past week, pain has interfered with your: MOOD	○	○	○	○	○	○	○	○	○	○	○
Choose the one number that describes how much, during the past week, pain has interfered with your: WALKING ABILITY	○	○	○	○	○	○	○	○	○	○	○
Choose the one number that describes how much, during the past week, pain has interfered with your: NORMAL WORK (includes both work outside the home and housework)	○	○	○	○	○	○	○	○	○	○	○
Choose the one number that describes how much, during the past week, pain has interfered with your: RELATIONS WITH OTHER PEOPLE	○	○	○	○	○	○	○	○	○	○	○
Choose the one number that describes how much, during the past week, pain has interfered with your: SLEEP	○	○	○	○	○	○	○	○	○	○	○
Choose the one number that describes how much, during the past week, pain has interfered with your: ENJOYMENT OF LIFE	○	○	○	○	○	○	○	○	○	○	○

Other than HIV, do you have any other health problems?
○ Yes
○ No

If yes, specify
_____

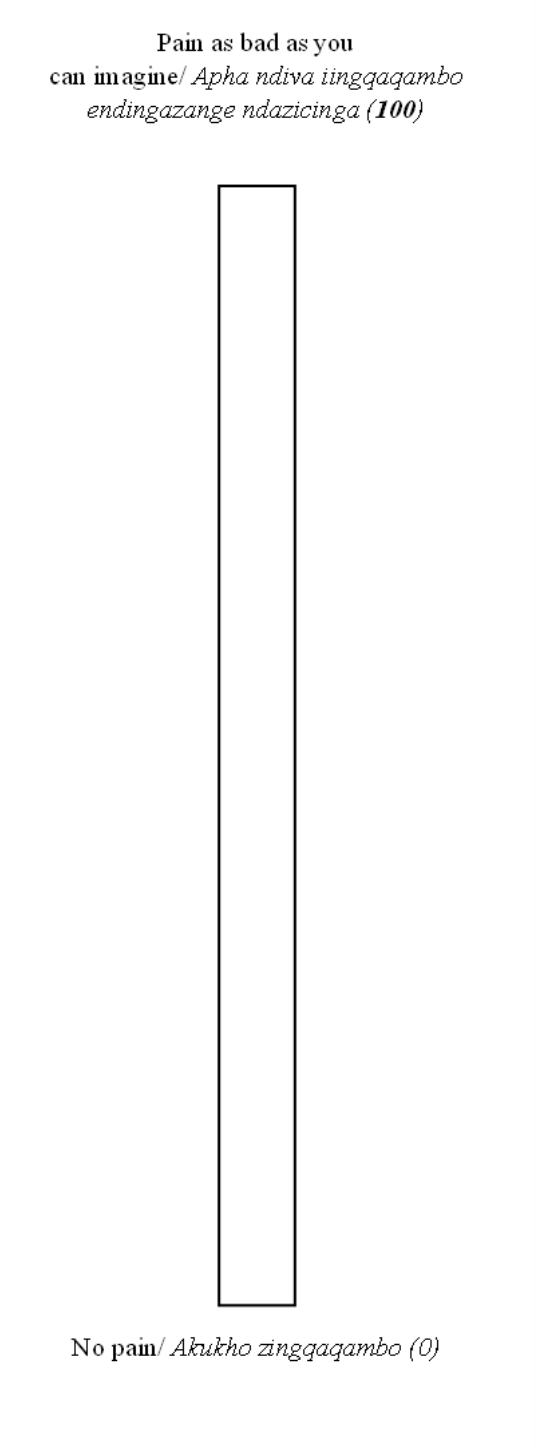
Have you had to go to hospital in the past 6 months?
○ Yes
○ No

If yes, specify
_____

Have you had any infections in the past 6 months?
○ Yes
○ No

Appendix G: electronic VAS

Description: The English and isiXhosa version of the electronic VAS.



Appendix H: hopkins symptom checklist-25

Description: listed below are some symptoms of strain that people sometimes have. Please read each one carefully and check the answer which best reflects how much that symptom has bothered you during the past month.

	Questions	1 Not at all	2 A little	3 Quite a bit	4 Extremely
HOP1	Suddenly scared for no reason				
HOP2	Feeling fearful				
HOP3	Faintness, dizziness, or weakness				
HOP4	Nervousness or shakiness inside				
HOP5	Heart pounding or racing				
HOP6	Trembling				
HOP7	Feeling tense or keyed up				
HOP8	Headaches				
HOP9	Spells of terror or panic				
HOP10	Feeling restless, can't sit still				
HOP11	Feeling low in energy--slowed down				
HOP12	Blaming yourself for things				
HOP13	Crying easily				
HOP14	Loss of sexual interest or pleasure				
HOP15	Poor appetite				
HOP16	Difficulty falling asleep, staying asleep				
HOP17	Feeling hopeless about the future				
HOP18	Feeling blue				
HOP19	Feeling lonely				
HOP20	Feeling trapped or caught				
HOP21	Worrying too much about things				
HOP22	Feeling no interest in things				
HOP23	Thoughts of ending your life				
HOP24	Feeling everything is an effort				
HOP25	Feelings of worthlessness				

Appendix I: study information sheet and consent form

CONSENT TO PARTICIPATE IN AN ADDITIONAL COMPONENT OF A RESEARCH STUDY

Title of the main Research Project: An observational study of shared variation and reciprocal influences: distress, immune function and pain in HIV

Title of the additional component: Relationships between induced skin sensitivity, inflammatory reactivity, clinical pain and distress in people living with HIV.

PRINCIPAL INVESTIGATORS: Drs Tory Madden, John A. Joska, Romy Parker and Jonny Peter
ADDRESS: Department of Anaesthesia and Perioperative Medicine, D23, Groote Schuur Hospital, Anzio Road, Observatory, 7925
Tel: 021 650 3683

KEY INFORMATION: ADDITIONAL PROCEDURE

You are being invited to take part in an additional part of the main research project. I will read through this form with you, to describe the details of the additional part. I will only give you the information on this additional part when the information is different to the main study that we have just discussed. As for the main study, you are free to choose to participate or not, and to change your mind and stop participating, without any negative effects.

What is this additional procedure about?

In this part, we are interested in whether the system that protects you from disease and illness influences sensitivity to painful and non-painful events. The purpose of the research is to understand how sensitivity, distress, pain, and the system that protects you from disease and illness interact in people living with HIV.

What will happen in this additional procedure?

- If you agree to participate and sign the form, you will be asked to stay for a longer period of time at your first study visit, so that you can participate in this section of the research.
- First, you will be shown some instruments that are used to test the sensitivity of your skin. The researcher will touch these instruments to your skin and ask you to report what you feel each time. Each touch takes about 1 second, and will be repeated several times. One of the tests also involves a small, very brief electrical stimulation that lasts for less than 1 second each time. Some of the tests might be painful.
- Next, you will receive 5 intense electrical stimulations. Each stimulation lasts for 1 second, and they are separated by 9-second breaks. This whole process lasts about 1 minute. The intense stimulations can be quite painful and require some effort to tolerate, so the researcher will help you to count your way through them. If they are too painful and you want to stop, you can say ‘STOP’ to the researcher and they will stop the stimulations immediately.
- Finally, the researcher will test the sensitivity of your skin 4 more times over the next hour. The tests will be the same as the ones you received before the intense electrical stimulation, but they may feel a bit different.
- This additional procedure will take about an extra 90 minutes.
- This procedure only happens once, at Visit 1 of the main research study.
- Your skin may be a little red and sensitive for a few hours after the procedure. For most people, the sensitivity resolves by the next day. If you have any concerns, you should tell the researcher.

What will your responsibilities be?

- You must give honest answers to the questions the researcher asks, and follow the researcher's instructions. The researcher's main priority is to keep you safe, so it is important that you answer questions truthfully and follow his/her instructions.
- If you don't understand something the researcher says, it is important that you tell him/her so that he/she can explain again.

Are there risks involved in this additional part of the research study?

Electrical stimulation has some risks:

- If you are pregnant, or if you have heart problems, an electrical implant (e.g. a pacemaker), or a metal implant (e.g. screws/plates) in your arm, then it is not safe for you to receive electrical stimulation. It is very important that you answer all the researcher's questions truthfully, so that we can make sure that it is safe for you to participate.
- The brief tests can be uncomfortable or a little painful.
- The intense electrical stimulation is quite painful. The researcher will help you count your way through the 5 events of intense electrical stimulation.
- There is a very small risk of burn to your skin from the electrical stimulation. However, we use very strict controls to prevent this.
- Your skin will probably be sensitive for some hours after the procedure.
- These are the main risks. You should feel free to mention your feelings or concerns to any member of the study team.

Will you be paid to take part in this additional part and are there any costs involved?

No, you will not be paid to take part in this additional part, but the extra time will be covered with an extra R150 sent by the cashless system, on top of your reimbursement for the main study.

All other information, including the study contact information, remains the same as for the main study.

Explain-back task to confirm comprehension and informed status
Additional procedure

Ask volunteer to answer each of the following questions (in any order) to confirm understanding:

Now that we have been through that information, I need to ask you a few questions so that I can be sure I have explained well enough. If any of the questions are difficult, I can explain again.

NB: if participant unable to answer a certain question, stop there and revise information with them. Return to questions. ALL questions must be answered correctly for participant to be included. If 3 attempts to explain and question fail, participant must be excluded for lack of comprehension.		
Question	**Answer**	***Tick***
Is your clinical care affected by your decision to participate in this additional part of the study?	No	
What are the risks of participating in this additional procedure?	Painful procedure May make my skin red and sensitive for a short time Small risk of burn	
What part of your body will we use for this additional procedure?	One forearm	
How much extra time will this additional procedure take?	About 90 minutes	
Are you allowed to change your mind about participating in this part of the study?	Yes	
If you find the intense electrical stimulation too painful, what should you say to the researcher?	STOP	

You can withdraw from the study at any time without negative consequences and continue receiving care at this clinic.

CONSENT FORM: ADDITIONAL PROCEDURE

Declaration by participant

By signing below, I, ……..……. agree to take part in the additional procedure entitled: "Relationships between induced skin sensitivity, inflammatory reactivity, clinical pain and distress in people living with HIV".

I declare that:
• I have read or had read to me this additional information and consent form and it is written in a language with which I am comfortable.
• I have had a chance to ask questions and all my questions have been adequately answered.
• I understand that taking part in this study is voluntary and I have not been pressurised to take part.
• I may choose to leave the study at any time and will not be penalised or prejudiced in any way.
• I may be asked to leave the study before it has finished, if the researcher feels it is in my best interests, or if I do not follow the study plan, as agreed to.

Confirm consent to the following procedures with participant's initials:

• _____ Skin sensitivity procedure

Signed at Town Two, Khayelitsha on [date] ……...…….. 2021.

.....
Signature of participant

Declaration by researcher

I (name) ……..…… declare that:

• I explained the information in this document to ……..
• I encouraged him/her to ask questions and took adequate time to answer them.
• I am satisfied that s/he adequately understands all aspects of the research, as confirmed by his/her answers to the questions.
• I will maintain confidentiality at all times.

Signed at Town Two, Khayelitsha on [date] ……...…….. 2021.

.....
Signature of researcher

Appendix J: english and isixhosa procedure scripts

Description: English script

Preparation

Step 1: Disinfect all working areas in the lab.

Step 2: Check that all equipment is available, then set up equipment.

a. Stimulation equipment checklist:
   - DS7A
   - Surface electrode
   - Electrode
   - PC and touch screen monitor

b. QST equipment
   - Pinprick stimulators (128 and 256 Nm),
   - Von Frey 32mN, 128mN
   - Brush (Prime Art Bianco Flat 16)

c. Additional items:
   - marker pens,
   - mark-up template,
   - Electrode stickers,
   - Alcohol balls and
   - Gauze swabs.
   - Stopwatch
   - Alcohol solution to clean pinpricks

Step 3: Connect and set up equipment.

- DS7A input connected at the back, turned on, settings: voltage max, x1, current 0.01, pulse width 2000ms.
- Connect surface electrode and electrode to DS7A using the Red and black cable.
- Connect the NI device to the DS7A: AI to OUT on DS7A and AO to IN on DS7A.
- Connect NI Device via USB port into PC (pc running affect5)
- Open Affect5 file for HFS (UKOMELELA)
- Set out QST equipment, remove covers.
- Prepare marker pens, mark-up template, electrode stickers, alcohol balls and gauze swabs.
- Prepare forms: info sheet, consent, screening, participation record.

Step 4: Ensure you have hard copies of IFC documents and study information sheet.

Step 5: Switch on the switch on the plug once everything is connected.

[Insert Participant code]

Would you like me to speak in English or isiXhosa? **[if isiXhosa]** isiXhosa isn't my home language, but I will try my best. Please let me know if anything I say doesn't make sense.

Introduction

Welcome and thank you for participating in this additional procedure. Before we begin, please switch off your phone and make sure there are no alarms or alerts that will turn it back on during the study. I will be reading the instructions from this script to ensure that all procedures are performed correctly.

Which arm did Sister Mtingeni take blood from? Ok, we will use the other arm. Please remove any items on your arm or wrist.

[Write arm allocation on screen and participant record sheet (L or R)]

This procedure will take about 90 minutes. First, I will explain the procedure. Next, we will do some skin tests on this [indicate] forearm. Then, you will receive intense electrical stimulation to the same forearm.

There will be a 30-minute break. After that, we will repeat the skin tests several times in the 40 minutes. So, it is skin tests, intense electrical stimulation, 30-minute break, and more skin tests. Do you have any questions?

[complete participation record]

Demographic Info

I am not allowed to know anything about your medical information. So, when I ask you the following questions, please answer YES/NO only.

[Remember if any of these questions are true for participant then they will not participate in the study]

- Do you have any pain in this arm today? If yes schedule for another day.
- Have you taken any medication beside your routine medication?
- [Females only] Are you pregnant or do you suspect you might be pregnant?
- Do you have a metal or electrical implant, such as a pacemaker?
- Do you have problems with skin healing?

[check for tattoos on the arm in the area between the two electrodes]

[Give info sheet and consent form]

Sister Mtingeni already read you the study information sheet and the consent form. Are you still comfortable with what those said? You are allowed to change your mind at any time. I want you to be familiar with the tools we will be using in this procedure **[like this brush].** This is the electrode we will use for stimulation. If you look closer, you will see that it has small blunt pins. It will sit on your skin like this **[indicate using other electrode]** but because it is blunt it is not painful, and it will not penetrate your skin.

[when done] Do you have any questions?

Even though you have signed consent, you're allowed to withdraw from the study at any time during the procedure – you just have to tell me you want to stop. If you withdraw, you can decide whether you will allow us to use the data that we have collected or if you'd like us to destroy it.

Before we start, I want to check whether you need to go to the bathroom? Unfortunately, you will not be able to leave the room to use the bathroom during the procedure. You will have to wait until the procedure is finished.

Skin marking

Now I'm going to draw dots on your arm like this **[show picture]** for the testing procedure. Is that ok? **[mark up using foam template, with A towards the cubital fossa; centre of radial lines approx 8cm from cubital fossa but NOT on prominent vein or scarred area.]**

[label radial lines A-H]

[Wipe forearm and upper arm with alcohol balls]

As I showed you, this is the electrode we use. [show electrode to participant] I will place it on your skin like this, and it will stay there for the whole procedure. This electrode has short pins, but they are not painful, and they do not penetrate your skin.

[Strap electrode goes around upper arm. Disc electrode approx 8cm from cubital fossa]

~Press Enter~

Detection threshold testing

[TURN switch of DS7A UP]

Now, for this next test, I will ask you to tell me when you feel the electrical stimulus. I will give you the electrical stimulus, starting at an intensity of zero. I will slowly increase the intensity until you tell me that you can feel it. Please say "yes" if you feel it, even a little bit. It will feel like a very tiny pinprick in this area **[indicate to electrode]**

[Step up by 1, step down by 0.5, then step up by 0.2]

[find threshold for arm, write current on monitor/check screen, set DS7A to that level and then flick switch to x10].

[Leave switch on]

~Press Enter~

Scale Intro

We use this scale in the procedure, so it is important that you understand it. It helps us report what you are feeling because every person feels differently.

[Show VAS on the screen and paper]

- This scale runs from this line on the left **[gesture],** which is no pain, to this line on the right **[gesture]**, which is pain as bad as you can imagine.
- If you select here on the left point **[indicate]** it means the trial was not painful even if you felt something.
- If you select a rating anywhere on the flat line **[indicate]** it means the trail was painful.
- If you select here on the right point, it means the trail was pain as bad as you can imagine.
- So, this means if you select here on the left point, it was not painful even if you felt something. As you move closer to the right side, it means it becomes more painful. Then if you select here on the right point it means it was the most intense pain as bad as you can imagine.

So, each time I test your skin with a tool, like this brush **[indicate]** I will ask you to rate each stimulation on the scale, on the computer using your finger. You will use your finger to select a point on the scale after each stimulation. Please ensure that you press your finger directly on the line of the scale and not in the space above or below the line.

Now, I am going to give you an opportunity to practice using the scale. Please gently press your finger anywhere directly on the line of the scale.

[Check if VAS rating on screen changes from 999 to another value]

Test battery intro

Each time I test your skin, we will use 5 different tools. Each tool uses a different kind of stimulus: we can:

a) Touch you lightly with this filament (demo VFF)

b) Brush your skin clearly with a brush (demo)

c) Press a tiny, blunt-ended metal rod against your skin (demo both)

d) Or give you a single electrical stimulus but I won't do that now because I am still explaining the tolls to you.

- You can feel that the sensations evoked by the different tools can be quite distinct in nature. After I touch you with each tool, I will ask you to rate what you feel on the scale.
- Do not try to rate the different tools relative to one another. Each time we use a new tool, do not try to compare it to the previous tool. Just consider each test by itself and start again with the scale.
- Remember that a rating here on the left point, it was not painful even if you felt something. As you move closer to the right side, it means it becomes more painful. Then if you select here on the right point it means it was pain as bad as you can imagine.
- Please stay with these reference points during the whole experiment! I will ask you to close your eyes when we test your skin. Then I will ask you to open them so you can give me a rating on the screen.

Now I will give you a chance to practice giving ratings for each tool. When we start the experiment, you'll need to give your rating quite quickly, but for now we have more time. When testing I will touch you three times with each tool [besides the brush], and I want you to give me a rating for the average of the 3 stimuli.

~Press SPACE

Baseline testing

Now that we have practiced, we can begin. First, we do 3 rounds of the test battery, which usually takes about 6 minutes. After that we will give you the Intense electrical stimulation. Then you will have a 30-minute break, and lastly, we will spend the rest of the time repeating the skin tests.

Temporal summation

We will only do the next procedure once

a. Temporal summation with 256mN pinprick

[instructions for temporal summation]

Okay, so we are going to perform the next task only once in this procedure.

1. Now, I am going to touch you once with this tool [show 256mN pinprick] then I am going to ask you to please give me a rating.
2. Okay, now I will touch you 10 times in a row. Then I will ask you to please give me a rating for how the LAST stimulus felt.

[You will need these phrases: Please give me a rating for this; average of 3]

HFS

Note clock time of HFS onto participation record.

[Get ready to start stopwatch when the first HFS train starts]

Now it's time for the intense electrical stimulation. Most people find these moderately painful. The stimulation takes 1 minute in total, but it is split up into 5 trains. Each train lasts one second, and then you get a 9-second break, and the screen will ask you to rate how that train felt. So, you'll have one second of stimulation, then 9 seconds' break to give a rating, one second of stimulation, 9 seconds' break to give a rating - and so on. After this, you will have a 30-minute break.

Watch the screen during this time, because it will show you when the next train is coming and ask you to give ratings after each train. I want you to concentrate on keeping your arm glued to the table when the first train starts because some people pull their arm back as a reflex and then you could pull out cables, which can be a problem. Is it alright if I put my hand like that [indicate - hand on palm] just to remind you to keep your hand glued to the table?

As a safety precaution, I will keep my finger on the safety switch so that if you decide you want to pull out of the study you can say 'STOP' and I will immediately flick the switch down to deactivate the stimulator. I've never had someone need to pull out, but I'll be ready in case you need me to stop it. Remember: there are only 5 trains, and each one lasts one second before you get 9 seconds' break - so just count yourself through.

[Ensure switch on DS7A is turned up]

[hand on output switch, start stopwatch [see light on amplifier]., check that participant gives ratings]

[count them through the trains e.g. okay that was the first one….]

[Turn off safety switch after this procedure]

30 minutes break

Now we will take a 30-minute break. During this time, you are not allowed to leave this room. I would like to chat with you, but unfortunately, I am not allowed to because we want to keep things standardised. So here are some information pamphlets that you could look at during this break.

Surface area testing

[give instructions but screen at 30 minutes]

Follow up testing. [Turn on safety switch]

Now, I will also use the pinprick **[show]** to test for an area of higher sensitivity. I'm going to apply the pinprick in two spots, and I want you to tell me if there is a very obvious difference in the sensation (physical feeling).

Does it feel different if I touch you here **[distal]…** and here **[adjacent to electrode]**?

[if no, repeat from proximal to electrode]

[If participant is unsure, repeat test]

Ok, now I want to map out the area of higher sensitivity. So, I will apply the pinprick repeatedly along each of these radial lines, moving towards the electrode. If you feel a distinct change in sensation, please say "now". Please close your eyes. [test. Repeat this instruction at each mapping time point.]

Binding Check

[Write P or N for pain status]

[Next, complete confidence scale]

[Next, to participant: What do you think the study was about**?]**

[note participants direct quote]

After this, Turn Participant screen off, thank them for participating in the study and return to complete blinding check

Description: isiXhosa procedure scripts. Please note only instructions for the assessor are written in English.

Intshayelelo

Wamkelekile kwaye siyabulela ngokuthatha in-xa-xheba kolu-hlolo lwa-nge-zele-le-kileyo. Phambi kokuba siqalise, ndicela ucime ifowuni yakho uqinisekise ukuba akukho alarm enokuyivula ngeli xesha sisa-xa-ke-kileyo ngalo. Ndizakufunda imi-yalelo kwesi-script uku-qinise-kisa ukuba zonke iin-kqubo zenziwe kakuhle.

Yeyiphi ingalo atsale kuyo igazi uSister Mtingeni? Kulungile ke siza-ku-sebenzisa enye ingalo. Ndicela ukhulule izinto ozinxibileyo kulengalo.

Oku kuza-ku-thaba-tha ka-nga-nge-mizuzu eyi90. Okokuqala ndizakucacisela ngolu-hlolo. Okulandelayo sizakwenza uhlolo lwe-sikhumba apha kulengalo. Emva koko uzakufumana (intense electronical stimulation) Kule ngalo inye.

Sizakuphumla ka-nga-nge-mizuzu eyi30. Emva koko, size siphi-nda-phinde ukuhlola isikhumba sakho amathuba amaninzi kwimizuzu enga-mashumi amane elandelayo.

Demographic Info [ask for info for participation record sheet]

Andi-vumele-kanga ukwazi nantoni na ngolwazi lwempilo yakho. Ngoko ke, xa ndikubuza lemibuzo ilandelayo, nceda uphendule EWE /HAYI kuphela.

- Ingaba-unazo na intlungu kule ngalo namhlanje?
- Ingaba akhona na amanye amayeza owa-seli-leyo ngaphandle kwamayeza akho esi-qhelo?
- **[females only**] Ingaba ukhulelwe okanye uyacabanga ukuba ukhulelwe?
- ngaba unayo na i-implant yentsimbi okanye yom-bane, njenge-pacemaker yentliziyo?
- Ingaba unazo na iing-xaki noku-phili sa kwesikhumba?

[check for tattoos on the arm in the area between the two electrodes]

[Give info sheet and consent form]
USister Mtingeni ebe-sekufundele ulwazi no-xwe-bhuuu lwesi-vumelwano malunga nolu-phando. Usa-khululekile malunga no-lolwazi? Uvumelekile ukuba utshin-tshe ingqondo yakho na-nga-liphi na ixesha.

Ndifuna uqhelane nezi-xhobo esizaku-zisebenzisa kolu hlolo (njengale brashi). Le yi-electrode esiza-ku-yisebenzisa xa sivu-selela isikhumba. Xa uyi-qwa-la-sela uzakubona ukuba (ine-naliti) ineenaliti ezincinci ezinga-hlabiyo. Izaku-hlala kwi-sikhumba sakho kanje, kodwa kuba ibu-thuntu ayikho buhlungu kwaye ayi-zu-ku-gqobhoza isikhumba sakho. **[inne naliti ezincane ezibuthuntu]**

[When done] Unayo imibuzo?

Phambi kokuba siqale,inga-ba uyafuna uku-sebenzisa indlu yanga-sese? Nge-lishwa awuzu-kwazi ukuya kuyo xa sele siqa-lile ngohlolo. Kuza ku-funeka sigqibe ngohlolo pha-mbi kokuba uphume.

Unayo imibuzo?

Nango-na ubu-tyi-kitye isi-vume-lwano, uvu-me-lekile ukuba ur-hoxe (ugoxe) kuphando nanini na nge-lixesha lolu-hlolo ---- kufuneka nje undixelele ukuba ufuna ukuyeka. Ukuba uyar-hoxa, ungathatha isi-gqi-bo sokuba uyasi-vumela na ukuba sise-benzise olu-lwazi sendilu-fumene okanye ufuna silutsha-balalise.

Skin marking (A-H)

Ngoku ndi-za-ku-zoba amachaphaza ne-mi-gca engaweni yakho uku-ku-va-va-nya. Inga-ba kulungile okoy?
[mark up using foam template, with A towards the cubital fossa; centre of radial lines approx 8cm from cubital fossa but NOT on prominent vein or scarred area.]

[Wipe arm with alcohol balls]
Nje-ngo-ko ndi-ku-bonisile le yi(electode) esiyi-sebenzisayo. Ndiza-kuyi-beka apha kwi-sikhumba sakho ngo-luhlobo, kwaye izakuhlala apho kulo lonke olu-vava-nyo. Sizakubeka ibenye kulengalo.
[Strap electrode goes around upper arm. Disc electrode approx 8cm from cubital fossa.]

Detection threshold testing

[TURN switch of DS7A UP]

Ngoku, kolu-vavanyo lula-ndelayo, ndicela undixelel xa usiva uku-vu-selela kom-bane.
Ndizakuku-stimulator ngombane , ndi-qa-lela kumandla ase-zantsi. Ndizaku-mana ndi-wandisa amandla om-bane undi-xelele ukuba uya-weva. Nce-da uthi ‘ewe” xa uyiva , nokuba kukancinci. Izaku-vakala ngo-kungathi uhlaty-wa nge-naliti kancinci apha **[indicate electrode]**
[Step up by 1, step down by 0.5, then step up by 0.2]

[find threshold for arm, write current on monitor/check screen, set DS7A to that level and then flick switch to x10].[Leave switch on]

Scale Intro
Sisebenzisa esi-si-kali koluva-vanyo, ngoko ke ku-balu-lekile ukuba usiqonde. Isi-nceda nge-ngxe-lo (NGE-NXELO) yokuba uva ntoni kuba abantu beva ngo-kwa-hluki-leyo.
[Show VAS on the screen and paper]

Esiskali sisuka kulo mgca, obonisa ukuba **akukho zintlungu**, ukuya kule cala, **obonisa intlungu embi onokuba nayo.**

- Ukuba utyumba lo mgca unga-se-kho-hlo eku-gqibeleni, oko kuthetha ukuba uvavanyo be-lunge-bu-hlun-gwanga. Useno-kwenzeka ukuba ikhona into oyivileyo , kodwa ibinge-bu-hlun-gwanga
- Naliphi na inqa-naba eli-kulo **mgca u(flat) lithetha ukuba uvavanyo belibuhlungu**.
- Ukuba utyumba lo mgca unga-se-kunene eku-gqibeleni, uveintlungu embi onokuba nayo.
- Lonto ithetha ukuba ukuba ukhetha eku-gqi-beleni kwe-licala lasekhohlo, beku-ngebu-hlungwanga noba ikhona into oyivileyo. Nje-ngo-kuba uye uson-dela kwe-licala
  Lase-kunene, ithetha ukuba kuya kusiba buhlungu. Ze ukuba ukhetha apha eku-gqibeleni kwelicala lase-kunene ibiyeyona ntlungu embi onokuba nayo.

Ngalo lonke ixesha ndi-vava-nya isikhumba sakho ngesi-xhobo, [**nje-ngale brashi**] ndi-za-ku-kucela ukuba ureyithe istimulation ngasinye apha esikalini e-khom-puyutheni usebenzisa umnwe wakho.
Ndicela uqini-seke ukuba uci-ni-zela umnwe wakho nqo emgceni wesikali, ungaci-nezeli nga-phezulu okanye ngaph-antsi komgca.
Ngoku, ndiza kukunika ithuba **loku-ziqhe-lanisa** noku-sebe-nzisa isikali. ndiyacela ucofe umnwe wakho naphina kumgca wesikali. [**Check if nrs rating on screen changes from 999 to another value]**

Test battery intro
[clean my arm]

Qho xa ndi-vavanya isikhumba sakho, sizakusebenzisa iintlo-bo eziyi5 eza-hlu-kileyo zo-vava-nyo. Uhlobo nga-lunye lo-vavanyo lu-sebenzisa u-hlobo olwa-hlukileyo lwe(stimulus). Singa:

- Kubamba kancinci ngale(filament)
- Hli-ki-hla isikhumba sakho nge(brush)
- Cine-zela kancinci iskhumba sakho ngale naleetin ebutuntu
- Okanye siku(stimulate) kanye_kodwa andizukuyenza ngoku lonto.

- Ungeva ukuba iimva-kalelo ezi-vuselelwe zim-vavanyo eza-hlu-kene-yo (**eza-hluk-leyo**) azi-fani. Siza-kuku-cela ukuba usibonise indlela oziva ngayo kwisikali.
- Ungazami uk-ureyitha iim-vavanyo ezo-hlu-keneyo ngoku-fanayo. Qho xa sise-benzisa uvavanyo olutsha, zama ungalu-fanisi nolu-dlu-lileyo. Uvavanyo nga-lunye lubone nje-ngo-lutsha uze uqalele ngesikali.
- Khumbula ukuba apha kwelicala lase-kho-hlo eku-gqibeleni, **ibi-ngekho buhlungu nokuba ubuve into. Nje-ngo-kuba usondela kwelicala**, ithetha ukuba **kuya kusiba buhlungu**.Ukuba ukhetha apha kweli cala ,ithetha **ukuba ibi-zezona ntlungu zin-gamandla oye wanazo**

Ngoku uzakufumana ithuba lokuzama ukureyitha uvavanyo nga-lunye. Xa siqala ngo-vavanyo, kuza-kufuneka ureyithe ngoku-kha-wuleza, kodwa okwa-ngoku sisenalo ithuba ela-ne-leyo. Xa ndisenza uvavanyo ndiza-ku-kuthinta amaxesha amathathu ngesi-xhobo ngasinye, , kwaye ndifuna undinike um-li-nga-niselo ongu3. **[perform full test on arm]**

[practice complete]

Singa-qlisa ke-ngoku. Kuqala siza-kwenza imi-jikelo emi-thathu yoku-jonga olu-vava-nyo, edla ngo-ku-thatha imizuzu eyi6. Emva koko ndizakukunika istimulation sombane esinamandla, ndize nd-ikuphe imizuzu eyi30 yokuphumla, size si-chithe i-awari si-phinda-phinda uvavanyo. [**Start baseline testing including estim]**
[you will need these phrases: please give me a rating for this; average of 3]

Temporal summation

Kulungile, sizokwenza lomsebenzi ulandelayo kube kanye qha koluvavanyo.

1. Ngoku, Ndizokuncinda kanye ngesisixhobo[bonisa u 256mN we pinprick] emvakoko ndizokucela undiphe inani.
2. Kulungile, Ngoku ndizokuncinda amatyeli alishumi ngokulandelana. Emvakoko ndizokucela ukuba undiphe inani elizochaza ukuba istimulation sokugqibela usive sinjani.

[Get ready to start stopwatch when the first HFS train starts]

Baseline testing

Ngoku lixesha lestimulation sombane. Abantu abaninzi ba-ku-fumanisa ku-buhlungu oku. Istimulation sithatha umzuzu omnye xa sisonke ,kodwa sa-hlu-lwe sazi-train eziyi5. Itrain nganye iba ngu-mzuz-wana omnye, uze ufumane ikhefu lemi-zuzwana eyi9 size iscreen siku-buze ukuba ureyithe ukuba ivakele njani na lo-train. Ukuthi ke izakuba ngu-mzuzwana omnye we-stimulation, ibe yimizuzwana eyi9 yekhefu unikeze ireyithi, iphinde ibe ngumzuzwan we-stimulation, ibe-yimizuzwana eyi9 yekhefu nokunika ireyithi njalo-najlo. Emva koko uzakuphumla.

Ndicela ujonge i-screen nge-lithuba, kuba sizakukubonisa ukuba i-train elandelayo iza nini kwaye sikucele ukuba unike ireyithi emva kwe-train nganye. Ndicela ugxile (**uxile**) eku-gci-neni ingalo yakho incamathele etafileni xa kuqala uku-stimulatwa kuba abanye abantu baye bayitsale ingalo yabo ba-khuphe ii-cables, lonto ke engaba yingxaki. Ingaba kulungile na xa ndibeka isandla sam njena **[indicate]** uku-khum-buza ukuba ugcine isandla sakho sin-cama-thele etafileni?

Nje-ng-esiqinisekiso sokhu-seleko, ndiza kugcina umnwe wam ukwi-qho-sha lokhu-seleko ukwenzela ukuba uziva ufuna ungasa-funi uku-qhu-bekeka nophando ungathi YEKA kwaye ndi-za-kwu-icima kwa-ngoko umatshini. Akuzange kuke kubekho mntu orho-xayo okanye othi mandi-yeke, kodwa ndi-zakuba ready ukuba uthi mandiyeke. Khumbula ziyi5 qha ezi-train kwaye inye ithatha umzuzwana omnye phambi kokuba ufumane ikhefu lemi-zuzwana eyi9, ngoko ke ungazibalela wena ngokwakho.

[hand on output switch, start stopwatch, check that participant gives ratings]
[Ensure switch on DS7A is turned up]

HFS

Start stopwatch at first train [**see light on amplifier**].
Note clock time of HFS onto participation record.
Record VAS rating for each train. **[Turn off safety switch after this procedure]**

B R E A K

30 minutes
25 min in

Ngoku sizakuthatha ikhefu lemizuzu eyi30. Ngelithuba awuvu-melekanga ukuphuma kweli-gumbi. Ndingathanda uku-ncokola nawe kodwa nge-lishwa andi-vumele-kanga kuba sifuna ukugcina yonke into ifana. Nanga amaphepha olwazi onokube uwa-jonga ngeli xesha lekhefu.

Follow up testing

Repeat test battery (**Turn on safety switch**)

Surface area testing

[give instructions but screen at 30 minutes]

Ngoku ndizakusebenzisa i-filament ukuvavanya indawo enem-vakalelo ephe-zulu. Ndiza-kuku-hlaba kwiindawo ezimbini kwa-ye ndicela undixelele ukuba ukhona umehluko ocaci-leyo kwimvakalelo (imvakalelelo yomzimba)

Ingaba ivakala ngokwa-hlu-kileyo xa ndikubamba apha…… **[distal]…** nalapha? **[adjacent to electrode]?** [if **no**, repeat from proximal to electrode]

[If participant is unsure, repeat test]
Kulungile ke, ngoku ndifuna uku-makisha indawo enem-vakalelo ephezulu. Ndiza-kuku-hlaba ngalefilament ngoku-phinde-neyo kule migca, ndisiya kwi-electrode. Ukuba uva umehluko cacileyo kwim-vakalelo, nceds uthi “ngoku”. Ndicela uvale amehlo akho.
[test. Repeat this instruction at each mapping time point.]

Appendix K: study protocol

Title:

Study protocol: Does distress predict central sensitisation in people living with HIV?

Authors:

1. Luyanduthando Mqadi, BSc Physiology & Psychology (NWU), BMedScHons Neuroscience (Physiology), Pain team, Department of Psychiatry and Mental Health, University of Cape Town, Cape Town. Luyanduthando.mqadi@gmail.com. ORCID: 0000-0003-0542-2221
2. Gillian J Bedwell, BSc (Physiotherapy), PG Dip (Interdisciplinary Pain Management), MSc (Physiotherapy), Pain team, Department of Anaesthesia and Perioperative Medicine, University of Cape Town, Cape Town. gillbedwell@gmail.com. ORCID: 0000-0003-4522-5679
3. Peter Kamerman, BSc (Hons) (Wits), PhD (Wits), School of Physiology, University of Witswatersrand
4. John Joska, MBChB, MMed (psych), PhD, FC Psych (SA), University of Cape Town
5. Romy Parker, BSc(Phys), BSc(Med)(Hons)Ex.Sci.(Phys), MSc(Pain), PhD. Pain team, Department of Anaesthesia and Perioperative Medicine, University of Cape Town, Cape Town
6. Victoria J Madden, BSc (Physiotherapy), PhD. Pain team, Department of Anaesthesia and Perioperative Medicine, University of Cape Town, Cape Town. HIVMental Health Research Unit, Department of Psychiatry and Mental Health, University of Cape Town, Cape Town. torymadden@gmail.com. ORCID: 0000-0002-5357-4062.

Protocol version 1.1 (revised 6 April 2022, after basic tidying of data; before data inspectionor analysis)

Abbreviations:

SH	Secondary Hyperalgesia
CS	Central sensitisation
PSD	Psychosocial distress
VAS	Visual Analogue Scale
HSCL-25	Hopkins Symptom Checklist-25
PID	Participant study ID
HIV	Human Immunodeficiency Virus

INTRODUCTION

Pain and psychosocial distress (PSD) are prevalent and intersecting problems in people withHIV. People with HIV report persistent pain more frequently than their HIV-negative peers (1). Pain is commonly comorbid with PSD (2), which includes symptoms of depression and anxiety. Several studies show that people with HIV in sub-Saharan Africa frequently report PSD: the prevalence of distress symptoms ranges from 13% - 50% (3-8) for depression and 11% - 29% (9-11) for anxiety. PSD and pain typically result in similar adverse health outcomes, such as impaired functioning (12, 13) and reduced quality of life (13, 14).

Although pain is widely acknowledged to contribute to distress (15, 16), there is evidence that the opposite is also true: distress contributes to pain (17-21). It is plausible that PSD contributes to pain by increasing the efficiency with which excitatory neural signalling is transmitted across synapses within the dorsal horn – i.e. by promoting "central sensitisation"(CS)(22).

Central sensitisation is a hallmark of many persistent pain conditions (23). Although CS manifests in multiple ways, a common feature is secondary hyperalgesia – an increased sensitivity to noxious input, such that a stimulus that is usually painful (in an un-sensitised condition) becomes even more painful, including in areas adjacent to the actual site of tissueinjury (24). It is the increase in magnitude and spatial distribution of excitatory signalling in CS that is thought to support SH (25). Clinical SH can be experimentally modelled in humans using procedures that induce a robust, reproducible, short-lived SH (26, 27). This study will use this model of inducing experimental SH to test whether the severity of self- reported PSD is associated with experimentally induced SH, in people with HIV. Given that CS is a known feature of persistent clinical pain, we will also test whether the association between PSD and induced SH differs in people with HIV who report persistent pain (self- reported pain on most days for more than 3 months (28))—who would be expected to have some CS—and people with HIV who are pain-free. We will estimate SH using two outcomes:surface area (**primary outcome**), and magnitude (**secondary outcome**) of SH. We hypothesise that both *surface area* and *magnitude* of experimentally induced SH will be positively associated with self-reported PSD, and that the persistent pain group will have greater surface area and magnitude of experimentally induced SH than the group without pain.

METHOD

Study overview

We will conduct a single-blinded, case-controlled, experimental study with participants who report persistent pain as cases and those who are pain-free as controls. This study dovetailswith a larger 'parent' study that includes the assessments of pain and PSD. Two researcherswill collect data: Researcher 1 will perform assessments in the parent study (assessment of eligibility, pain status, and PSD), and Researcher 2 will perform the assessments in this current study (induction and assessment of SH) and the analysis of data. We will use the method of high frequency electrical stimulation (HFS) (27) to induce experimental SH in bothgroups. We will assess the surface area of the experimentally induced SH using an 8-radial-lines approach (29), and the magnitude of experimentally induced SH by comparing ratings of pinprick stimuli on an electronic vertical visual analogue scale (VAS) before and after the induction.

Blinding

Researcher 1's involvement will support blinding of Researcher 2 to pain status (hereafter referred to as group membership) and PSD scores. Researcher 2 will be blinded to group membership during data analysis. Researcher 1 will be unaware of the study hypotheses; Researcher 2 will be aware of the hypotheses. Participants will be blinded to the study's aims and hypotheses as well as their classification into different groups. We will assess forunblinding for Researcher 2 and participants at the end of the procedure (see *Phase 7: Blinding assessment* below).

Ethical approval

Ethical clearance has been granted by the Faculty of Health Sciences Human Research Ethics Committee of the University of Cape Town, under the parent study protocol number 764/2019 and sub-study 277/2021. This current study is to be used in a dissertation towardsa Masters by research degree. The principles of the Declaration of Helsinki will be adhered to throughout the study (30). This study will enrol only eligible, informed, and consenting participants, and participants will remain free to withdraw at any point during the procedure, without negative consequences.

Participants and recruitment

Recruitment of participants to this study will be an additional step after recruitment to the parent study. The parent study will match participants across groups for sex and age (within5 years): matching limitations are unnecessary for the current study. The parent study will

include participants if they are able to give fully informed consent and engage in the test battery. Participants who consent to participate in the parent study will be invited to participate in the current study, subject to the eligibility criteria in **Table 1**. Participants will becompensated for involvement in the parent study and will also receive an additional ZAR 150[1] for their involvement in the current study, given the extra time requirement and discomfort of the procedure.

Table 1: Parent study and current study eligibility criteria.

Inclusion
Between the ages of 18 and 65 years.[a]
Fluent in spoken English or isiXhosa.[a]
HIV positive with evidence of viral suppression (test showing viral load < 50 copies/ml within preceding three months).[a]
With persistent pain (pain on most days for more than 3 months (28)) or no pain and thosewho report persistent pain will be included based on pain frequency and duration, with no regard to the severity, location, or presumed cause of their pain.[a]
Exclusion
Reported being advised by a medical practitioner to avoid stressful situations.[b]
Self-reported neurological or cardiovascular problems, or pregnancy.[a]
Acute psychosis [c], high risk of suicidality, substance abuse disorder, or obvious cognitivedisorder (identified by screening).[a]
Has reported the following on arm to be tested [a]: • current pain • problems with sensation • electrical or metal implants • tattoos (between electrode locations)

[a] Additional for the current study.

[b] Researcher 1 clarified the relevance of the advice to the induction procedure and used clinical judgement to determine safety.

[c] To exclude individuals in need of acute mental health care or unfit to consent.

[1] Approximately US$9.21 – US$11.18 between February and November 2021

Sample size

Previous studies using experimentally induced SH in healthy controls without between-group comparisons have used samples of 7-20 (27, 31); one study with a between-group comparison used n = 15 per group (32). Given that the current study uses participants with known health concerns and that we had no data to inform an estimated effect size, we aimedfor a somewhat larger but still pragmatic sample. The parent study will recruit and assess 100 individuals, and we anticipate that approximately 60 of those will be eligible and willing to also participate in the current study. Therefore, we aim to obtain full datasets for 60 participants (n = 30 participants per group).

Stimuli

We will deliver HFS via specialised electrodes (Figure 1) to induce SH on the skin. The cathode consists of 10 blunt steel pins arranged in a circle. Using a double adhesive sticker,Researcher 2 will attach the cathode to the anterior surface of the participant's forearm, approximately 8 cm distal to the cubital fossa, avoiding any prominent vasculature. The anode consists of a conductive fabric band strapped around the participant's upper mid-arm.To induce SH, we will administer five trains of 100 Hz stimulation at a current of 10 times theparticipant detection threshold (33). Each train lasts for 1 second, followed by a 9-second break between each train. The sensory effect of the HFS is centred on the skin area below the cathode. A software programme called Affect5 (updated from Affect4 (34)) will be used to trigger electrical pulses from a constant current electrical stimulator (DS7A; Digitimer Limited, Hertfordshire, UK) set to maximum voltage (400V), 2000μs pulse width, a square pulse shape and negative polarity.

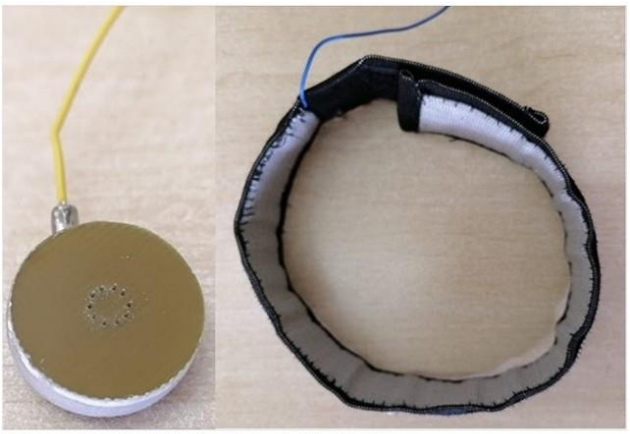


Figure 1: *Cathode (left) and anode (right).*

Arm receiving HFS

During the parent study, blood samples will be collected from each participant. Given that venepuncture induces sensitivity in the area where blood is drawn, the opposite arm will be

used in the current study procedure. In certain cases, the participant may have contraindications to electrical stimulation (e.g. metal implant) of the chosen arm, or difficulties with blood draw may require venepuncture of both arms. Therefore, an arm usedfor venepuncture may be used for the current study if there is a 7-day 'washout period' between the blood draw and the current procedure.

Grouping variable: clinical persistent pain

Researcher 1 will assess participants pain status using two questions adapted from the BriefPain Inventory (BPI) (35). The first question (*"*Throughout our lives, most of us have had painfrom time to time (such as minor headaches, sprains, and toothaches). Have you had pain other than these everyday kinds of pain *during the last week?"*) asks participants to disclose pain status during the last week. A positive response to the first question will be followed by a second question asking participants to disclose pain status on *most days for the past 3 months*. Both questions will be used to classify participants into group membership based onpain status. The BPI is a validated and reliable self-report measure that has been used to study several populations with HIV and persistent pain in South Africa and has been translated and validated in isiXhosa (35).

OUTCOME MEASURE

Electronic Visual Analogue Scale (VAS)

Given the anticipated range of education level amongst participants, we will use a touch- based electronic vertical VAS to obtain participants' ratings during sensory testing. This scale is anchored at the bottom with "no pain" and at the top with "worst pain you can imagine". The anchors will be translated into isiXhosa for Xhosa-speaking participants. Thescale is user-friendly: the participant will swipe a stylus pen across the screen to give a rating. In response, a central bar on the scale fills up with red to the point where the participants swiped, providing visual feedback on the rating. The rating will be electronicallyrecorded, and the computer will transform the analogue feedback into a numerical rating between 0 (bottom anchor) and 100 (upper anchor). Visual analogue scales are validated and reliable and require only brief training, and are well suited to determining pain intensity in diverse adult populations (36).

Exploratory predictor variable

Psychosocial distress

The 25-item Hopkins Symptom Checklist (HSCL-25) is a validated self-report measure of symptoms of depression and anxiety (37) that has been used in South Africa (38), includingin people with HIV (39). To our knowledge, the HSCL-25 has not been formally validated in

isiXhosa. We translated it to optimise local relevance of the language with a formal forward- and-back-translation process (40). The HSCL-25 consists of 25 items, of which 10 focus on anxiety and 15 focus on depressive symptoms (37). Participants indicate the extent to whicheach symptom (e.g., 'crying easily', 'poor appetite', 'feeling lonely') applies to them within thepast month on a 4-point scale (1 = 'not at all' to 4 = 'extremely'). This study will use the total score, which provides an ordinal measure of PSD.

OUTCOMES

Primary outcome: surface area of secondary hyperalgesia

A von Frey filament (VFF: Marstock nervtest, Germany) exerting a force of 128 mN will be used to estimate the surface area of experimentally induced SH on eight radial lines (29). This von Freyfilament is stiff and typically feels sharp when applied to the volar surface of the forearm. Before the induction of SH, Researcher 2 will use a template to mark eight radial lines on the anterior surfaceof the participant's forearm. The radial lines originate from the stimulation site and are arranged at 45 degrees to each other (Figure 2). After the induction of SH, Researcher 2 will perform a two-step process to determine the surface area of SH. First, she will screen for an area of higher sensitivity by asking, "Does it feel different if I touch you here [1st stimulus, distal to the cubital fossa (see area marked 'A' on Figure 2)] and here [2nd stimulus, adjacentto electrode (see area marked 'A0' on Figure 2)]?" If the participantis uncertain, the screening will be repeated on the same radial line.If the participant reports no distinct difference in sensation betweenthe two points on radial line 'A' Researcher 2 will screen from 'E to E0'. If there is no distinct difference on radial line 'E', the surface area will be assumed to be 0. This screening process was included after pilot testing showed that participants sometimes report no SH on VAS ratings, but may also sometimes report a vague transition in sensation on surface area mapping, despite some uncertainty about whether the sensation actually differs. This can result in the mapping out of a supposed surface area of SH that does not actually exist. If the participant reports a distinct difference in sensation on one of the screening tests ('A vs A0' or 'E vs E0'), the surface area of SH will be mapped outby applying the von Frey filament along each radial line, moving towards the electrode (Figure 3).

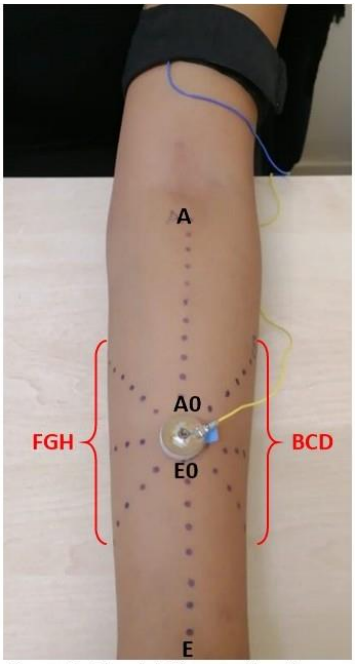


***Figure 2**: The eight lines on the skin indicate radial lines that are 45° apart, and the black dots forming the radial lines are 1cm apart. A to A0 and E to E0 indicates the points for SH screening. No screening is done over radial lines BCD and FGH.*

We recognise the limitations that such a method presents, specifically the risk of missing anarea of SH on the lateral (see 'BCD' in Figure 2) and medial ('FGH' in Figure 2) sides of the forearm where we will not screen for SH. Procedural requirements prevented immediate estimation of SH magnitude and we wished to prioritise clarity, therefore we chose to err on the side of under-estimating the surface area of SH rather than running the risk of over- estimating the surface area, and we considered this when selecting the analytical approach.

During this procedure, the participant will be asked to keep their eyes shut and verbally indicate when they feel a distinct difference in sensation to a stimulus, in comparison to the preceding stimulus (that will have been delivered at a point that is further from the electrode).This point of transition will be recorded as the boundary of the SH for each radial line. To calculate the total estimated surface area of SH, we will sum the surface areas of the 8 triangles, with *surface area = 1/2ab (sin 45°)*, where *a* and *b* are the lengths of the sides of a triangle adjacent to the 45-degree angle.

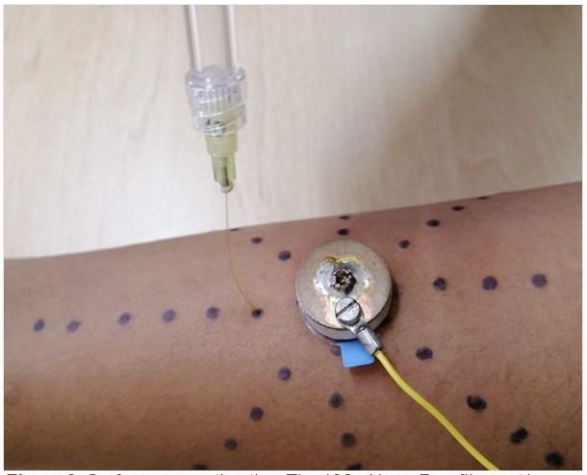


Figure 3: *Surface area estimation. The 128mN von Frey filament is applied starting at the most distal point and moving proximal to the electrode.*

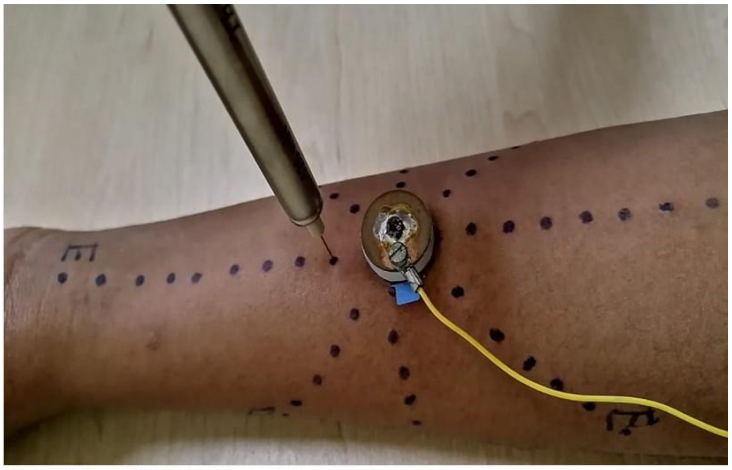


Figure 4: *Assessing the magnitude of SH.*

Secondary outcome

Magnitude of SH

We will use two weighted, blunt-ended metal rods called pinprick stimulators (PinPrick, MRCsystems, Heidelberg, Germany) to assess the magnitude of SH (Figure 4). The pinprick stimulators exert either 128mN or 256mN when correctly applied to the volar surface of the participant's forearm, parallel to gravity. We will stimulate the skin with the chosen force for
~1 second, provide two more repetitions, and then ask for a VAS rating for an average of thethree stimulations. These ratings will be obtained before and after the induction of SH at phases 3 and 6b (see procedure below). The mean of the ratings for the two weights at eachtime point will be taken forward to analysis. For each time point after induction, this mean will

be expressed as a percentage of the equivalent mean as obtained before induction. Thus, the procedure will yield three estimates of SH magnitude for each participant, consisting of induced sensitivity expressed as a percentage of baseline sensitivity to pinprick stimulation.

Exploratory outcomes

Knowing that HFS is typically painful, we want to know how painful the induction is in this sample. Participants will rate each train of HFS on the electronic VAS, yielding 5 HFS ratings(1 per train) from each participant.

Four additional outcomes will be assessed, although none of these data will be used in the current study. At each assessment time point, participants will provide VAS ratings of stimulation with (1) a 32mN von Frey filament (VFF: Marstock nervtest, Germany), (2) a lightbrush (Prime Art Bianco Flat 16), and (3) a single electrical stimulus with the same settings as used for the HFS. Participants will also provide (4) VAS ratings of (a) a single 256mN stimulation and (b) the 10th of ten 256mN pinprick stimuli delivered at the same location.

PROCEDURE

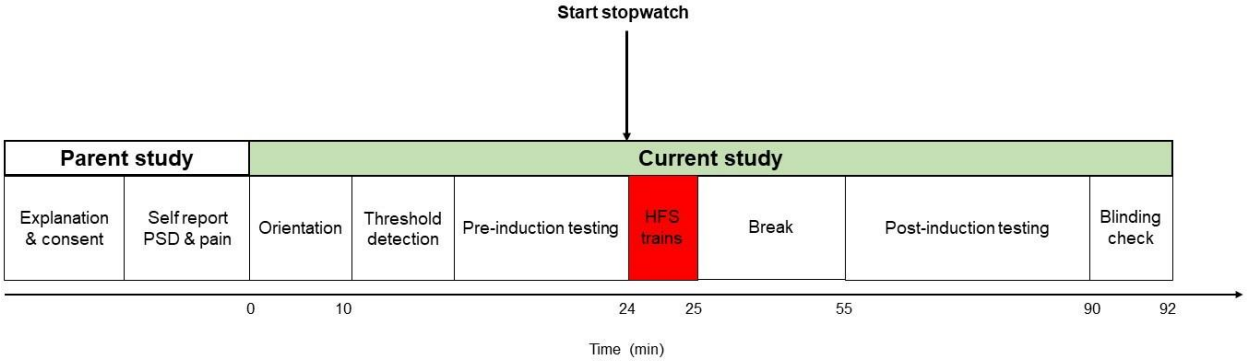


Figure 5: *Study procedures (not drawn to scale). The white and light green boxes at the top of the figure delineate the timing of the procedures by which occur in a separate session (parent study) and the current session (current study). Red shading represents the SH induction, given to one forearm that is free of clinical pain. The current study assessments were separated from parent study assessments by either a short break, or up to 7 days.*

Figure 5 shows the study procedures, which include 9 phases: explanation and consent, PSD and pain assessment (completed in the parent study), orientation to experimental stimuli, threshold detection, pre-induction sensory tests, HFS trains, post-induction sensorytests, and blinding assessment.

Relevant procedures of the parent study

Researcher 1 will screen for exclusion criteria and complete the informed consent process for this study as part of the parent study. Participants will choose if they wish to communicate in English or isiXhosa. During the informed consent process, Researcher 1 will

use the explain-and-explain-back method to confirm understanding, and to detect cognitive difficulties that might prevent the participant from understanding the study information and procedure. If participants show continuous difficulties understanding the study information, they will be excluded from participating in this study. Once the participant can confirm theirunderstanding of the study information and consents to participate, they will sign the informed consent form, receive venepuncture and complete a battery of self-report assessments including the HSCL-25 and BPI. The induction and testing of SH will then be conducted in a separate session, in a separate room, with Researcher 2.

Procedures of the current study

Phase 1: Orientation phase:

The participant will be asked to sit on a chair across a table from Researcher 2. The researcher will follow a script to ensure consistency in all experimental procedures. Participants will be asked to sanitise their hands, and the anterior surface of the participant'sforearm and upper mid-arm will be cleaned with alcohol. Next, Researcher 2 will briefly introduce the sensory modalities (128mN, 256mN, single electrical stimulus, brush and VFF)and participants will receive a chance to ask questions. The participant will confirm consent to participate. Skin markup will be completed as per Figure 2. The electrical stimulus currentwill be calibrated for each participant (*see Threshold detection*). The electronic, touch-VAS will be explained to the participant. After this, Researcher 2 will demonstrate sensory testing,first on her forearm, then on the participant's forearm, to show the different stimuli evoked bythe different sensory modalities. Participants will receive an opportunity to practice giving touch-VAS ratings for each sensory modality. All practice ratings will be recorded, but the data will not contribute to the analysis.

Phase 2: Threshold detection:

The electrical stimulus will be calibrated to the individual detection threshold. For this, repeated electrical stimuli are delivered with gradual increases in current from 0 mA, by steps of 0.10 mA, until the participant first reports feeling the electrical stimulus. Next, the current is decreased by 0.5 mA steps until the participant reports no longer feeling the electrical stimulus. Finally, the current is increased by 0.2 mA until the participant detects it again. This final current will be used as the detection threshold. We will multiply this current by 10 to yield the stimulation intensity used in the experimental induction of SH and all singleelectrical stimuli during sensory testing (27).

Phase 3: Pre-induction sensory tests:

The participant will receive three rounds of sensory testing with modalities used in the following order: 128mN pinprick, 256mN pinprick, single electrical stimulus, brush and VFF.

We have chosen to consistently present the pinprick stimulations first because the ratings of these modalities reflect the magnitude of SH and are therefore of greater interest than the ratings of the subsequent stimulus modalities. For three of the five sensory modalities (128mN pinprick, 256mN pinprick, and 32mN VFF), one rating will be recorded for an average of 3 stimulations, with each stimulation lasting ~1 second. For the remaining two sensory modalities (electrical stimulus and brush), one rating will be recorded for one stimulation; the electrical stimulation lasts for 2 ms, and the brush stimulation lasts for ~1 second. The whole test battery (5 modalities) is repeated for three rounds during the pre- induction phase, and the ratings are averaged across the three rounds. Finally, a 256mN pinprick will be used to obtain ratings of a single stimulus and the 10^{th} of ten stimuli deliveredat the same location. Every rating will be recorded on the electronic VAS.

Phase 4: HFS induction:

The participant will receive five one-second trains of 100 Hz electrical stimulation (HFS) separated by nine seconds (offset to onset). Upon the onset of the first train, Researcher 2 will start a stopwatch to support accurate timing of follow-up assessments. The participant will be asked to rate each stimulation train using the touch-VAS. As a safety precaution, the participant will be given a bail-out word, “stop” or “yeka” (isiXhosa), to indicate immediate withdrawal if they wish to stop the trains. The researcher will keep their finger on the activation switch of the electrical simulator so that, if the participant uses the bail-out word, the stimulator will immediately be deactivated. All mid-procedure withdrawals will be reportedwith reasons as given by participants. Participants will provide a VAS rating after each train. The HFS phase will yield one VAS rating for each of the 5 HFS trains from each participant.

Phase 5: Break:

There will be a 30-minute break after the HFS to allow SH to develop (41, 42).

Phase 6: Post-induction sensory tests:

Surface area of SH

Surface area testing will be performed at 30, 45 and 60 minutes after HFS. Thus, we willrecord three estimations of SH surface area.

Magnitude of SH

We will conduct sensory testing at 35, 50, and 65 minutes after the onset of the HFS trains.Thus, we will obtain three post-induction ratings for each sensory modality, for each participant.

Phase 7: Blinding assessment

The study is designed to maintain blinding of Researcher 2 to the group membership and PSD results of each participant, and participants to the study's aims and hypotheses, including the existence of a differential group membership. In addition, Researcher 2 will beblinded to group membership during analyses. In this phase, we will assess for unblinding.

Blinding check for Researcher 2

After each procedure and before conducting the participant's blinding assessment, Researcher 2 will answer two questions on the study computer. First, she will guess eachparticipant's group, i.e., the 'pain' group or 'no pain' group. Second, she will rate her confidence in her guess about group membership on a five-point Likert scale of "not at allconfident", "not confident", "neutral", "confident", and "extremely confident".

Blinding check for participants:

To assess unblinding of participants, Researcher 2 will ask the participant what they think the study is about. We wish to maximise sensitivity of this blinding check. Therefore, if the participant is uncertain, to avoid a situation where we fail to detect unblinding (for example, ifthe participant has a suspicion but cannot clearly form the language to convey it), they will be asked to guess. The researcher will follow up on any indication that blinding may have been broken and use conservative criteria to judge whether the participant remained blindedto study purpose at this time point (i.e. if uncertain, Researcher 2 will judge blinding as broken, rather than intact).

DATA ANALYSIS

Statistical analyses will be conducted using the most recent versions of R (The R Foundationfor Statistical Computing) and R Studio (Integrated Development Environment for R), based on the pilot study analysis script provided and locked with this protocol. The pilot script uses the packages tidyverse (43), readxl (44), gridExtra (45), here (46), kableExtra (47), ggstatsplot (48), pracma (49), BI (50), dplyr (51), readr (52), arsenal (53).

Demographic data will be handled descriptively and presented in a table. All data on SH magnitude and surface area will be entered into and recorded by Affect5, i.e. directly into soft copy. Once the study is complete, all scripts used for data analysis will be available as supplementary information. In general, data will be visualised to assess distribution assumptions, and hypothesised relationships, before formal analysis. We will conduct and interpret all analyses with the analyst blinded to group membership. At the final stage of theanalyses, the analyst will be unblinded in order to assess researcher blinding. Here, we present our planned methods, assuming data that satisfy the assumptions of our planned tests. However, the final choice of test will depend on the actual data. A p-value of less than 0.05 will be used as the threshold for statistical significance.

Confounding variables

We anticipate possible confounding by the following variables: (i) calibrated current (although informal communication with other laboratories that use HFS suggests the

magnitude of electrical current used for the induction of SH does not predict the outcomes), (ii) painfulness of the HFS induction, and (iii) delay between PSD assessment and the current study procedures. Regarding (iii), some participants underwent SH induction ("current study procedures" above) on a different day to the parent study assessment of self-reported PSD; therefore, we will assess for an influence of days between PSD assessment and SH induction on the study outcomes. In each case, we will test for a correlation between the potential confounding variable and our study outcomes, and include the potential confounding variable as a covariate in the analyses if a significant correlation is found.

Primary analysis:

We will have three estimates of the surface area of SH for each participant at 30, 45 and 60 minutes after induction. Considering that we are interested in the surface area of SH across the whole period of assessment, we will plot surface area of SH over time (assuming a valueof 0 at the pre-induction time point) and calculate the area under the line for each participantacross the three time points using the trapezoidal integration rule. We will use the 'area_under_curve' function from the bayestestR package(54) in R to compute the total surface area of SH found within each participant. To determine whether the surface area of experimentally induced SH is predicted by the severity of self-reported PSD and pain status (i.e. group membership), we anticipate the following model structure being useful: Surface area ~ β_0 + β_1(PSD) + β_2(group) + β_3(PSD x group) + error.

Secondary analysis:

We will have three estimates of the magnitude of SH for each participant at 30, 45, and 60 minutes after induction. We will calculate the mean ratings for the two weights: this will result in one estimate of SH magnitude for each of the three post-induction time points. For each time point after induction, the mean ratings will be expressed as a percentage of the mean rating for the two weights as obtained before induction. Thus, each participant will provide three estimates of SH magnitude, expressed as a percentage of baseline sensitivityto pinprick stimulation. Similar to our approach for surface area, we will plot SH magnitude over time and compute the area under the line for a single estimate of SH magnitude for each participant. To determine whether the magnitude of experimentally induced SH is predicted by the severity of self-reported PSD and group, we anticipate the following model structure being useful: Magnitude ~ β_0 + β_1(PSD) + β_2(group) + β_3(PSD x group) + error.

Exploratory analysis:

We intend to manipulate skin sensitivity using HFS to induce SH. HFS is typically painful; wewant to know how painful the induction is. We will report the 5 ratings of the HFS trains (1 per train) from each participant to describe the painfulness and explore the relationship

between participant-reported painfulness of HFS and the (a) surface area and (b) magnitudeof SH. We will also explore for a difference in painfulness of HFS between groups.

Blinding check:

We will calculate and report the percentage of participants who correctly guessed the aim ofthe study and the number of participants for whom Researcher 2 accurately guessed group membership. There is a 50% chance that Researcher 2 can accurately guess the participant's group membership; however, this does not necessarily demonstrate blinding failure (55). Therefore, we will use three statistical methods, James' blinding index (BI), Chi-square goodness of fit test, and perform sensitivity analyses (where necessary), to determine the blinding of Researcher 2 to group membership. The James' BI shows a variation of the kappa coefficient, which is sensitive to the degree of 'disagreement' and not 'agreement', and it does so by prioritising 'do not know (DK)' responses (56). Given that Researcher 2 gives her best guess to participant's group membership, we will exclude 'DK' options from the blinding check and DK responses will be equivalent to '0' in the analysis.

We will also use a Chi-square goodness of fit test as an additional assessment of blinding ofResearcher 2 to group. Finally, we will identify the cases for which guesses were accurate and for which the researcher reported a confidence of 4 (i.e. confident) and 5 (i.e. extremelyconfident). This process will highlight individual cases for which blinding may have been broken. We will conduct sensitivity analyses to examine the influence of cases in which blinding was possibly broken on the findings of the primary and secondary analyses.

REFERENCES

1. Madden VJ, Parker R, Goodin BR. Chronic pain in people with HIV: a commoncomorbidity and threat to quality of life. Pain Management. 2020;10(4):253-60.
2. Duric V, Clayton S, Leong ML, Yuan L-L. Comorbidity Factors and Brain MechanismsLinking Chronic Stress and Systemic Illness. Neural Plast. 2016;2016:5460732-.
3. Lawler K, Mosepele M, Seloilwe E, Ratcliffe S, Steele K, Nthobatsang R, et al. Depression among HIV-positive individuals in Botswana: a behavioral surveillance. AIDSand Behavior. 2011;15(1):204-8.
4. Kaharuza FM, Bunnell R, Moss S, Purcell DW, Bikaako-Kajura W, Wamai N, et al.Depression and CD4 cell count among persons with HIV infection in Uganda. AIDS and Behavior. 2006;10(1):105-11.
5. Kagee A, Martin L. Symptoms of depression and anxiety among a sample of SouthAfrican patients living with HIV. AIDS care. 2010;22(2):159-65.
6. Farley J, Miller E, Zamani A, Tepper V, Morris C, Oyegunle M, et al. Screening for hazardous alcohol use and depressive symptomatology among HIV-infected patients in Nigeria: prevalence, predictors, and association with adherence. Journal of the InternationalAssociation of Physicians in AIDS Care. 2010;9(4):218-26.
7. Monahan PO, Shacham E, Reece M, Kroenke K, Ong'Or WO, Omollo O, et al. Validity/reliability of PHQ-9 and PHQ-2 depression scales among adults living with HIV/AIDSin western Kenya. Journal of general internal medicine. 2009;24(2):189-97.
8. Myer L, Smit J, Roux LL, Parker S, Stein DJ, Seedat S. Common mental disorders among HIV-infected individuals in South Africa: prevalence, predictors, and validation of briefpsychiatric rating scales. AIDS patient care and STDs. 2008;22(2):147-58.

9. Nel A, Kagee A. The relationship between depression, anxiety and medication adherence among patients receiving antiretroviral treatment in South Africa. AIDS care.2013;25(8):948-55.
10. Yunusa MA, Njoku CH, Obembe A. Self-reported adherence to treatment: a study ofsocioeconomic factors and psychiatric morbidity among male and female patients with HIV infection in Sokoto, Nigeria. Nigerian Journal of Medicine. 2014;23(1):33-9.
11. Pefura-Yone EW, Soh E, Kengne AP, Balkissou AD, Kuaban C. Non-adherence toantiretroviral therapy in Yaounde: prevalence, determinants and the concordance of two screening criteria. Journal of infection and public health. 2013;6(4):307-15.
12. Sabin CA, Harding R, Bagkeris E, Nkhoma K, Post FA, Sachikonye M, et al. Pain inpeople living with HIV and its association with healthcare resource use, well being and functional status. AIDS. 2018;32(18).
13. Deshmukh NN, Borkar AM, Deshmukh JS. Depression and its associated factorsamong people living with HIV/AIDS: Can it affect their quality of life? J Family Med Prim Care. 2017;6(3):549-53.
14. Marcus KS, Kerns RD, Rosenfeld B, Breitbart W. HIV/AIDS-related Pain as a ChronicPain Condition: Implications of a Biopsychosocial Model for Comprehensive Assessment and Effective Management. Pain Medicine. 2000;1(3):260-73.
15. Edwards RR, Dworkin RH, Sullivan MD, Turk DC, Wasan AD. The Role of Psychosocial Processes in the Development and Maintenance of Chronic Pain. J Pain.2016;17(9 Suppl):T70-92.
16. Turk DC, Fillingim RB, Ohrbach R, Patel KV. Assessment of psychosocial andfunctional impact of chronic pain. J Pain. 2016;17(9):T21-T49.
17. Rosenfeld B, Breitbart W, McDonald MV, Passik SD, Thaler H, Portenoy RK. Pain inambulatory AIDS patients. II: Impact of pain on psychological functioning and quality of life. Pain. 1996;68(2-3):323-8.
18. Rotheram-Borus MJ. Variations in perceived pain associated with emotional distressand social identity in AIDS. AIDS Patient Care STDS. 2000;14(12):659-65.
19. Lagana L, Chen X-H, Koopman C, Classen C, Kimerling R, Spiegal D. Depressive symptomatology in relation to emotional control and chronic pain in persons who are HIV positive. Rehabilitation Psychology. 2002;47(4):402.
20. Richardson JL, Heikes B, Karim R, Weber K, Anastos K, Young M. Experience of pain among women with advanced HIV disease. AIDS Patient Care STDS. 2009;23(7):503-11.
21. Miaskowski C, Penko JM, Guzman D, Mattson JE, Bangsberg DR, Kushel MB. Occurrence and characteristics of chronic pain in a community-based cohort of indigentadults living with HIV infection. J Pain. 2011;12(9):1004-16.
22. Woolf CJ. Evidence for a central component of post-injury pain hypersensitivity.Nature. 1983;306(5944):686-8.
23. Staud R. Biology and therapy of fibromyalgia: pain in fibromyalgia syndrome. ArthritisRes Ther. 2006;8(3):208.
24. Moseley GL, Butler DS. Explain pain supercharged: NOI; 2017.
25. Cayrol T, Pitance L, Roussel N, Mouraux A, Van den Broeke E. No evidence ofwidespread mechanical pressure hyperalgesia after experimentally induced central sensitization through skin nociceptors. Pain Rep. 2018;Latest Articles:1.
26. Klein T, Magerl W, Rolke R, Treede R-D. Human surrogate models of neuropathicpain. Pain. 2005;115(3).
27. Klein T, Magerl W, Hopf H-C, Sandkühler J, Treede R-D. Perceptual correlates ofnociceptive long-term potentiation and long-term depression in humans. J Neurosci. 2004;24(4):964-71.
28. Blyth FM, March LM, Brnabic AJM, Jorm LR, Williamson M, Cousins MJ. Chronicpain in Australia: a prevalence study. PAIN. 2001;89(2).
29. Madden VJ, Bedwell GJ, Chikezie PC, Rice ASC, Kamerman PR. A systematic review of experimental methods to manipulate secondary hyperalgesia in humans: protocol.Systematic Reviews. 2019;8(1):208.

30. WorldMedical A. Handbook of WMA Policies 2019 [Available from: https://www.wma.net/policy/hb-e-version-2019-2/.
31. van den Broeke E, Vanmaele T, Mouraux A, Stouffs A, Biurrun-Manresa J, TortaDJRSOS. Perceptual correlates of homosynaptic long-term potentiation in human nociceptive pathways: a replication study. 2020;8(1):200830.
32. Van Den Broeke E, Geene N, Van Rijn C, Wilder-Smith O, Oosterman J. Negative expectations facilitate mechanical hyperalgesia after high-frequency electrical stimulation ofhuman skin. European journal of pain. 2014;18(1):86-91.
33. Klein T, Stahn S, Magerl W, Treede RD. The role of heterosynaptic facilitation inlong-term potentiation (LTP) of human pain sensation. Pain. 2008;139(3):507-19.
34. Spruyt A, Clarysse J, Vansteenwegen D, Baeyens F, Hermans D. Affect 4.0: a freesoftware package for implementing psychological and psychophysiological experiments. Experimental psychology. 2010;57(1):36-45.
35. Parker R, Jelsma J, Stein DJ. Pain in amaXhosa women living with HIV/AIDS: across-sectional study of ambulant outpatients. BMC Women's Health. 2017;17(1):31.
36. Melzack R, Katz JJTop. Pain assessment in adult patients. 2006;5:291-304.
37. Ventevogel P, De Vries G, Scholte WF, Shinwari NR, Faiz H, Nassery R, et al. Properties of the Hopkins Symptom Checklist-25 (HSCL-25) and the Self-Reporting Questionnaire (SRQ-20) as screening instruments used in primary care in Afghanistan.Social psychiatry and psychiatric epidemiology. 2007;42(4):328-35.
38. Kagee AJE, Health. Symptoms of distress and posttraumatic stress among SouthAfrican former political detainees. 2005;10(2):169-79.
39. Kagee A, Saal W, Bantjes J. Distress, depression and anxiety among personsseeking HIV testing. AIDS care. 2017;29(3):280-4.
40. Brislin RW. Back-Translation for Cross-Cultural Research. Journal of Cross-CulturalPsychology. 1970;1(3):185-216.
41. Biurrun Manresa J, Kæseler Andersen O, Mouraux A, van den Broeke EN. High frequency electrical stimulation induces a long-lasting enhancement of event-related potentials but does not change the perception elicited by intra-epidermal electrical stimulidelivered to the area of increased mechanical pinprick sensitivity. PLoS One. 2018;13(9):e0203365.
42. Cayrol T, Pitance L, Roussel N, Mouraux A, van den Broeke EN. No evidence ofwidespread mechanical pressure hyperalgesia after experimentally induced central sensitization through skin nociceptors. Pain Rep. 2018;3(6):e691-e.
43. Wickham H. tidyverse: Easily Install and Load the 'Tidyverse' [Available from:https://cran.r-project.org/web/packages/tidyverse/index.html.
44. Wickham HB, J. . readxl part of the tidyverse1.3.1.9000 [Available from: https://readxl.tidyverse.org/.
45. Auguie BA, A. gridExtra: Miscellaneous Functions for "Grid" Graphics 2017 [Availablefrom: https://cran.r-project.org/web/packages/gridExtra/index.html.
46. Müller K, Bryan J. here: A Simpler Way to Find Your Files 2020 [Available from: https://cran.r-project.org/web/packages/here/index.html.
47. Zhu H. kableExtra: Construct Complex Table with 'kable' and Pipe Syntax 2021[Available from: https://cran.r-project.org/web/packages/kableExtra/index.html.
48. Patil I. ggstatsplot: 'ggplot2' Based Plots with Statistical Details 2021 [Available from:https://cran.r-project.org/web/packages/ggstatsplot/index.html.
49. Borchers HW. pracma: Practical Numerical Math Functions 2021 [Available from:https://cran.r-project.org/web/packages/pracma/index.html.
50. Schwartz M. BI: Blinding Assessment Indexes for Randomized, Controlled, ClinicalTrials 2021 [Available from: https://cran.r-project.org/web/packages/BI/index.html.
51. Wickham H. dplyr: A Grammar of Data Manipulation 2022 [Available from: https://dplyr.tidyverse.org, https://github.com/tidyverse/dplyr.
52. Wickham H. readr: Read Rectangular Text Data 2022 [Available from: https://readr.tidyverse.org, https://github.com/tidyverse/readr.

53. Heinzen E. arsenal: An Arsenal of 'R' Functions for Large-Scale Statistical Summaries 2021 [Available from: https://github.com/mayoverse/arsenal, https://cran.r-project.org/package=arsenal, https://mayoverse.github.io/arsenal/.
54. Signorell A. Area Under the Curve [Available from: https://search.r-project.org/CRAN/refmans/DescTools/html/AUC.html.
55. Sharpe L, Ryan B, Allard S, Sensky T. Testing for the integrity of blinding in clinical trials: how valid are forced choice paradigms? Psychother Psychosom. 2003;72(3):128-31.
56. Bang H, Ni L, Davis CE. Assessment of blinding in clinical trials. Control Clin Trials. 2004;25(2):143-56.

Appendix L: pilot and formal analysis scripts

[Scroll down for analyses scripts]

Pilot data analysis

Luyanduthando Mqadi, Tory Madden, Peter Kamerman, and Gill Bedwell

13 February, 2023

Contents

Research question

Does distress predict central sensitisation in people living with HIV?

Background

This is a pilot script for later adaptation to test whether the severity of self-reported psychosocial distress (PSD) is positively associated with experimentally induced secondary hyperalgesia (SH), in people living with HIV who report persistent pain or are pain-free. We plan to adapt this script once we are doing our full data analysis, but this pilot version locks in our intentions.

Research objectives

In a sample of people living with HIV who report persistent pain or being pain-free, we will:

- Assess PSD;
- Induce SH;
- Assess whether the surface area of experimentally induced SH is predicted by the severity of self-reported PSD and group ***(primary analysis)***;
- Assess whether the magnitude of experimentally induced SH is predicted by the severity of self-reported PSD and group ***(secondary analysis)***.

Hypotheses

We hypothesise that both surface area and magnitude of experimentally induced SH will be predicted by self-reported PSD and that the persistent pain group will have greater surface area and magnitude of SH than the group without pain.

```
# knitr setup
knitr::opts_chunk$set(tidy.opts = list(tidy = TRUE, warning = FALSE,
  message = FALSE, fig.align = "center", fig.path = "formal_analysis/figures",
  fig.retina = 1))

# Load packages
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --

## v ggplot2 3.3.6     v purrr   0.3.5
## v tibble  3.1.8     v stringr 1.4.1
## v tidyr   1.2.1     v forcats 0.5.2
## v readr   2.1.3
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(readxl)
library(gridExtra)
```

```
##
## Attaching package: 'gridExtra'
##
## The following object is masked from 'package:dplyr':
##
##     combine
```

```
library(readr)
library(here)
```

```
## here() starts at C:/Users/user/Desktop/training_repo
```

```
library(kableExtra)
```

```
##
## Attaching package: 'kableExtra'
##
## The following object is masked from 'package:dplyr':
##
##     group_rows
```

```
library(ggstatsplot)
```

```
## You can cite this package as:
##      Patil, I. (2021). Visualizations with statistical details: The 'ggstatsplot' approach.
##      Journal of Open Source Software, 6(61), 3167, doi:10.21105/joss.03167
```

```
library(pracma)
```

```
##
## Attaching package: 'pracma'
##
## The following object is masked from 'package:purrr':
##
##     cross
```

```
library(bayestestR)
library(BI)
library(arsenal)
library(magrittr)
```

```
##
## Attaching package: 'magrittr'
##
## The following object is masked from 'package:arsenal':
##
```

```
##     set_attr
##
## The following objects are masked from 'package:pracma':
##
##     and, mod, or
##
## The following object is masked from 'package:purrr':
##
##     set_names
##
## The following object is masked from 'package:tidyr':
##
##     extract
```

```
# Extract PSD (Hopkins-25) and pain status information to maintain blinding of the data analyst
psd_data <-
  readRDS(here::here("Lu_Mqadi/self_report_pilot_data/dummy_psd.rds"))
names(psd_data)[names(psd_data) == "id"] <- "pid"
names(psd_data)[names(psd_data) == "group_reported"] <-
  "pain_status"# rename columns

psd_data %<>%
  mutate(across(3:27,
                as.numeric)) %>%
  mutate(age = as.numeric(age)) %>%
  mutate(hop_tot = rowMeans(select(., 3:27))) %>%
  mutate(age = case_when(pid == "UKO-996" ~ 2021 - 1987,
  # Participant's age erroneously recorded as the year of birth - manual corrected here.
  # To be corrected in the database before formal analysis to ensure accuracy.
                         TRUE ~ age))
# Extract additional data on the arm used for HFS, electrical current and binding
hfs_current <-
  readRDS(here::here(
    "Lu_Mqadi/self_report_pilot_data/dummy_current_blinding.rds"
  )) %>%
  select(id, arm, current, blinded, group_guess, assessor_confidence) %>%
  mutate(calibrated_current = (current / 10)) %>%
  select(-current) %>%
  dplyr::rename(pid = id)
names(hfs_current)[names(hfs_current) == "blinded"] <-
  "participants_blinded"

# Now create a new data frame for all the extracted data
df <- psd_data %>%
  left_join(hfs_current) %>%
  select(
    pid,
    age,
    sex,
    pain_status,
    hop_tot,
    arm,
    participants_blinded,
    group_guess,
```

```
    assessor_confidence,
    calibrated_current
  )
```

```
## Joining, by = "pid"
```

Demographic characteristics

1. Determine the total number of participants enrolled and identify the sex of all participants.
2. Summarise descriptive statistics for demographic characteristics of all participants.
3. Create table to summarise demographic characteristics.

```
# number of participants enrolled
sample.n <- length(df$pid)
paste0(sample.n, " participants enrolled")
```

[1] "6 participants enrolled"

```
# participants who identify as male or female
males <- df %>%
  filter(sex == "male") %>%
  nrow()
paste0(males, " male participants")  # males
```

[1] "5 male participants"

```
females <- df %>%
  filter(sex == "female") %>%
  nrow()
paste0(females, " female participant")  # females
```

[1] "1 female participant"

```
# participants who reported persistent pain or no pain
n_group_a <- df %>%
  filter(pain_status == "a") %>%
  nrow()
paste0(n_group_a, " participants in group 'a'")  # group 'a'
```

[1] "1 participants in group 'a' "

```
# 1 participant; note that there is only 1 participant in
# group 'a'; therefore, some statistical tests may be
# unavailable for this group due to insufficient sample
# size (size < 3).

n_group_b <- df %>%
  filter(pain_status == "b") %>%
  nrow()
paste0(n_group_b, " participants in group 'b'")  # group 'b'
```

[1] "5 participants in group 'b'"

```
# summarise 'age' data for all participants and round off
# to 0 or 1 decimal place where necessary
controls_age <- arsenal::tableby.control(numeric.stats = c("N",
  "mean", "sd", "median", "range"), cat_stats = c("countpct"),
  stats.labels = c(list(N = "number of participants", mean = "mean",
    standard_deviation = "sd", median = "median", range = "range")))
table_one <- tableby(pain_status ~ sex + age, data = df, control = controls_age)
summary(table_one, text = TRUE) %>%
  kbl(caption = "Demographic characteristics of participants (N=6)",
    booktabs = T) %>%
  kable_styling(latex_options = c("striped", "HOLD_position"))
```

Table 1: Demographic characteristics of participants (N=6)

	a (N=1)	b (N=5)	Total (N=6)	p value
sex				0.624
- female	0 (0.0%)	1 (20.0%)	1 (16.7%)	
- male	1 (100.0%)	4 (80.0%)	5 (83.3%)	
age				0.953
- number of participants	1	5	6	
- mean	42.000	43.000	42.833	
- SD	NA	14.663	13.121	
- median	42.000	34.000	38.000	
- range	42.000 - 42.000	31.000 - 60.000	31.000 - 60.000	

```
# summarise 'current' data for all participants and round
# off to 1 decimal place.
controls_current <- arsenal::tableby.control(numeric.stats = c("N",
  "mean", "sd", "median", "range"), cat_stats = c("countpct"),
  stats.labels = c(list(N = "number of participants", mean = "mean",
    standard_deviation = "sd", median = "median", range = "range")))
table_two <- tableby(pain_status ~ sex + calibrated_current,
  data = df, control = controls_current)
summary(table_two, text = TRUE) %>%
  kbl(caption = "Individually calibrated current (mA)", booktabs = T) %>%
  kable_styling(latex_options = c("striped", "HOLD_position"))
```

Table 2: Individually calibrated current (mA)

	a (N=1)	b (N=5)	Total (N=6)	p value
sex				0.624
- female	0 (0.0%)	1 (20.0%)	1 (16.7%)	
- male	1 (100.0%)	4 (80.0%)	5 (83.3%)	
calibrated_current				0.305
- number of participants	1	5	6	
- mean	1.300	1.680	1.617	
- SD	NA	0.295	0.306	
- median	1.300	1.800	1.700	
- range	1.300 - 1.300	1.200 - 1.900	1.200 - 1.900	

Potential confounders

Relationship between the individually calibrated electrical current and SH outcomes

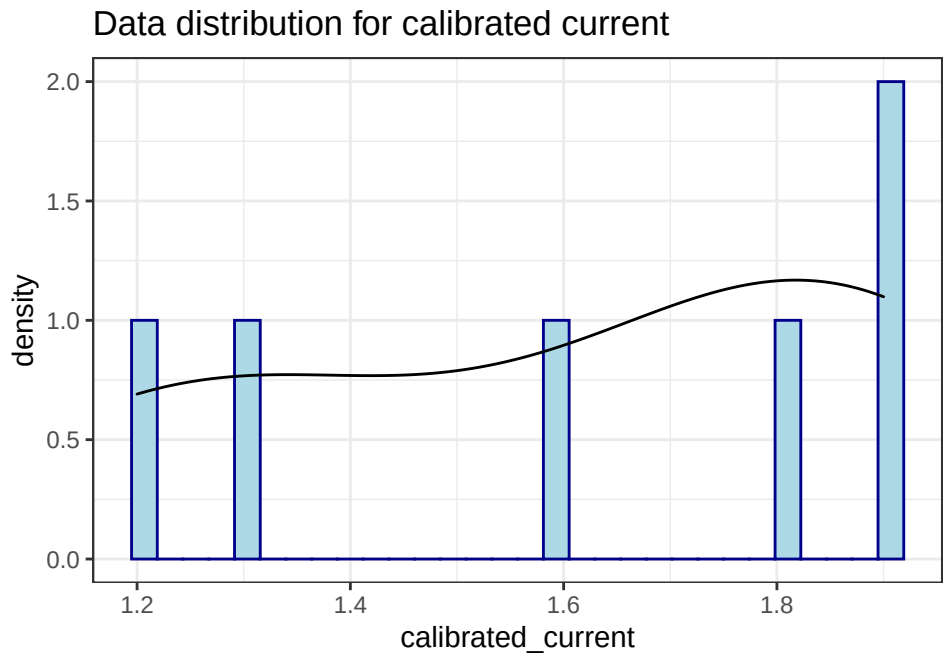
Previous studies have found no relationship between the magnitude of electrical current used for the induction of SH and the outcomes, i.e. surface area and magnitude of SH. We will assess the relationship between the calibrated current and the outcomes.

Surface area of SH and calibrated current

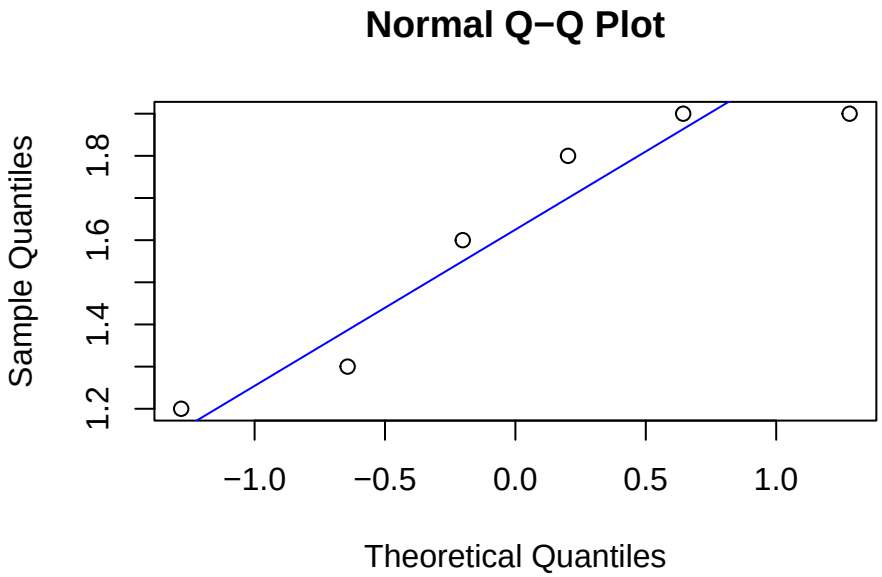
1. Formally check the distribution of the 'calibrated current' data using histogram, Q-Q plot, and Shapiro-Wilk's test of normality.
2. Assess the relationship between the 'surface area' of SH and electric current:
    - Import SH 'surface area' data;
    - Calculate the area under the line (AUL) for each participant across the three time points using the (click link) **trapezoidal integration rule** ;
    - Plot 'surface area' and electric current data;
    - Formally check the distribution of the surface area data using a histogram, Q-Q plot and Shapiro-Wilk's test of normality;
    - Use Pearson's correlation test to assess the relationship between the surface area of SH and electric current.

```
# assess whether the individual calibration approach
# confounded the outcomes.  formally check distribution of
# the electric current data

# histogram
ggplot(df, aes(x = calibrated_current)) + geom_histogram(color = "darkblue",
  fill = "lightblue") + labs(title = "Data distribution for calibrated current") +
  theme_bw() + geom_density()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
# interpretation: clearly not a normal distribution; no
# need to run subsequent plots/tests to verify this, but we
# will include the code in case formal data histogram is
# not conclusive.

# Q-Q plot
qqnorm(df$calibrated_current, pch = 1)
qqline(df$calibrated_current, col = "blue")
```



```r
# Shapiro-wilk's test
with(df, shapiro.test(calibrated_current))  # interpretation:
```

```
##
##  Shapiro-Wilk normality test
##
## data:  calibrated_current
## W = 0.86412, p-value = 0.2038
```

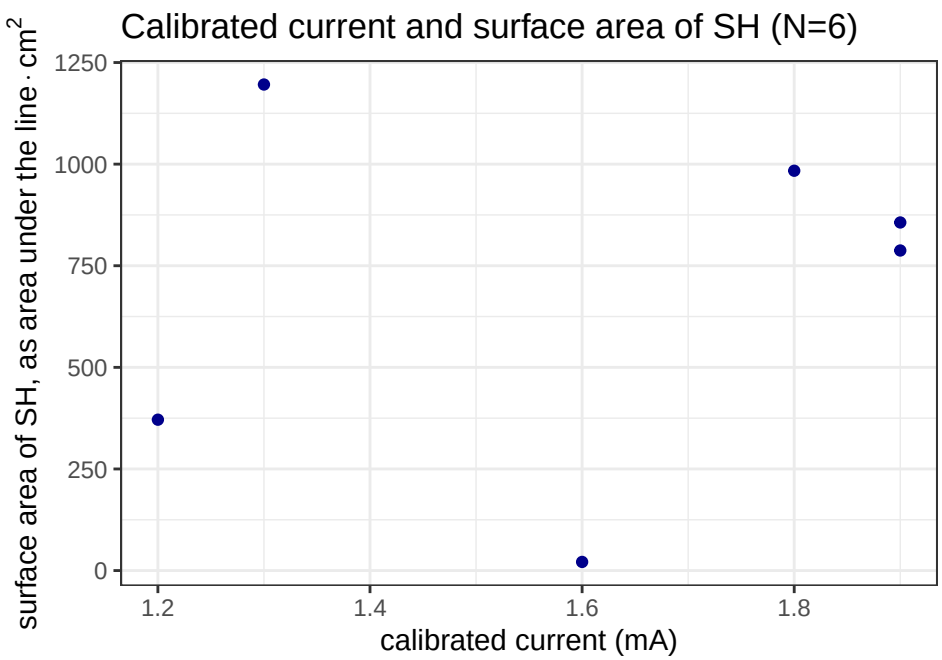
```r
# p-value = 0.20 data are normally distributed.

# import SH surface area data
df1 <- read_excel(here::here("Lu_Mqadi/SH_pilot_data/area_all_test_data.xlsx")) %>%
  select(pid, time_point, radial_line, SA_point)

# convert time column
df1 %<>%
  separate(col = time_point, sep = "_", into = c("delete",
    "time")) %>%
  mutate(time = as.numeric(time)) %>%
  select(-delete) %>%
  dplyr::rename(length_1 = SA_point)

# calculate length 2 of the triangles
df1 %<>%
  group_by(pid, time) %>%
  mutate(length_2 = case_when(radial_line == "A" ~ lead(length_1,
    1), radial_line == "B" ~ lead(length_1, 1), radial_line ==
    "C" ~ lead(length_1, 1), radial_line == "D" ~ lead(length_1,
    1), radial_line == "E" ~ lead(length_1, 1), radial_line ==
    "F" ~ lead(length_1, 1), radial_line == "G" ~ lead(length_1,
    1), radial_line == "H" ~ lag(length_1, 7))) %>%
  ungroup()

# calculate surface area for each triangle
sin_angle <- sin((deg2rad(45)))
df1 %<>%
  mutate(triangle_sa = (length_1 * length_2 * sin_angle)/2)

# calculate total surface area of SH at t 30, 45, and 60
# minutes
primary_outcome <- df1 %>%
  select(-radial_line, -length_1, -length_2) %>%
  dplyr::ungroup() %>%
  dplyr::group_by(pid, time) %>%
  dplyr::summarise(area = sum(triangle_sa)) %>%
  ungroup()
```

```
## `summarise()` has grouped output by 'pid'. You can override using the `.groups`
## argument.
```

```
# now calculate AUL for each participant from t30 to t60
# and remove duplicates.
aul <- primary_outcome %>%
  dplyr::group_by(pid) %>%
  mutate(sa = area_under_curve(time, area, method = "trapezoid")) %>%
  select(pid, sa) %>%
  ungroup %>%
  distinct()

df %<>%
  left_join(aul)
```

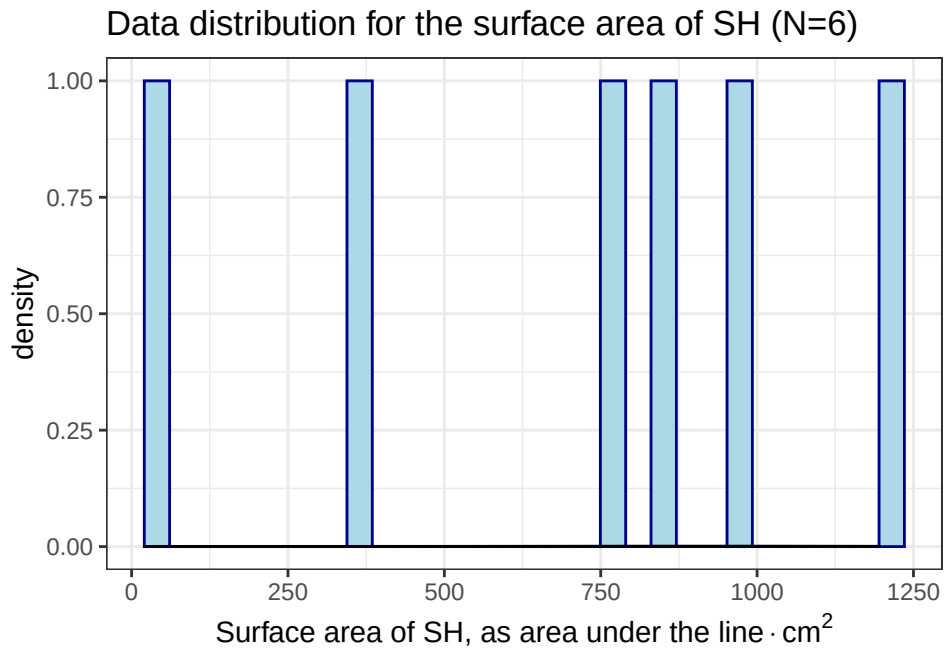
```
## Joining, by = "pid"
```

```
# plot 'surface area' and electric current data
ggplot(df) + aes(x = calibrated_current, y = sa) + geom_point(colour = "darkblue") +
  labs(title = "Calibrated current and surface area of SH (N=6)",
    x = "calibrated current (mA)", y = expression("surface area of SH, as area under the line" %.%
      cm^{
        2
      })) + theme_bw()
```

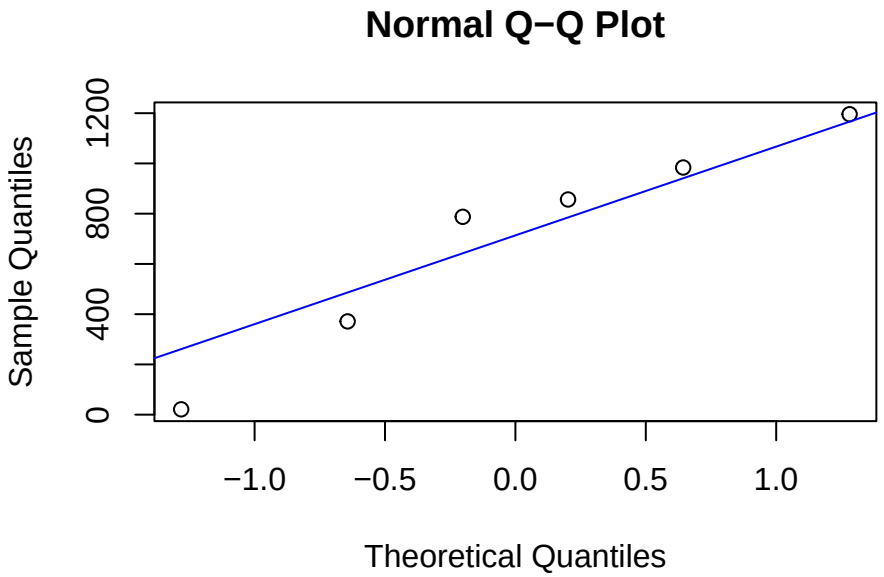



```
# formally check distribution of the surface area data
# histogram
ggplot(df, aes(x = sa)) + geom_histogram(color = "darkblue",
  fill = "lightblue") + labs(title = "Data distribution for the surface area of SH (N=6)",
  x = expression("Surface area of SH, as area under the line" %.%
    cm^{
      2
    })) + theme_bw() + geom_density()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```r
# Q-Q plot
qqnorm(df$sa, pch = 1)
qqline(df$sa, col = "blue")  # interpretation:
```




```r
# Shapiro-wilk's test
with(df, shapiro.test(sa))
```

```
##
##  Shapiro-Wilk normality test
##
## data:  sa
## W = 0.93929, p-value = 0.6535
```

```
# p-value = 0.6535, data are normally distributed; we will
# use Pearson's correlation test to test for the
# relationship between electric current and the surface
# area of SH. we will also use Kendall's and Spearman's
# correlation tests depending on the distribution of the
# surface area data in the formal dataset.

# Pearson's correlation test
cor.test(df$calibrated_current, df$sa, method = c("pearson"))
```

```
##
##  Pearson's product-moment correlation
##
## data:  df$calibrated_current and df$sa
## t = 0.34627, df = 4, p-value = 0.7466
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.7439656  0.8627159
## sample estimates:
##       cor
## 0.1705987
```

```
# r = 0.17, p = 0.75; the test indicates no significant
# relationship between the electric current and surface
# area of SH. Therefore, we will not include electric
# current as a covariate when analysing the surface area of
# SH.

# Kendall's correlation test
cor.test(df$calibrated_current, df$sa, method = c("kendall"))  # tau = 0
```

```
##
##  Kendall's rank correlation tau
##
## data:  df$calibrated_current and df$sa
## z = 0, p-value = 1
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
## tau
##   0
```

```
# Spearman's correlation test
cor.test(df$calibrated_current, df$sa, method = c("spearman"))  # rho = 0.12
```

```
##
##  Spearman's rank correlation rho
```

```
##
## data:  df$calibrated_current and df$sa
## S = 30.942, p-value = 0.8268
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.1159542
```

Magnitude of SH and calibrated current

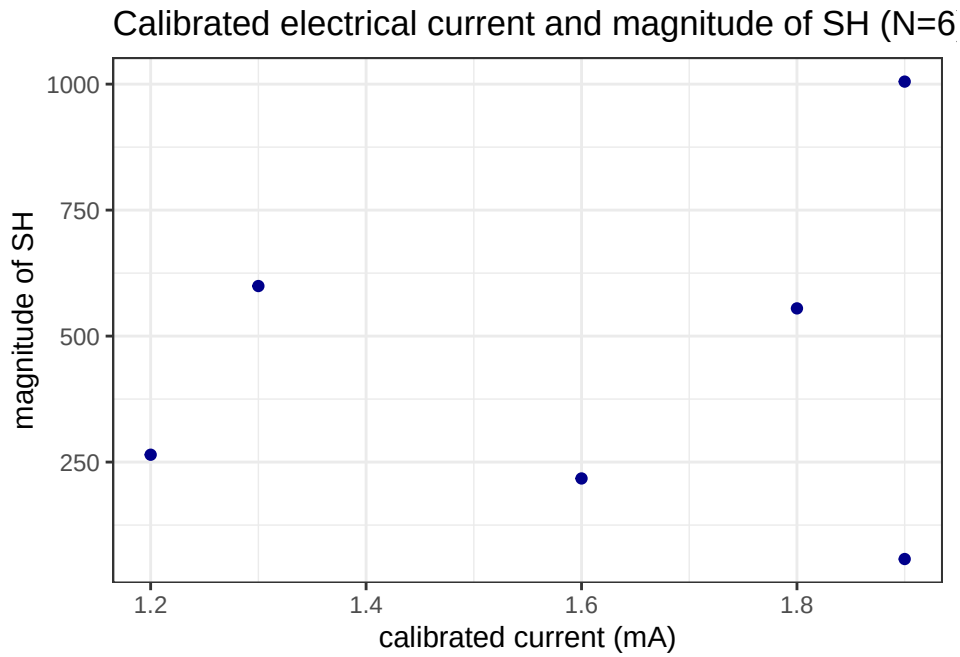
1. Assess the relationship between SH 'magnitude' and electric current:
    - Import SH 'surface area' data;
    - Calculate the area under the line (AUL) for each participant across the three time points using the trapezoidal integration rule;
    - Plot the 'magnitude' and electric current data;
    - Formally check the distribution of the magnitude data using histogram, Q-Q plot and Shapiro-Wilk's test of normality;
    - Use Pearson's correlation test to assess the relationship between the electric current and magnitude of SH.

```
# prepare vector listing all files in folder that start with "UKO"
vector_response_files <-
  list.files(here("Lu_Mqadi/SH_pilot_data/")) %>%
  str_subset(pattern = "^~", negate = TRUE) %>%
  str_subset(pattern = "UKO-") #list all relevant files in folder
vector_response_files <-
  paste0("Lu_Mqadi/SH_pilot_data/", vector_response_files)

# prepare empty data frame
secondary_outcome <- data.frame()

# use for loop to rbind all data together
for (i in vector_response_files) {
  id_pre <- str_remove(i, "Lu_Mqadi/SH_pilot_data/") %>%
    str_remove(".xlsx")

  p <- read_excel(here(i)) %>%
    mutate(pid = id_pre) %>%
    select(pid, time_point, modality, rating) %>%
    filter(time_point %in% c("baseline_point", "time_35", "time_50", "time_65")) %>%
    filter(modality %in% c("pp_128", "pp_256"))

  secondary_outcome <- rbind(secondary_outcome, p)
}

# convert modality column
secondary_outcome %<>%
  mutate(modality = ifelse(modality == 'pp 0',
                           'pp_0',
                           modality)) %>%
  group_by(pid, time_point, modality) %>%
  mutate(rating = mean(rating) + 0.1) %>%
  # calculate mean for baseline and add small constant to avoid zero values
```

```
  unique() %>%
  ungroup() %>%
  group_by(pid, modality) %>%
  # group and pivot to express follow-up ratings as % of baseline rating
  pivot_wider(names_from = time_point,
              values_from = rating) %>%
  mutate(
    perc_35 = time_35 / baseline_point,
    perc_50 = time_50 / baseline_point,
    perc_65 = time_65 / baseline_point
  ) %>%
  select(pid, modality, contains("perc")) %>%
  group_by(pid, modality) %>%
  pivot_longer(cols = 3:5,
               names_to = "time_point",
               values_to = "perc_change") %>%
  group_by(pid,
           time_point) %>%
  pivot_wider(names_from = modality,
              values_from = perc_change) %>%
  ungroup() %>%
  rowwise %>%
  mutate(sh = mean(c_across(contains("pp_")))) %>%
  # find the mean for the two pinprick weights
  select(pid,
         time_point, sh) %>%
  separate(col = time_point,
           sep = '_',
           into = c('delete', 'time')) %>%
  select(-delete) %>% # delete useless column
  mutate(time = as.numeric(time))

aul2 <- secondary_outcome %>%
  group_by(pid) %>%
  mutate(mag = area_under_curve(time,
                                sh,
                                method = "trapezoid")) %>%
  select(pid, mag) %>%
  distinct()

df %<>% left_join(aul2)
```

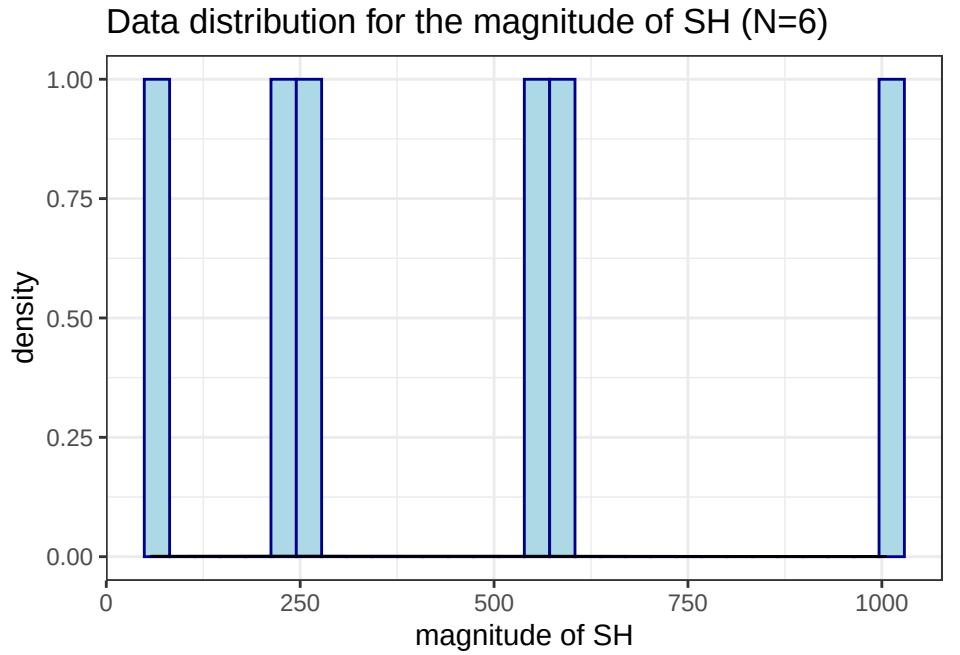
```
## Joining, by = "pid"
```

```
# plot 'magnitude' and electric current data
ggplot(df) +
  aes(x = calibrated_current,
      y = mag) +
  geom_point(color = "darkblue") +
  labs(title = "Calibrated electrical current and magnitude of SH (N=6)",
       y = "magnitude of SH",
       x = "calibrated current (mA)") +
  theme_bw()
```

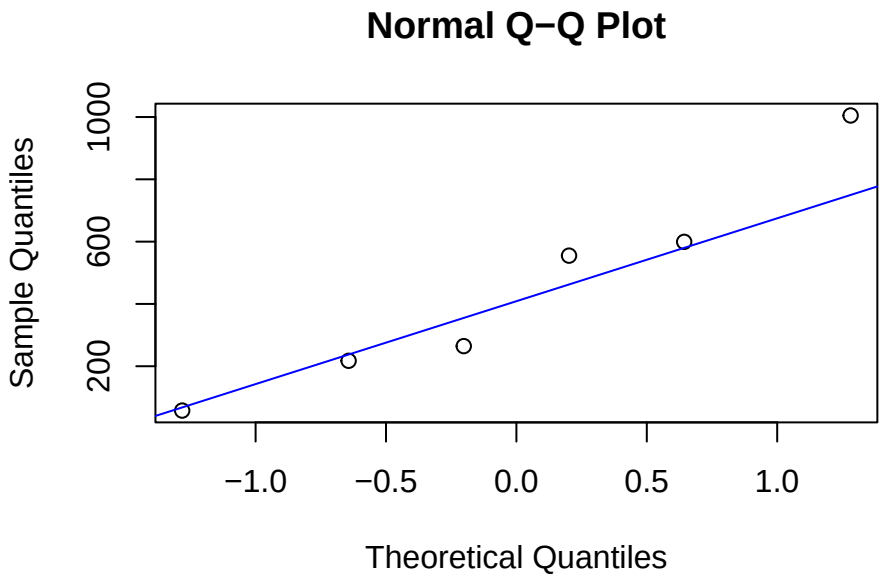



```
# formally check distribution of the magnitude data
# histogram
ggplot(df, aes(x = mag)) +
  geom_histogram(color = "darkblue", fill = "lightblue") +
  labs(title =  "Data distribution for the magnitude of SH (N=6)", x = "magnitude of SH")  + theme_bw()
  geom_density()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



```r
# Q-Q plot
qqnorm(df$mag, pch = 1)
qqline(df$mag, col = "blue") # interpretation:
```




```r
# Shapiro-wilk's test
with(df, shapiro.test(mag))
```

```
##
##  Shapiro-Wilk normality test
##
## data:  mag
## W = 0.94304, p-value = 0.6837
```

```r
# p-value = 0.68, data are normally distributed; use Pearson's correlation test to test for the
# relationship between current and the magnitude of SH.
# we will also use Kendall's and Spearman's correlation tests
# depending on the distribution of the magnitude data in the formal dataset.

# Pearson correlation test
cor.test(df$calibrated_current, df$mag, method = c("pearson"))
```

```
##
##  Pearson's product-moment correlation
##
## data:  df$calibrated_current and df$mag
## t = 0.38635, df = 4, p-value = 0.7189
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  -0.7350356  0.8676709
## sample estimates:
```

```
##       cor
## 0.1896672
```

```
# r = 0.19 , p = 0.72; the test indicates a positive weak relationship between
# the electric current and the magnitude of SH.
# therefore, we will exclude electric current as a covariate when analysing the magnitude of SH.

# Kendall's correlation test
cor.test(df$calibrated_current, df$mag, method = c("kendall")) # tau = 0.5520524
```

```
##
##  Kendall's rank correlation tau
##
## data:  df$calibrated_current and df$mag
## z = 0, p-value = 1
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
## tau
##   0
```

```
# Spearman's correlation test
cor.test(df$calibrated_current, df$mag, method = c("spearman")) # rho = 0.6667367
```

```
##
##  Spearman's rank correlation rho
##
## data:  df$calibrated_current and df$mag
## S = 35, p-value = 1
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
## rho
##   0
```

```
# the individual calibration approach did not confound the surface area (primary outcome)  or
# the magnitude (secondary outcome) of SH. therefore, we will
# exclude electric current as a covariate when # analysing the
# data for the surface area and magnitude of SH.
```

Relationship between HFS painfulness and SH outcomes

We will check for a relationship between the participant-reported painfulness of HFS (5 ratings per participant) and the outcomes (surface area and magnitude of SH) or group, and include painfulness of the HFS as a covariate if indicated by these results. To do this we will:

1. Formally check the distribution of the 'HFS VAS ratings' using histogram, Q-Q plot and Shapiro-Wilk's test of normality.
2. Plot sample and individual HFS VAS ratings.
3. Calculate and tabulate the the mean, median, SD, min, max, and confidence interval for the HFS VAS ratings.
4. Assess the relationship between:
    - SH 'surface area' and HFS painfulness;

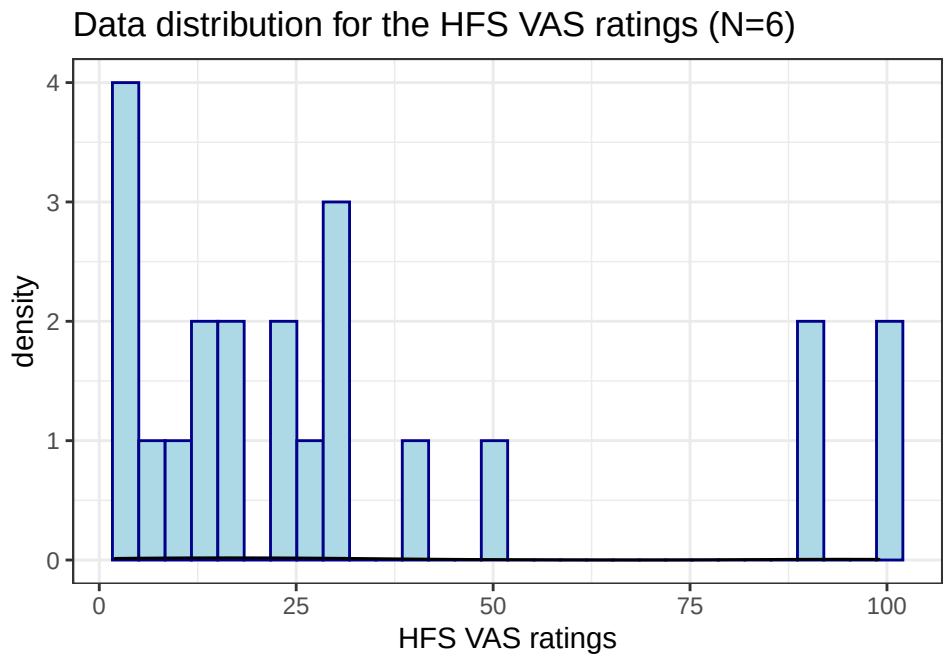
- SH 'magnitude' and HFS painfulness.

```
# prepare data
hfs_ratings <- read_excel(here("Lu_Mqadi/SH_pilot_data/hfs_ratings_all_test_data.xlsx")) %>%
  select(pid, train_counter, rating) %>%
  mutate(rating = as.numeric(rating)) %>%
  mutate(train_counter = as.numeric(train_counter))  # convert to numeric
names(hfs_ratings)[names(hfs_ratings) == "rating"] <- "train_rating"  # rename columns

hfs <- df %>%
  select(pid, pain_status, sex, sa, mag)
hfs_pain <- hfs_ratings %>%
  left_join(hfs)  # join HFS ratings with SH outcomes
```

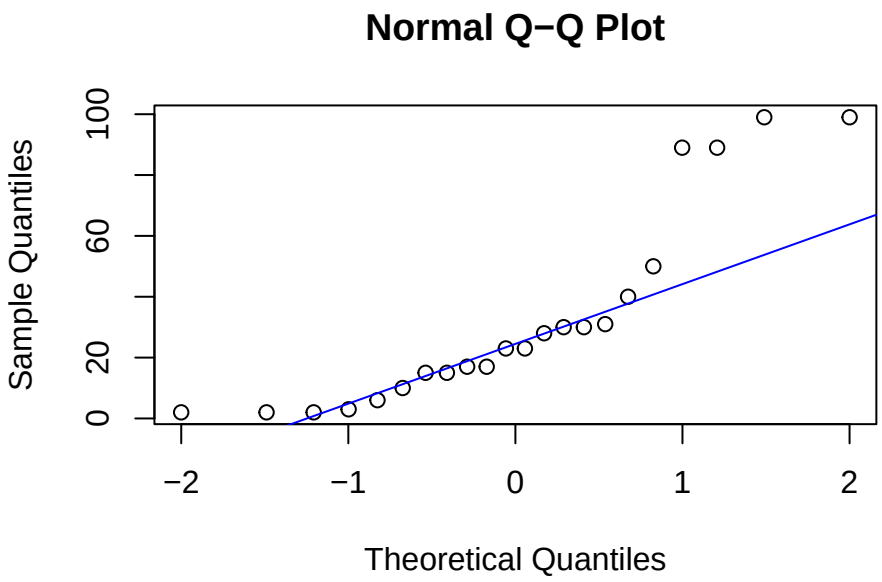
```
## Joining, by = "pid"
```

```
# formally check distribution of the HFS data histogram
ggplot(hfs_pain, aes(x = train_rating)) + geom_histogram(color = "darkblue",
  fill = "lightblue") + labs(title = "Data distribution for the HFS VAS ratings (N=6)",
  x = "HFS VAS ratings") + theme_bw() + geom_density()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
# Q-Q plot
qqnorm(hfs_pain$train_rating, pch = 1)
qqline(hfs_pain$train_rating, col = "blue")  # interpretation:
```

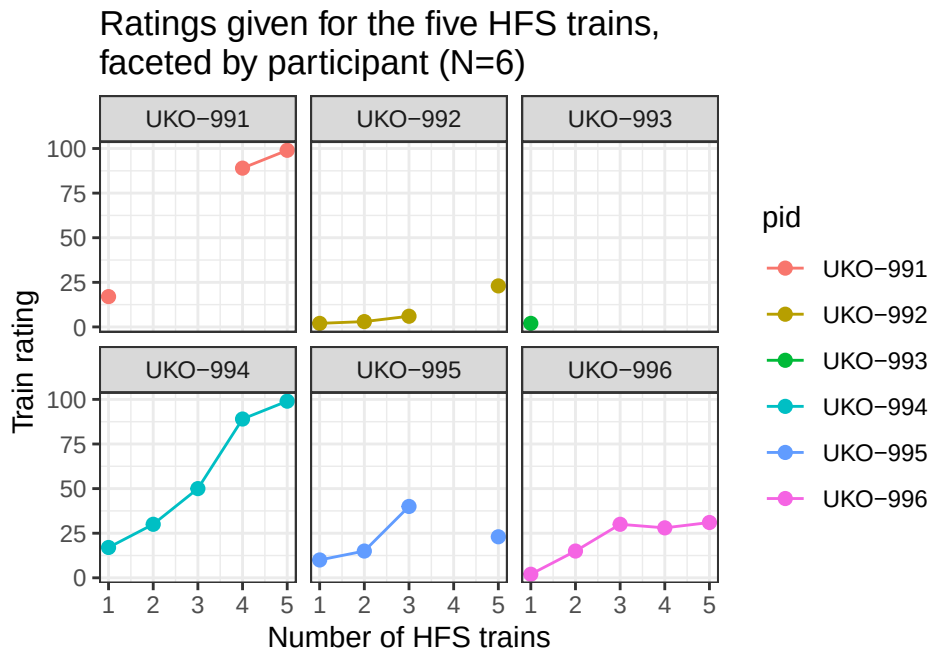



```
# Shapiro-wilk's test
with(hfs_pain, shapiro.test(train_rating))
```

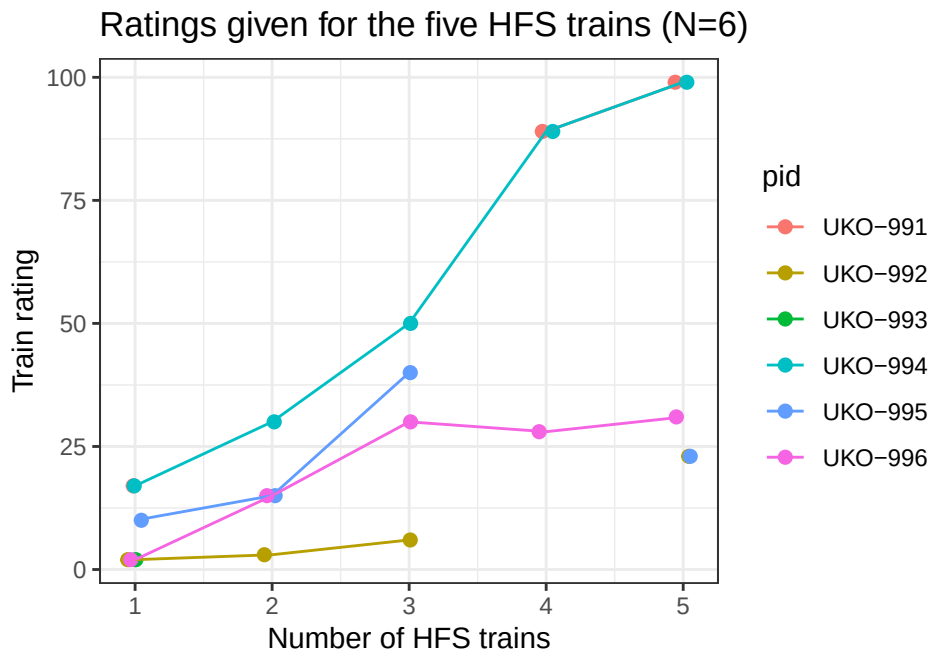
```
##
##  Shapiro-Wilk normality test
##
## data:  train_rating
## W = 0.80347, p-value = 0.0005672
```

```
# p-value = 0.0005672, data are not normally distributed;
# considering that this is a pilot script, we will run the
# Pearson's, Spearman's and Kendall's correlations tests
# depending on the distribution of the HFS VAS ratings in
# the formal dataset.

# plot HFS VAS ratings; please note missing values on plot
ggplot(hfs_ratings) + aes(x = train_counter, y = train_rating,
  group = pid, colour = pid) + geom_point(data = hfs_ratings,
  aes(x = train_counter, y = train_rating, colour = pid, fill = pid),
  shape = 21, size = 2) + geom_line() + facet_wrap(~pid) +
  labs(title = "Ratings given for the five HFS trains,\nfaceted by participant (N=6)",
    x = "Number of HFS trains", y = "Train rating") + theme_bw()
```

```
## geom_path: Each group consists of only one observation. Do you need to adjust
## the group aesthetic?
```




```
ggplot(hfs_ratings) + aes(x = train_counter, y = train_rating,
  group = pid, colour = pid) + geom_point(data = hfs_ratings,
  aes(x = train_counter, y = train_rating, colour = pid, fill = pid),
  shape = 21, size = 2, position = position_jitter(width = 0.06,
    height = 0)) + geom_line() + labs(title = "Ratings given for the five HFS trains (N=6)",
  x = "Number of HFS trains", y = "Train rating") + theme_bw()
```




```
# calculate the confidence interval
sample.mean <- round(mean(hfs_ratings$train_rating, na.rm = TRUE),
```

```
  1)  # sample mean
paste0(sample.mean, " sample mean")
```

```
## [1] "32.7 sample mean"
```

```
sample.sd <- round(sd(hfs_ratings$train_rating, na.rm = TRUE),
  1)  # sample SD
paste0(sample.sd, " sample standard deviation")
```

```
## [1] "32.2 sample standard deviation"
```

```
sample.se <- sample.sd/sqrt(sample.n)  # standard error
paste0(sample.se, " sample error")
```

```
## [1] "13.1455949529364 sample error"
```

```
alpha <- 0.05
degrees.freedom <- sample.n - 1
paste0(degrees.freedom, " degrees of freedom")
```

```
## [1] "5 degrees of freedom"
```

```
t.score <- qt(p = alpha/2, df = degrees.freedom, lower.tail = FALSE)  # now calculate t-scores
paste0(t.score, " t-score corresponding to the CI")
```

```
## [1] "2.57058183563631 t-score corresponding to the CI"
```

```
margin.error <- t.score * sample.se  # margin of error
paste0(margin.error, " margin of error")
```

```
## [1] "33.7918276046507 margin of error"
```

```
lower.bound <- sample.mean - margin.error  # CI lower bound
upper.bound <- sample.mean + margin.error  # CI upper bound
CI <- print(c(lower.bound, upper.bound))
```

```
## [1] -1.091828 66.491828
```

```
# we are 95% confident that the population mean falls
# between -1.09 and 66.49

# summarise HFS ratings for all participants and round off
# to 1 decimal place.
hfs_ratings %>%
  summarise(Mean = round(mean(train_rating, na.rm = TRUE),
    1), Median = round(median(train_rating, na.rm = TRUE),
    1), SD = round(sd(train_rating, na.rm = TRUE), 1), Min = min(train_rating,
    na.rm = TRUE), Max = max(train_rating, na.rm = TRUE)) %>%
  kbl(caption = "Descriptive statistics for HFS VAS ratings for all participants (N=6).",
    booktabs = T) %>%
  kable_styling(latex_options = c("striped", "HOLD_position"))
```

Table 3: Descriptive statistics for HFS VAS ratings for all participants (N=6).

Mean	Median	SD	Min	Max
32.7	23	32.2	2	99

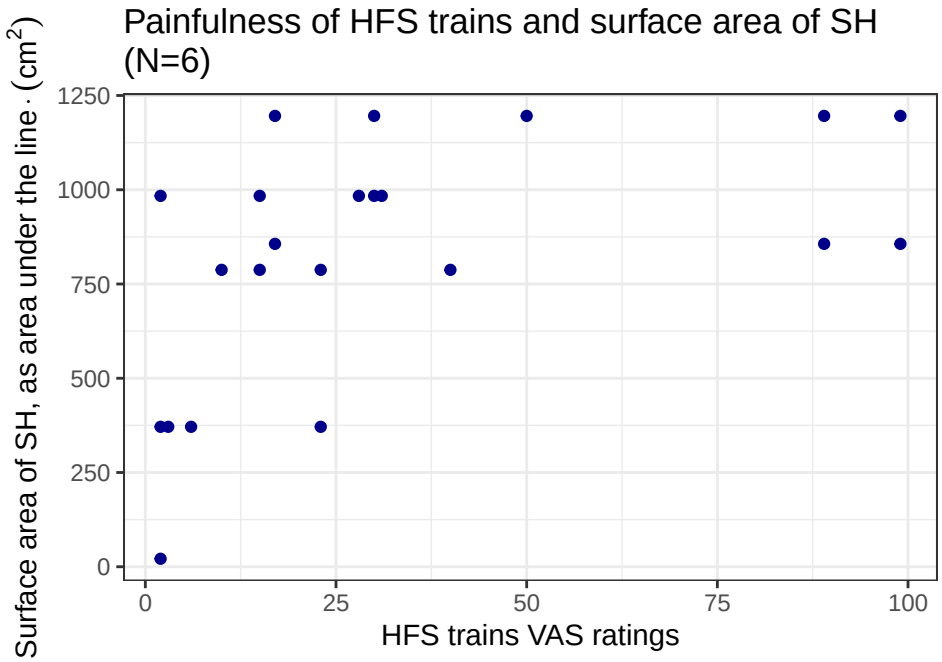
```
# calculate and tabulate mean HFS ratings by group
hfs_ratings %>%
  right_join(df) %>%
  select(pid, pain_status, train_rating) %>%
  group_by(pain_status) %>%
  summarise(mean = round(mean(train_rating, na.rm = TRUE),
    1), SD = round(sd(train_rating, na.rm = TRUE), 2)) %>%
  kbl(caption = "Descriptive statistics for HFS VAS ratings by group (i.e, pain status) (N=6)",
    booktabs = T) %>%
  kable_styling(latex_options = c("striped", "HOLD_position"))
```

```
## Joining, by = "pid"
```

Table 4: Descriptive statistics for HFS VAS ratings by group (i.e, pain status) (N=6)

pain_status	mean	SD
a	57.0	35.94
b	25.6	28.27

```
# relationship between HFS painfulness and surface area of
# SH.  plot 'surface area' and HFS ratings
ggplot(hfs_pain) + aes(x = train_rating, y = sa) + geom_point(colour = "darkblue") +
  labs(title = "Painfulness of HFS trains and surface area of SH \n(N=6)",
    x = "HFS trains VAS ratings", y = expression("Surface area of SH, as area under the line" %.%
      (cm^{
        2
      }))) + theme_bw()
```




```
# Pearson correlation test
cor.test(hfs_pain$train_rating, hfs_pain$sa, method = c("pearson"))
```

```
##
##  Pearson's product-moment correlation
##
## data:  hfs_pain$train_rating and hfs_pain$sa
## t = 2.5154, df = 20, p-value = 0.02055
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.08650964 0.75565994
## sample estimates:
##       cor
## 0.4902376
```

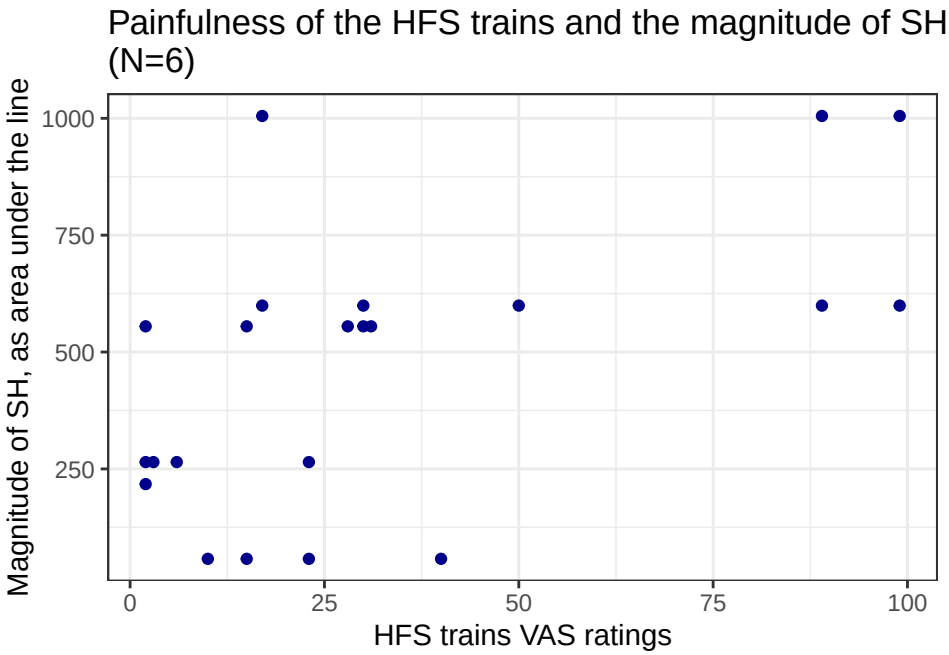
```
# r = 0.49 , p = 0.02; the test indicates a moderate
# positive relationship between the HFS pain and the
# surface of SH.

# Kendall's correlation test
cor.test(hfs_pain$train_rating, hfs_pain$sa, method = c("kendall"))
```

```
##
##  Kendall's rank correlation tau
##
## data:  hfs_pain$train_rating and hfs_pain$sa
## z = 2.806, p-value = 0.005017
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## 0.4650153
```

```
# tau = 0.47 p = 0.005017

# Spearman's correlation test
cor.test(hfs_pain$train_rating, hfs_pain$sa, method = c("spearman"))
```

```
##
##  Spearman's rank correlation rho
##
## data:  hfs_pain$train_rating and hfs_pain$sa
## S = 756.18, p-value = 0.00531
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##       rho
## 0.5730222
```

```
# rho = 0.57 p = 0.00531

# the painfulness of the HFS trains may have confounded the
# surface area of SH (primary outcome).  ratings for the
# HFS trains will be included as a covariate when analysing
# the surface area of SH.

# relationship between HFS painfulness and magnitude of SH.
# plot 'magnitude' and HFS ratings
ggplot(hfs_pain) + aes(x = train_rating, y = mag) + geom_point(colour = "darkblue") +
  labs(title = "Painfulness of the HFS trains and the magnitude of SH \n(N=6)",
    x = "HFS trains VAS ratings", y = "Magnitude of SH, as area under the line") +
  theme_bw()
```



```r
# Pearson correlation test
cor.test(hfs_pain$train_rating, hfs_pain$mag, method = c("pearson"))
```

```
##
##  Pearson's product-moment correlation
##
## data:  hfs_pain$train_rating and hfs_pain$mag
## t = 3.0106, df = 20, p-value = 0.00691
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
##  0.1789620 0.7932745
## sample estimates:
##       cor
## 0.5584354
```

```r
# r = 0.56 , p = 0.007; the test indicates a moderate
# positive relationship between the HFS pain and the
# magnitude of SH.

# Kendall's correlation test
cor.test(hfs_pain$train_rating, hfs_pain$mag, method = c("kendall"))
```

```
##
##  Kendall's rank correlation tau
##
## data:  hfs_pain$train_rating and hfs_pain$mag
## z = 2.5745, p-value = 0.01004
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## 0.4266635
```

```r
# tau = 0.43 p = 0.01

# Spearman's correlation test
cor.test(hfs_pain$train_rating, hfs_pain$mag, method = c("spearman"))
```

```
##
##  Spearman's rank correlation rho
##
## data:  hfs_pain$train_rating and hfs_pain$mag
## S = 786.82, p-value = 0.007248
## alternative hypothesis: true rho is not equal to 0
## sample estimates:
##      rho
## 0.555719
```

```r
# rho = 0.56 p = 0.01

# the HFS trains may have also confounded the magnitude of
# SH (secondary outcome).  HFS ratings will be included as
# a covariate when analysing data for the magnitude of SH.
```

SH outcomes

Primary analysis

We have three estimates of the surface area of SH for each participant at 30, 45 and 60 minutes after induction. We are interested in the surface area of SH across the whole period of assessment (pre-induction, 30, 45 and 60 minutes after induction). To determine whether the surface area of experimentally induced SH is predicted by the severity of self-reported PSD and group. We will:

1. Formally check distribution of the raw surface are data using histogram, Q-Q plot and Shapiro-wilk's test of normality
2. Visualise surface area data with 1) PSD data, and 2) pain status data to assess distribution;
3. Formally check the distribution of the surface area (AUL) of SH data, by group using histogram, Q-Q plot and Shapiro-Wilk's test of normality;
4. Formally check distribution of the surface area data, by group (i.e. pain status), using Shapiro-Wilk test of normality;
5. Plot the surface area of SH (AUL) over pain status;
6. Determine if the surface area of experimentally induced SH is predicted by the severity of self-reported PSD and pain status (i.e. group) [Surface area ~ PSD + group + PSD*group] +

```
primary_outcome %<>%
  right_join(df) %>%
  select(pid, pain_status, time, area)
```

```
## Joining, by = "pid"
```

```
# plot the the surface area of SH over time by group
fig1 <- ggplot(data = primary_outcome) + aes(x = time, y = area,
  group = interaction(time, pain_status), fill = pain_status) +
  geom_boxplot() + geom_point(shape = 21, color = "black",
  fill = "black", size = 1) + labs(x = "Time (mins)", y = expression("Surface area of SH" %.%
  cm^{
    2
  })) + scale_y_continuous(limits = c(0, 100)) + scale_x_continuous(breaks = c(30,
  45, 60)) + theme_bw()
grid.arrange(fig1, bottom = "Figure 1: Surface area of SH for each time point after the induction \n,co
```

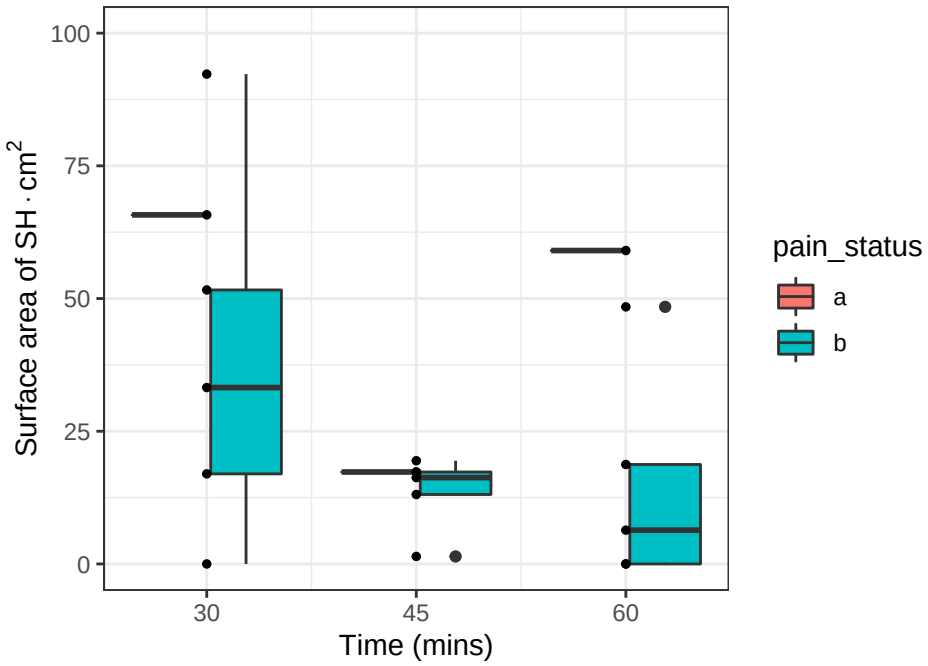


Figure 1: Surface area of SH for each time point after the induction ,coloured by group (N=6).Black dots indicate individual participant scores. Horizontal lines: 25th, 50th (median) and 75th percentile. Boxplot whiskers represent the maximum and minimum values.

```
# does surface area predict PSD and group i.e pain status?
# visualise surface area data against 1) PSD data, and 2)
# pain status data surface area vs PSD
fig2 <- ggplot(data = df) + aes(x = hop_tot, y = sa) + geom_point(color = "darkblue") +
  theme_bw() + labs(x = "Total PSD score", y = expression("Surface area of SH, as area under the line" 
  cm^{
    2
  }))
grid.arrange(fig2, bottom = "Figure 2: Relationship between self-reported PSD and surface \narea of SH"
```

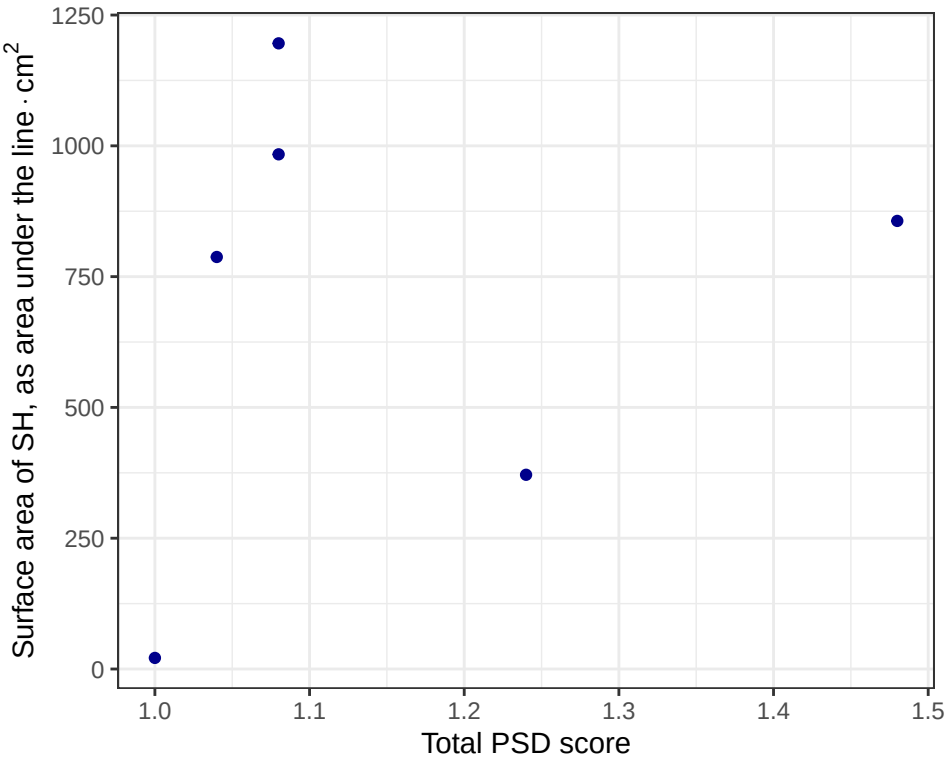


Figure 2: Relationship between self−reported PSD and surface area of SH

```
fig3 <- ggplot(data = df) + aes(x = pain_status, y = sa, fill = pain_status) +
  geom_jitter(color = "black", size = 1.5, alpha = 0.9) + geom_boxplot() +
  theme_bw() + labs(x = "Pain status", y = expression("Surface area of SH, as area under the line" %.%
  cm^{
    2
  }))
grid.arrange(fig3, bottom = "Figure 3: Relationship between pain status and surface area of SH.\n Note
```

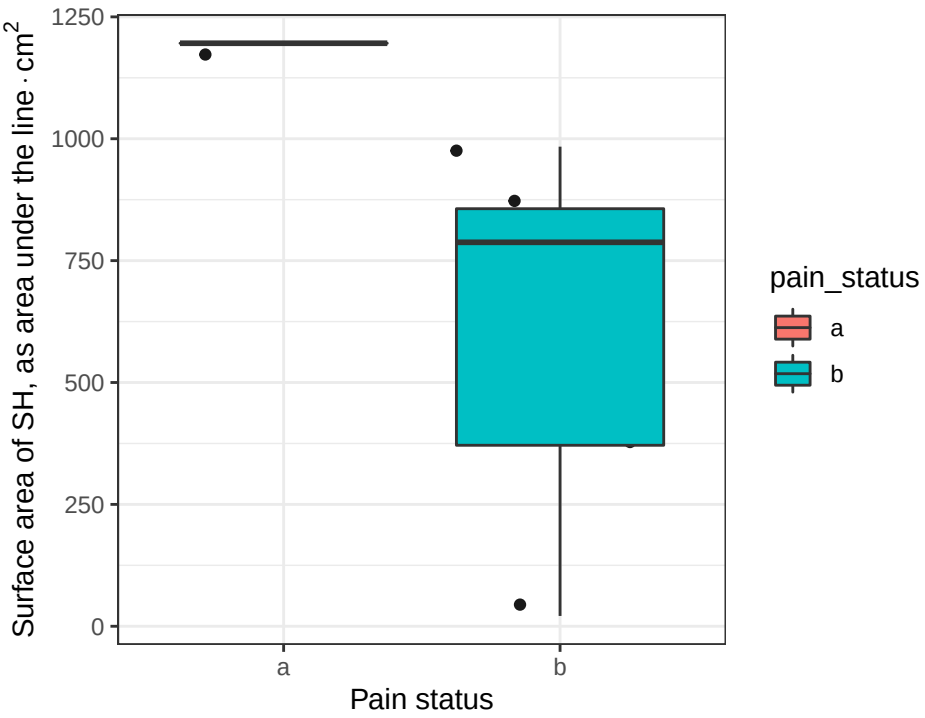
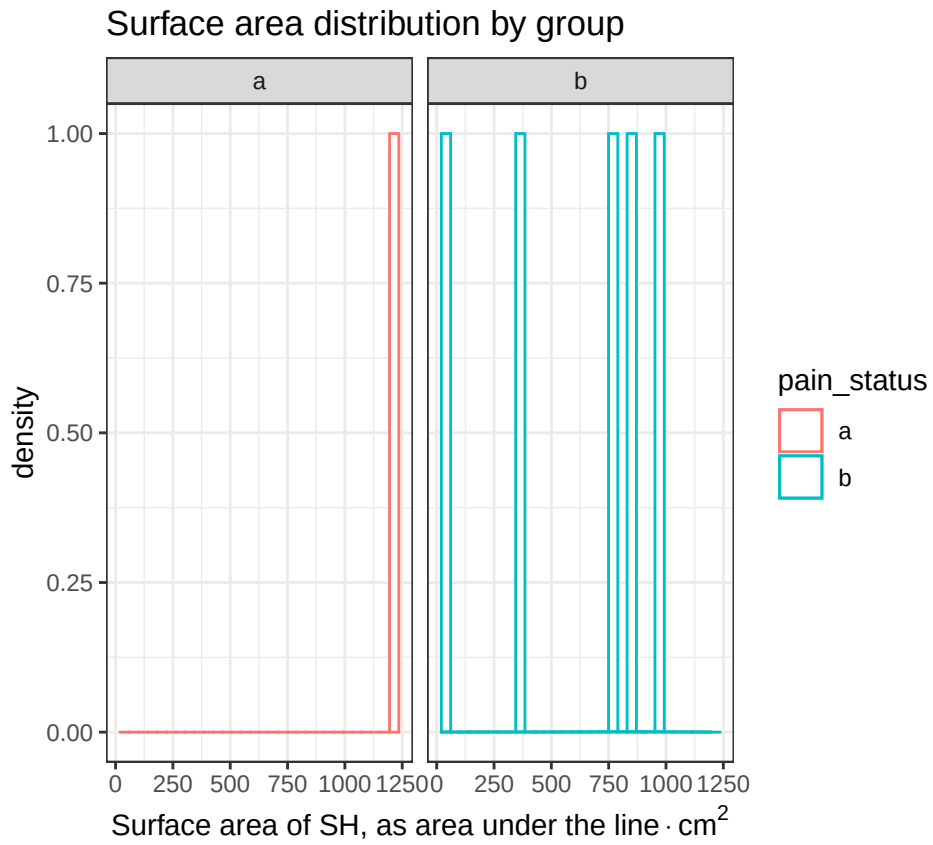


Figure 3: Relationship between pain status and surface area of SH. Note that pain status is either persistent pain or pain−free; interpretation blinded

```
# formally check distribution of the surface area data by
# group, surface area data are normally distributed, p
# -value = 0.6535 (see above)

# histogram
ggplot(df) + aes(x = sa, group = pain_status, color = pain_status) +
  geom_histogram(fill = "white") + geom_density() + facet_wrap(~pain_status) +
  labs(title = "Surface area distribution by group", x = expression("Surface area of SH, as area under
    cm^{
      2
    })) + theme_bw()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
# Shapiro-Wilk test with(df, shapiro.test(sa[pain_status ==
# 'a'])) insufficient participants in group 'a' to run
# test, sample size less than 3
with(df, shapiro.test(sa[pain_status == "b"]))
```

```
##
##  Shapiro-Wilk normality test
##
## data:  sa[pain_status == "b"]
## W = 0.90145, p-value = 0.418
```

```
# p-value = 0.42; data are normally distributed.

# prepare data frame to run model
df2 <- hfs_ratings %>%
  select(pid, train_rating) %>%
  left_join(df)
```

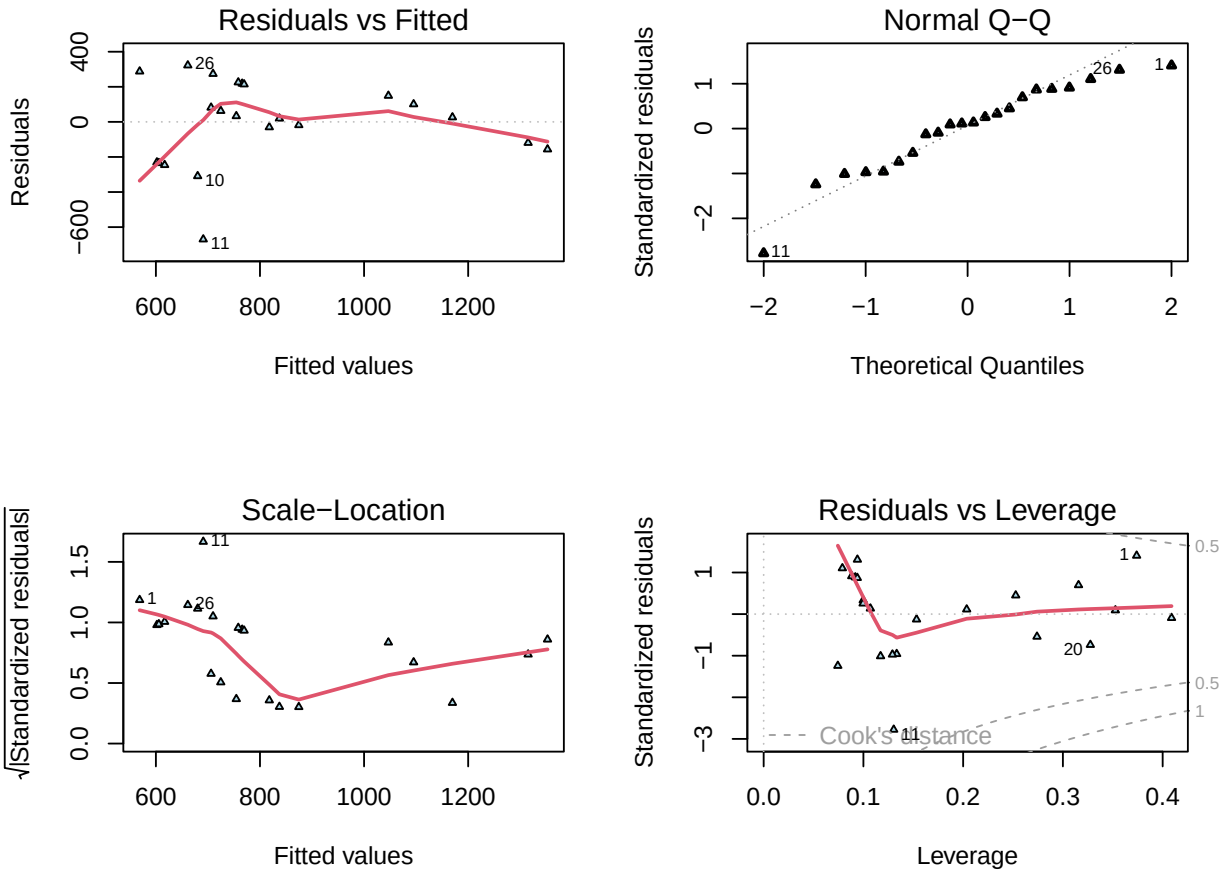
```
## Joining, by = "pid"
```

```
# enter data into model
model <- lm(sa ~ hop_tot + pain_status + train_rating + hop_tot *
  pain_status, df2)
summary(model)
```

```
##
## Call:
## lm(formula = sa ~ hop_tot + pain_status + train_rating + hop_tot *
##     pain_status, data = df2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -669.84 -147.30   29.56  197.99  322.41
##
## Coefficients: (1 not defined because of singularities)
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1384.208    446.679   3.099  0.00619 **
## hop_tot               -371.182    441.898  -0.840  0.41194
## pain_statusb          -329.434    162.987  -2.021  0.05838 .
## train_rating             3.729      2.199   1.696  0.10713
## hop_tot:pain_statusb        NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 258.6 on 18 degrees of freedom
##   (8 observations deleted due to missingness)
## Multiple R-squared:  0.4748, Adjusted R-squared:  0.3872
## F-statistic: 5.424 on 3 and 18 DF,  p-value: 0.007777
```

```
# suggests no relationship between surface area and 1)
# distress and 2) pain status. One unit increase in
# distress is associated with 371.182 units decrease in
# surface area AUC (%); being in pain_status group b
# (rather than group a) is associated with 329.434 unit
# decrease in the surface area AUC (%). 47% of the
# variation in surface area is explained by this model.
```

```
# plot model assumptions
par(mfrow = c(2, 2))
plot(model, pch = 24, bg = "lightblue", colour = "blue", cex = 0.5,
  lwd = 2)
```



Secondary analysis

We have one estimate of SH magnitude for each of the three post-induction time points (35, 50, 65). To determine whether the magnitude of experimentally induced SH is predicted by the severity of self-reported PSD and group. We will:

1. Formally check distribution of the raw magnitude data using histogram, Q-Q plot and Shapiro-wilk's test of normality.
2. Formally check distribution of the magnitude data, by group (i.e. pain status), using Shapiro-Wilk test of normality;
3. Plot the mean magnitude of SH over time (assuming a value of 0 at the pre-induction time point)
4. Plot the magnitude of SH (assuming a value of 0 at the pre-induction time point) and PSD;
5. Determine if the magnitude of experimentally induced SH is predicted by the severity of self-reported PSD and pain status (i.e. group) [Magnitude ~ PSD + group + hop_tot * pain_status] +

```
secondary_outcome %<>%
  right_join(df) %>%
  select(pid, pain_status, time, sh)
```

```
## Joining, by = "pid"
```

```
# plot the SH magnitude over time by group
fig4 <- ggplot(data = secondary_outcome) + aes(x = time, y = sh,
  group = interaction(time, pain_status), fill = pain_status) +
  geom_boxplot() + geom_point(shape = 21, color = "black",
  fill = "black", size = 1) + labs(x = "Time (mins)", y = "SH magnitude (% of rating before induction)"
  scale_y_continuous(limits = c(0, 100)) + scale_x_continuous(breaks = c(35,
  50, 65)) + theme_bw()
grid.arrange(fig4, bottom = "Figure 4: Magnitude of SH over time, coloured by group (N=6).\n Black dots
```

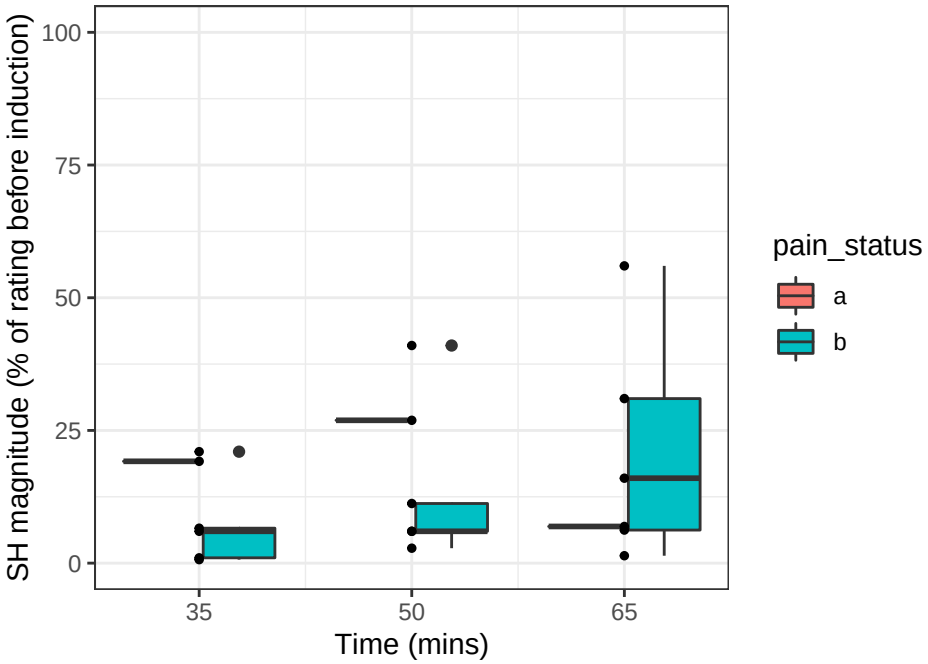


Figure 4: Magnitude of SH over time, coloured by group (N=6). Black dots indicate individual participant scores. Horizontal lines: 25th, 50th (median) and 75th percentile. Boxplot whiskers represent the maximum and minimum values

```
# does magnitude of SH predict PSD and group (pain status)?
# plot PSD vs mag
fig5 <- ggplot(data = df) + aes(x = hop_tot, y = mag) + geom_point(color = "darkblue") +
  theme_bw() + labs(x = "Total PSD scores", y = expression("Magnitude of SH (% of rating before inducti
grid.arrange(fig5, bottom = "Figure 5: Relationship between self-reported PSD and \nmagnitude of SH")
```

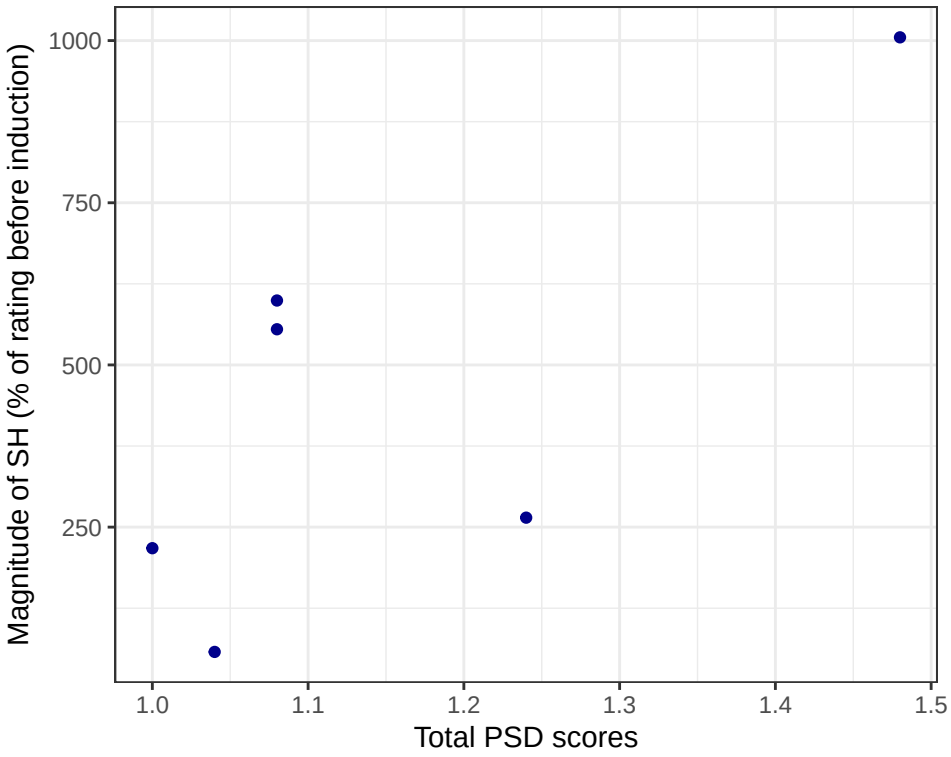


Figure 5: Relationship between self−reported PSD and magnitude of SH

```
# does magnitude of SH predict PSD and group (pain status)?
# plot PSD vs mag
fig6 <- ggplot(data = df) + aes(x = pain_status, y = mag, fill = pain_status) +
  geom_point(color = "darkblue") + geom_boxplot() + theme_bw() +
  labs(x = "Pain status", y = expression("Magnitude of SH (% of rating before induction)"))
grid.arrange(fig6, bottom = "Figure 6: Relationship between the participant's pain \nstatus (i.e., grou
```

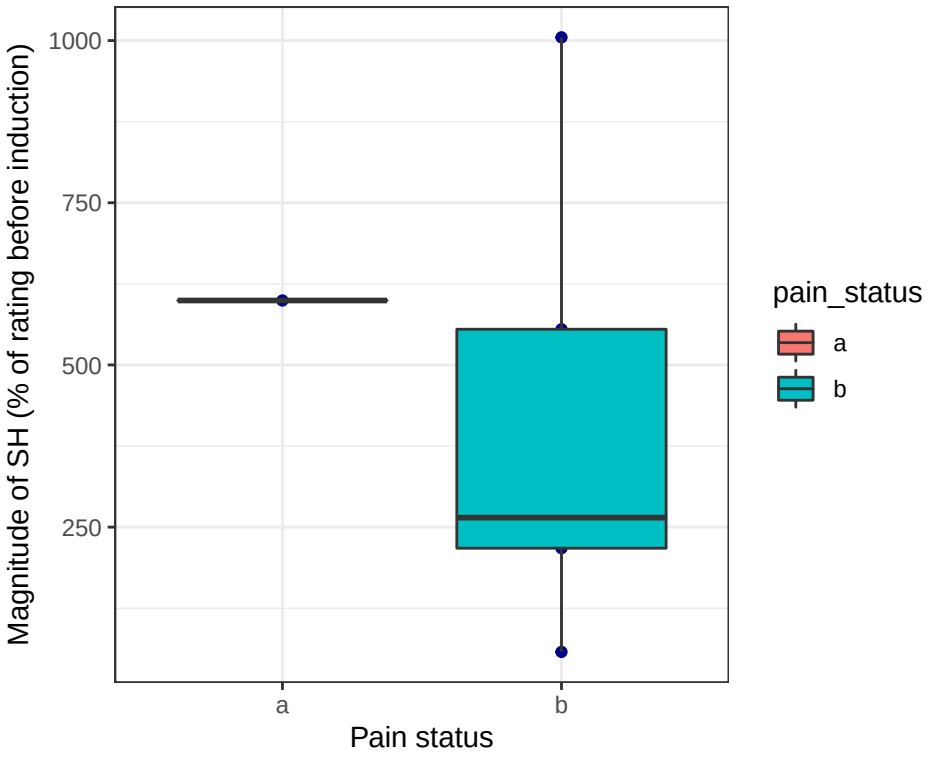


Figure 6: Relationship between the participant's pain status (i.e., group membership) and magnitude of SH

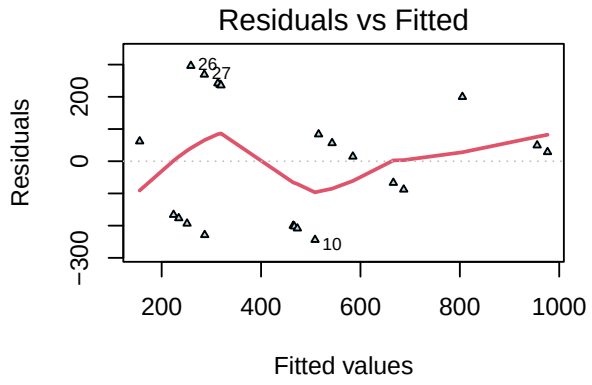
```
# enter data into model
model2 <- lm(mag ~ hop_tot + pain_status + hop_tot * pain_status +
  train_rating, df2)
# HFS painfulness included here as a covariate, because it
# was found to have a relationship with magnitude in the
# above testing.
summary(model2)  # No significant main or interaction effects.
```

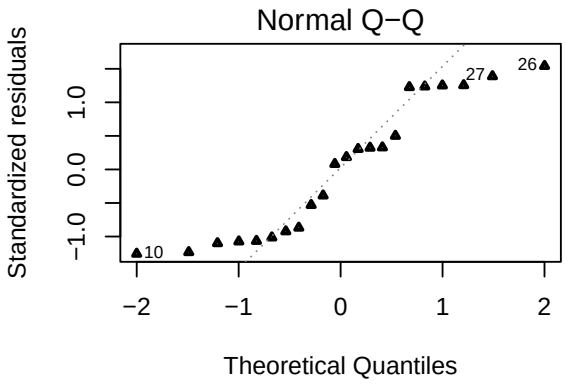
```
##
## Call:
## lm(formula = mag ~ hop_tot + pain_status + hop_tot * pain_status +
##     train_rating, data = df2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -244.03 -189.28   21.57  170.78  296.29
##
## Coefficients: (1 not defined because of singularities)
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)           -910.525    348.583  -2.612  0.01765 *
## hop_tot               1287.731    344.853   3.734  0.00152 **
## pain_statusb          -225.688    127.193  -1.774  0.09292 .
## train_rating             2.089      1.716   1.218  0.23911
## hop_tot:pain_statusb        NA         NA      NA       NA
## ---
```

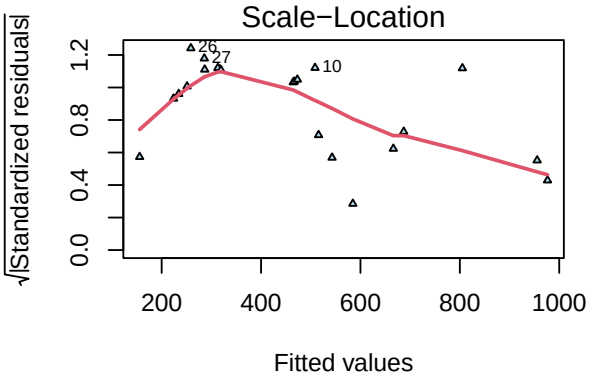
```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 201.8 on 18 degrees of freedom
##   (8 observations deleted due to missingness)
## Multiple R-squared:  0.6123, Adjusted R-squared:  0.5477
## F-statistic: 9.476 on 3 and 18 DF,  p-value: 0.0005626
```

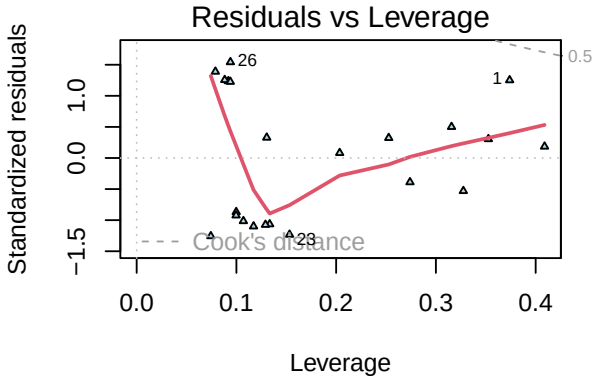
```
# suggests a relationship between the magnitude of SH and
# distress but no relationship with pain status. One unit
# increase in distress is associated with 1287.731 units
# increase in magnitude AUC (%); being in pain_status group
# b (rather than group a) is associated with -225.688-unit
# decrease in magnitude AUC (%).61% of the variation in
# magnitude is explained by this model.
```

```
# plot model assumptions
par(mfrow = c(2, 2))
plot(model2, pch = 24, bg = "lightblue", colour = "blue", cex = 0.5,
  lwd = 2)
```









Blinding assessment

Here, we want to know whether the participants and assessor remained blinded. So we will:

1. Calculate the percentage of participants who correctly guessed the aim of the study;
2. Calculate James' Blinding index;
3. Calculate the number of participants for whom the researcher correctly guessed group membership at a confidence level of 4 and 5;
    - tabulate the 'accurate' and 'inaccurate' responses for whom the researcher confidently guessed group membership.
    - conduct sensitivity analyses to examine the influence of any unblinded participants on the findings of the outcomes (where applicable).

```
# Were participants blinded to the aim of the study?
blinded <- df %>%
  select(pid, participants_blinded) %>%
  unique()

total_blinded <-
  sum(blinded$participants_blinded == 'y')
paste0(total_blinded, " of 6 participants were blinded")
```

```
## [1] "6 of 6 participants were blinded"
```

```
# number of blinded participants
percentage_blinded <-
  (total_blinded / sample.n) * 100
paste0(percentage_blinded, "% of the participants in the pilot study were blinded")
```

```
## [1] "100% of the participants in the pilot study were blinded"
```

```
# was the assessor blinded to the pain status of each participant?
assessor_blinding <- df %>%
  select(pid,
         pain_status,
         participants_blinded,
         group_guess,
         assessor_confidence) %>%
  unique()

# decode participants' pain status to show group membership
assessor_blinding %<>%
  mutate(
    pain_status = (
      case_when(pain_status == 'b' ~ 'pain',
                pain_status == 'a' ~ 'no pain')
  )) %>%
  mutate(group_guess = (
    case_when(
      group_guess == 'b' ~ 'pain',
      group_guess == 'a' ~ 'no pain')
```

```
  )) %>%
  mutate(accuracy = if_else(pain_status == group_guess,
                            'accurate',
                            'inaccurate'))

# James' BI
# there are four possible combinations to blinding outcomes,
# filter by each condition and calculate the value for each.

# condition 1: actual group (pain) vs guessed group (pain)
accurate_group_pain <- assessor_blinding %>%
  select(pid, pain_status, group_guess, accuracy) %>%
  filter(pain_status == "pain") %>%
  filter(accuracy == "accurate")
paste0("Assessor guessed correctly for ", nrow(accurate_group_pain), " participants in the pain group."
```

```
## [1] "Assessor guessed correctly for 5 participants in the pain group."
```

```
# condition 2: actual group (pain) vs guessed group (no pain)
inaccurate_group_pain <- assessor_blinding %>%
  select(pid, pain_status, group_guess, accuracy) %>%
  filter(pain_status == "pain") %>%
  filter(accuracy == "inaccurate")
paste0("Assessor guessed inaccurately for ", nrow(inaccurate_group_pain), " participants in the pain gr
```

```
## [1] "Assessor guessed inaccurately for 0 participants in the pain group."
```

```
# condition 3: actual group (no pain) vs guessed group (no pain)
accurate_group_nopain <- assessor_blinding %>%
  select(pid, pain_status, group_guess, accuracy) %>%
  filter(pain_status == "no pain") %>%
  filter(accuracy == "accurate")
paste0("Assessor guessed accurately for ", nrow(accurate_group_nopain), " participants in the no pain g
```

```
## [1] "Assessor guessed accurately for 1 participants in the no pain group."
```

```
# condition 4: actual group (no pain) vs guessed group (pain)
inaccurate_group_nopain <- assessor_blinding %>%
  select(pid, pain_status, group_guess, accuracy) %>%
  filter(pain_status == "no pain") %>%
  filter(accuracy == "inaccurate")
paste0("Assessor guessed accurately for ", nrow(inaccurate_group_nopain), " participants in the no pain
```

```
## [1] "Assessor guessed accurately for 0 participants in the no pain group."
```

```
# Generate table based on the above conditions
table_blind_con <- matrix(
  data = c(
    nrow(accurate_group_pain),
    nrow(inaccurate_group_pain),
```

```
    nrow(inaccurate_group_nopain),
    0, # DK
    nrow(accurate_group_nopain),
    0),# DK
  nrow = 3,
  ncol = 2,
  dimnames = list(c('pain', 'no pain', 'DK'), # DK = Don't know group = 0
                  c('pain', 'no pain')))

table_blind_con %>%
  kbl(caption = "Researcher blinding to group membership", booktabs = T) %>%
  kable_styling(latex_options = c("striped", "HOLD_position"))
```

Table 5: Researcher blinding to group membership

	pain	no pain
pain	5	0
no pain	0	1
DK	0	0

```
BI(table_blind_con)$JamesBI
```

```
##         Estimate Std. Error 95% LCL (2-Sided) 95% UCL (2-Sided)
## Overall        0          0                 0                 0
```

```
# BI= 0 and Upper confidence level = 0 which is less than 0.5.
# therefore, blinding of the researcher was unsuccessful.

# Chi-square goodness of fit
# Requires separate testing of each group (i.e., pain and no pain)

## pain group
chisq.test(x = c(5, 0), # Counts in guessed'pain' group
           p = c(0.5, 0.5)) # Probabilities assuming random guessing.
```

```
##
##  Chi-squared test for given probabilities
##
## data:  c(5, 0)
## X-squared = 5, df = 1, p-value = 0.02535
```

```
# X2 = 5, critical value lies between 0.05(3.841) and 0.02(5.412)
# therefore, it is likely that the distribution is true

## no pain group
chisq.test(x = c(0, 1),  # Counts in guessed 'no pain' group
           p = c(0.5, 0.5))
```

```
##
```

```
##  Chi-squared test for given probabilities
##
## data:  c(0, 1)
## X-squared = 1, df = 1, p-value = 0.3173
```

```
# X2 = 1, critical value probability lies between 0.10 (2.706) and 0.5 (3.841)
# therefore, it is unlikely that the distribution is true

# Two groups, therefore df = 1. With df = 1, the X2 value must exceed critical value
# to reach statistical significance. If the X2 value is less than this, p will be > 0.05,
# indicating that the distribution does not significantly deviate from the distribution
# that would be expected given chance guessing (which would be ~50% accuracy).

# determine the number of accurate responses
accurate.guess <-
  sum(assessor_blinding$accuracy == 'accurate')
paste0("the assessor accurately guessed the group membership of ", accurate.guess, " out of 6 participa
```

```
## [1] "the assessor accurately guessed the group membership of 6 out of 6 participants"
```

```
# there is a 50% chance that Researcher 2 can accurately guess the participant's group membership;
# however, this does not necessarily demonstrate blinding failure.
# therefore, we will quantify the accurate guesses for which the researcher
# reported a confidence interval of 4 (i.e. confident) and 5 (i.e. extremely confident)

# filter by accurate guesses and tabulate the guesses for which
# the assessor was "confident" and "extremely confident"
assessor_blinding %>%
  select(pid, group_guess, assessor_confidence, accuracy) %>%
  filter(accuracy == "accurate") %>%
  filter(assessor_confidence >= "4") %>%
kbl(caption = "Participants for whom the assessor accurately and confidently guessed
    group membership",
      booktabs = T) %>% kable_styling(latex_options = c("striped",
                                                        "HOLD_position"))
```

Table 6: Participants for whom the assessor accurately and confidently guessed group membership

pid	group_guess	assessor_confidence	accuracy
UKO-996	pain	4	accurate

```
# Researcher 2 accurately guessed group membership for all participants,

# filter by inaccurate guesses and tabulate the guesses for which
# the researcher was "confident"  (4) and "extremely confident" (5)
inaccurate_guess <- assessor_blinding %>%
  select(pid, group_guess, assessor_confidence, accuracy) %>%
  filter(accuracy == "inaccurate") %>%
  filter(assessor_confidence >= "4")

assessor_blinding
```

```
##       pid pain_status participants_blinded group_guess assessor_confidence
## 1 UKO-991        pain                    y        pain                   3
## 2 UKO-992        pain                    y        pain                   3
## 3 UKO-993        pain                    y        pain                   1
## 4 UKO-994     no pain                    y     no pain                   3
## 5 UKO-995        pain                    y        pain                   3
## 6 UKO-996        pain                    y        pain                   4
##   accuracy
## 1 accurate
## 2 accurate
## 3 accurate
## 4 accurate
## 5 accurate
## 6 accurate
```

```
# note the data frame does not have inaccurate responses; therefore, we have no
# inaccurate guesses for whom the researcher gave a confidence level of 4 or 5
as.data.frame(inaccurate_guess)
```

```
## [1] pid                 group_guess         assessor_confidence
## [4] accuracy
## <0 rows> (or 0-length row.names)
```

```
# tabulate inaccurate guesses
inaccurate_guess %>%
kbl(caption = "Participants for whom the assessor inaccurately but confidently guessed group membership
      booktabs = T) %>%
  kable_styling(latex_options = c("striped","HOLD_position"))
```

Table 7: Participants for whom the assessor inaccurately but confidently guessed group membership

pid	group_guess	assessor_confidence	accuracy

Formal data analysis

Luyanduthando Mqadi, Tory Madden, Peter Kamerman, and Gill Bedwell

22 November, 2022

Contents

Research question

Does distress predict experimentally induced secondary hyperalgesia (SH) in people living with HIV?

Background

In this script, we test whether the severity of self-reported psychosocial distress (PSD) is associated with experimentally induced SH in people living with HIV who report persistent pain or are pain-free. We will also test whether the association between PSD and induced SH differs in people with HIV who report persistent pain.

Research objectives

In people living with HIV who reported persistent pain or being pain-free, we:

- Assessed PSD;
- Induced SH;
- Assessed whether PSD was associated with the surface area of experimentally induced SH ***(primary analysis)*** and explored if this relationship was different between groups;
- Assessed whether PSD was associated with the magnitude of experimentally induced SH ***(secondary analysis)*** and explored if this relationship was different between groups.
- Assessed whether the surface area and magnitude of SH differed between groups.

Hypotheses

We hypothesised that both surface area and magnitude of experimentally induced SH will be positively associated with self-reported PSD, and that the persistent pain group will have greater surface area and magnitude of experimentally induced SH than the group without pain.

Load packages

Import data

Import master data list, with the following list structure:

1. Procedure details
2. Questionnaires data
3. Rating of HFS trains
4. Magnitude ratings
5. Surface area data

```r
## Import master data list
master_data <- readRDS('Lu_Mqadi/formal_analysis/sh-master.rds')

## prepare age and PSD data
distress_data <- data.frame(master_data[[2]]) %>%
  arrange(blinded_id) %>%
  mutate(blinded_id = as.character(blinded_id)) %>%
  dplyr::select(-hop_complete) %>%
  rename(pid = blinded_id,
         group = group_reported) %>%
# group reported is used as the grouping variable, but conflict between group assignment and
# pain status reported is noted in group_conflict.
  mutate(distress_tot = rowMeans(dplyr::select(., starts_with("hop")))) # compute Hopkins total

## extract additional data on the arm used for HFS, electrical current and binding
proc_details <- data.frame(master_data[[1]]) %>%
  arrange(blinded_id) %>%
  mutate(blinded_id = as.character(blinded_id)) %>%
  rename(pid = blinded_id,
         participants_blinded = blinded,
         assessor_group_guess = group_guess) %>%
  dplyr::select(
    pid,
    arm,
    current,
    participants_blinded,
    assessor_group_guess,
    assessor_confidence,
    participation_status
  ) %>%
  mutate(calibrated_current = (current / 100)) %>%
  dplyr::select(-current) %>%
  unique()

## pull participation status info
partic <- proc_details %>%
  mutate(pid = as.character(pid)) %>%
  dplyr::select(pid,
                participation_status) %>%
  mutate(participation_status = case_when(
      pid == "521" ~ "ineligible: screened into group b but was NOT ELIGIBLE at T0",
      TRUE ~ participation_status
    ))

## now create a new data frame for all the imported data
began_procedure <- distress_data %>%
  mutate(pid = as.character(pid)) %>%
  left_join(proc_details) %>%
  dplyr::select(
    pid,
    age,
    sex,
    group,
```

```
    group_conflict,
    distress_tot,
    arm,
    participants_blinded,
    assessor_group_guess,
    assessor_confidence,
    calibrated_current
  ) %>%
  right_join(partic)
```

```
## Joining, by = "pid"
## Joining, by = "pid"
```

```
## check for available data
data_available <-
  began_procedure %>%
  mutate(current_data = case_when(
    grepl("completed", participation_status) ~ "yes",
    grepl("missing", participation_status) ~ "yes",
    grepl("withdrew", participation_status) ~ "yes",
    grepl("excluded", participation_status) ~ "no",
    grepl("ineligible", participation_status) ~ "no"
  )) %>%
  mutate(sh_ratings = case_when(
    grepl("completed", participation_status) ~ "yes",
    grepl("missing post-procedure data", participation_status) ~ "yes",
    grepl("procedure aborted", participation_status) ~ "no",
    grepl("withdrew", participation_status) ~ "no",
    grepl("excluded", participation_status) ~ "no",
    grepl("ineligible", participation_status) ~ "no",
    grepl("ineligible", participation_status) ~ "no"
  )) %>%
  mutate(sa_data = case_when(
    grepl("completed", participation_status) ~ "yes",
    grepl("missing post-procedure data", participation_status) ~ "yes",
    grepl("procedure aborted", participation_status) ~ "no",
    grepl("withdrew", participation_status) ~ "no",
    grepl("excluded", participation_status) ~ "no",
    grepl("ineligible", participation_status) ~ "no"
  )) %>%
  mutate(post_proc_data = case_when(
    grepl("completed", participation_status) ~ "yes",
    grepl("missing post-procedure data", participation_status) ~ "no",
    grepl("procedure aborted", participation_status) ~ "no",
    grepl("withdrew", participation_status) ~ "no",
    grepl("excluded", participation_status) ~ "no",
    grepl("ineligible", participation_status) ~ "no"
  )) %>%
  dplyr::select(pid, current_data, sh_ratings, sa_data, post_proc_data)

rm(proc_details, partic)
```

Demographic characteristics

Participants' demographic information was separated into those (1) enrolled and those who (2) were taken for analysis. In each case, we identified the sex and the total number of participants in (1) and (2) and summarised descriptive characteristics in a table. Data carried forward to analysis will include participants who completed the study and those with missing post-procedure data.

```
## number of participants enrolled
sample.n <- length(began_procedure$pid)
glue("{sample.n} participants enrolled")
```

63 participants enrolled

```
## sex of enrolled participants
males <- began_procedure %>%
  filter(sex == "male") %>%
  nrow()
glue("{males} male participants enrolled")  # males
```

17 male participants enrolled

```
females <- began_procedure %>%
  filter(sex == "female") %>%
  nrow()
glue("{females} female participants enrolled")  # females
```

46 female participants enrolled

```
## participants excluded from analysis with reasons (1.
## withdrew, 2. excluded, 3. ineligible)

## participants who withdrew
withdrew.n <- began_procedure %>%
  filter(grepl("withdrew", participation_status)) %>%
  nrow()
glue("{withdrew.n} participants withdrew - see Table 1 below")
```

7 participants withdrew - see Table 1 below

```
## participants excluded
excluded.n <- began_procedure %>%
  filter(grepl("excluded", participation_status)) %>%
  nrow()
glue("{excluded.n} participants excluded - see Table 1 below")
```

4 participants excluded - see Table 1 below

```
## Ineligible
ineligible_n <- began_procedure %>%
  filter(grepl("ineligible", participation_status)) %>%
  nrow()
glue("{ineligible_n} participants ineligible")
```

1 participants ineligible

```
began_procedure %>%
  filter(grepl("excluded", participation_status) | grepl("withdrew",
    participation_status) | grepl("ineligible", participation_status)) %>%
  dplyr::select(pid, sex, group, participation_status) %>%
  kbl(caption = "Reasons for excluding participants from analysis (N = 12)",
    booktabs = T) %>%
  kable_styling(latex_options = c("striped", "HOLD_position",
    "scale_down"))
```

Table 1: Reasons for excluding participants from analysis (N = 12)

pid	sex	group	participation_status
139	female	a	withdrew: hfs too painful
200	male	a	excluded: procedure terminated
207	female	a	excluded: technical error
225	female	b	withdrew: hfs too painful
229	female	a	withdrew: hfs too painful
244	female	b	withdrew: hfs too painful
252	female	b	withdrew: hfs too painful
283	female	a	excluded: no 7 day delay between UKO studies
333	female	a	withdrew: hfs too painful
369	female	b	withdrew: hfs too painful
521	male	a	ineligible: screened into group b but was NOT ELIGIBLE at T0
568	male	b	excluded: technical error

```
## participants taken to analysis
analysed.n <- began_procedure %>%
  filter(grepl("completed", participation_status) | grepl("missing post-procedure data",
    participation_status)) %>%
  nrow()
glue("{analysed.n} participants data taken to analysis")
```

51 participants data taken to analysis

```
## create data frame with only participants whose data are
## being carried forward to analysis
analyse <- began_procedure %>%
  filter(grepl("completed", participation_status) | grepl("missing post-procedure data",
    participation_status))

## summarise demographic characteristics for all
## participants who began the procedure and round off to 0
## or 1 decimal place where necessary
began_proc_controls <- arsenal::tableby.control(numeric.stats = c("N",
  "mean", "sd", "median", "range"), cat_stats = c("countpct"),
  stats.labels = c(list(N = "number of participants", mean = "mean",
```

```
    standard_deviation = "sd", median = "median", range = "range")))
table_one <- tableby(group ~ sex + age + calibrated_current +
  distress_tot, data = began_procedure, control = began_proc_controls)
summary(table_one, text = TRUE) %>%
  kbl(caption = "Demographic characteristics of participants who began procedure",
    booktabs = T) %>%
  kable_styling(latex_options = c("striped", "HOLD_position"))  #%>%
```

Table 2: Demographic characteristics of participants who began procedure

	a (N=36)	b (N=27)	Total (N=63)	p value
sex				0.682
- female	27 (75.0%)	19 (70.4%)	46 (73.0%)	
- male	9 (25.0%)	8 (29.6%)	17 (27.0%)	
age				0.783
- number of participants	36	27	63	
- mean	44.083	43.444	43.810	
- SD	9.425	8.622	9.023	
- median	44.000	43.000	43.000	
- range	27.000 - 64.000	30.000 - 63.000	27.000 - 64.000	
calibrated_current				0.844
- number of participants	36	27	63	
- mean	0.141	0.138	0.140	
- SD	0.056	0.066	0.060	
- median	0.155	0.140	0.150	
- range	0.040 - 0.260	0.040 - 0.270	0.040 - 0.270	
distress_tot				< 0.001
- number of participants	36	27	63	
- mean	1.478	1.963	1.686	
- SD	0.449	0.634	0.584	
- median	1.360	2.000	1.480	
- range	1.000 - 2.520	1.080 - 3.360	1.000 - 3.360	

```
# save_kable('Tables/Demographic characteristics of
# participants who began procedure.png') # function wont
# knit to latex

## summarise demographic characteristics for all
## participants taken to analyses and round off to 0 or 1
## decimal place where necessary
analyse_controls <- arsenal::tableby.control(numeric.stats = c("N",
  "mean", "sd", "median", "range"), cat_stats = c("countpct"),
  stats.labels = c(list(N = "number of participants", mean = "mean",
    standard_deviation = "sd", median = "median", range = "range")))
table_two <- tableby(group ~ sex + age + calibrated_current +
  distress_tot, data = analyse, control = analyse_controls)
summary(table_two, text = TRUE) %>%
  kbl(caption = "Demographic characteristics of participants taken to analysis",
    booktabs = T) %>%
```

```
kable_styling(latex_options = c("striped", "HOLD_position"))
```

Table 3: Demographic characteristics of participants taken to analysis

	a (N=29)	b (N=22)	Total (N=51)	p value
sex				0.543
- female	22 (75.9%)	15 (68.2%)	37 (72.5%)	
- male	7 (24.1%)	7 (31.8%)	14 (27.5%)	
age				0.794
- number of participants	29	22	51	
- mean	43.103	42.455	42.824	
- SD	9.116	8.245	8.671	
- median	41.000	42.000	42.000	
- range	31.000 - 64.000	30.000 - 58.000	30.000 - 64.000	
calibrated_current				0.509
- number of participants	29	22	51	
- mean	0.140	0.128	0.135	
- SD	0.058	0.069	0.062	
- median	0.150	0.120	0.130	
- range	0.040 - 0.260	0.040 - 0.270	0.040 - 0.270	
distress_tot				< 0.001
- number of participants	29	22	51	
- mean	1.439	2.085	1.718	
- SD	0.447	0.628	0.618	
- median	1.360	2.100	1.480	
- range	1.000 - 2.520	1.080 - 3.360	1.000 - 3.360	

Summary of demographic characteristics:

Sixty-three participants (17 males; 46 females) were enrolled in this study and screened for eligibility. Participants could be excluded before or after participation, according to the eligibility criteria. Three participants were excluded before the analysis because of data unavailability resulting from technical errors (n=2), and the procedure was aborted (n =1). Two participants were regrettably included in the study despite reporting epilepsy but given that the epilepsy exclusion criterion was motivated by participant safety, rather than by anticipated effect on the findings, the data from these participants were retained. Two participants were excluded after participation because (i) the arm for SH induction had also been used for venepuncture without a 7-day washout period (n=1) and (ii) the participant was screened into group 'b' but was ineligible to participate (n=1). Seven participants withdrew from the study because the SH induction was too painful. We obtained complete datasets, however, blinding assessment data were lost for four participant due to technical errors. In the protocol we stated that we would recruit 60 participant (30 in each group), however, due to multiple errors of inclusion we end up with fifty-one participants (14 males, 37 females) aged 30 - 64 years (mean (SD) = 42.82 (8.7)) who completed the study. Here, we report only the data from these 51 participants, except for an exploratory analysis to check whether PSD differed between those who did and did not withdraw (see Exploratory analysis 2: withdrawal status). All participants reported that they had not taken pain medication in the 24 hours prior to the induction off SH.

SH data preparation

Surface area of SH

Here we imported surface area data and calculated the total surface area at 30, 45, and 60 minutes after the induction of SH. Next, we checked the distribution of these data using a histogram and visualised it in plots showing group and individual comparisons. After this, we calculated the area under the line (AUL) formed from the total surface area at the 3 designated time points. We used the **"Area Under the Curve"** function to calculate the AUL for each participant and used the trapezoidal integration approach as our method **(Yeh, 2002)**.

```
## available surface area data
available_sa <- data_available %>%
  filter(sa_data == "yes") %>% nrow()
glue("There area {available_sa} participants with surface area data available")
```

```
## There area 51 participants with surface area data available
```

```
 ## import surface area data
raw_sa <- data.frame(master_data[[5]]) %>%
  rename(pid = blinded_id,
         radial_line = modality) %>%
  mutate(pid = as.character(pid)) %>%
  dplyr::select(pid, date, time_point, radial_line, change_point) %>%
  separate(col = time_point,
           sep = '_',
           into = c('delete', 'time')) %>%
  mutate(time = as.numeric(time)) %>%
  dplyr::select(-delete) %>%
  dplyr::rename(length_1 = change_point) %>%
  group_by(pid, time) %>%
  mutate(
    length_2 = case_when
    (
      radial_line == 'A' ~ lead(length_1, 1),
      radial_line == 'B' ~ lead(length_1, 1),
      radial_line == 'C' ~ lead(length_1, 1),
      radial_line == 'D' ~ lead(length_1, 1),
      radial_line == 'E' ~ lead(length_1, 1),
      radial_line == 'F' ~ lead(length_1, 1),
      radial_line == 'G' ~ lead(length_1, 1),
      radial_line == 'H' ~ lag(length_1, 7))) %>%
  filter(!pid == 521) %>%
  ungroup()

## calculate surface area for each triangle
sin_angle <- 45*(3.141593/180)

# convert radians to degrees. Note, sin(x) function takes radians as input
# and not degrees
raw_sa %<>% mutate(triangle_sa = (length_1 * length_2 * sin_angle) / 2)

## calculate total surface area of SH at t 30, 45, and  60 minutes after the induction
```

```
raw_sa <- raw_sa %>%
  dplyr::select(-radial_line, -length_1, -length_2) %>%
  dplyr::group_by(pid, time) %>%
  dplyr::summarise(area = sum(triangle_sa)) %>% ungroup() %>%
  rename(time_sa = time) %>% distinct()
```

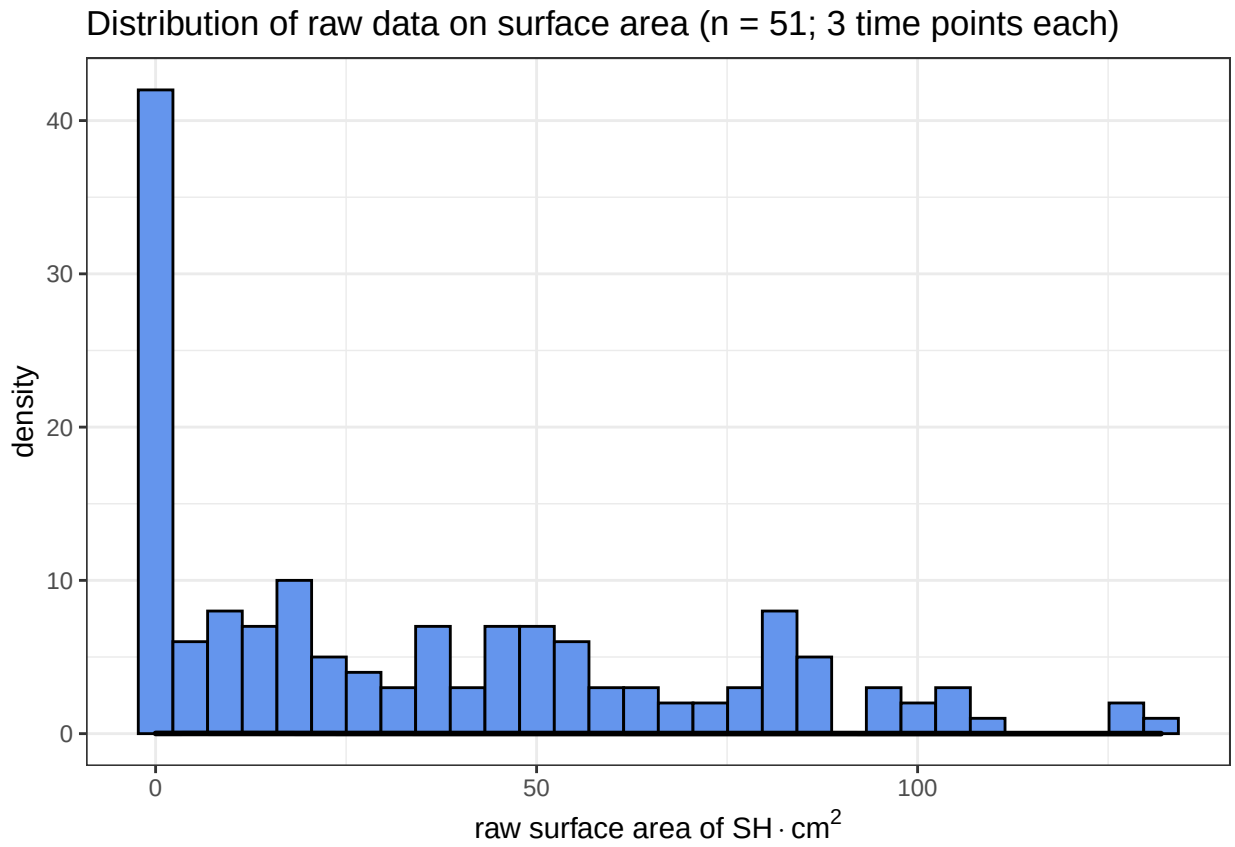
```
## `summarise()` has grouped output by 'pid'. You can override using the `.groups`
## argument.
```

```
raw_sa %<>%
  right_join(began_procedure) %>%
  filter(
    grepl("completed", participation_status) |
      grepl("missing post-procedure data", participation_status)) %>%
  dplyr::select(pid, group, time_sa, area, participation_status) %>%
  unique()
```

```
## Joining, by = "pid"
```

```
## check raw data distribution using histogram
ggplot(raw_sa) + aes(x = area) +
  geom_histogram(fill = "#6495ed", col = "black") + geom_density(lwd= 1) +
  labs(title =  "Distribution of raw data on surface area (n = 51; 3 time points each)",
       x = expression("raw surface area of SH" %.% cm ^ {2})) + theme_bw()
```

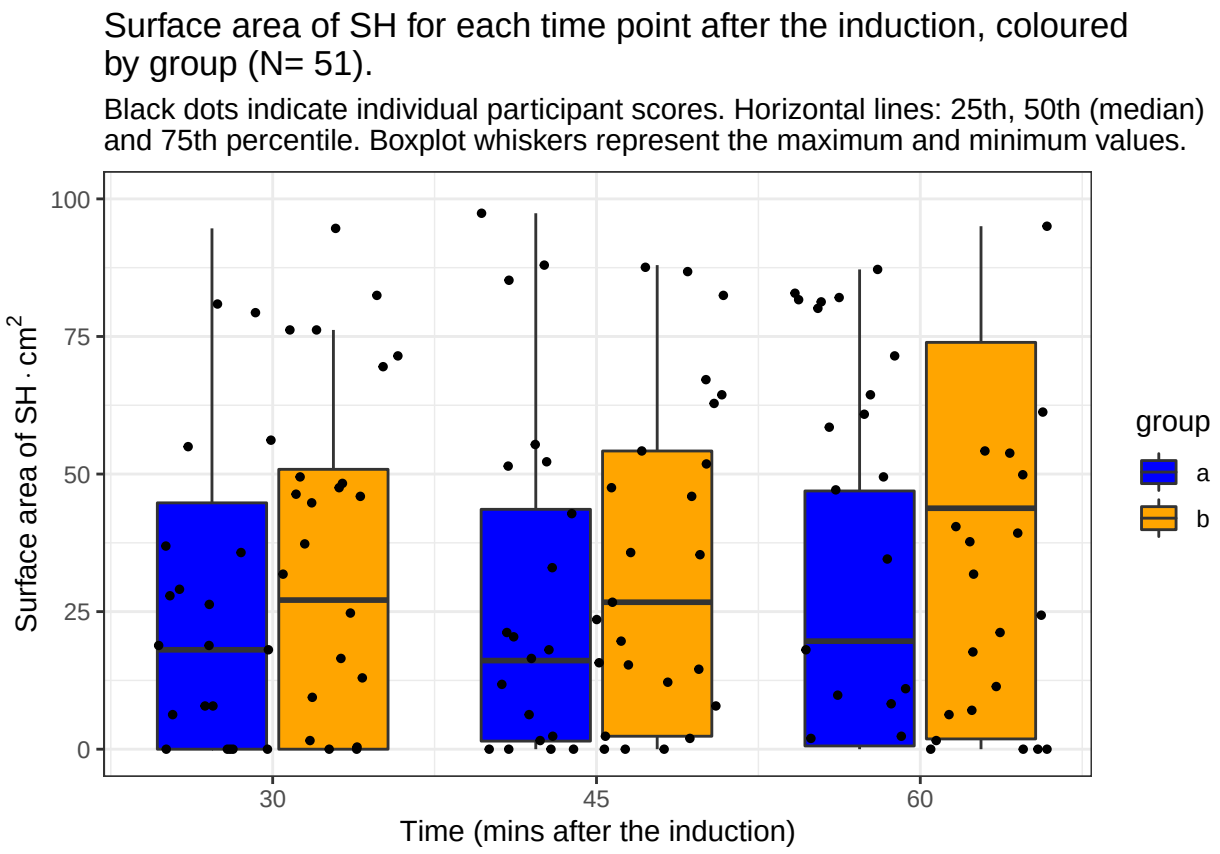
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



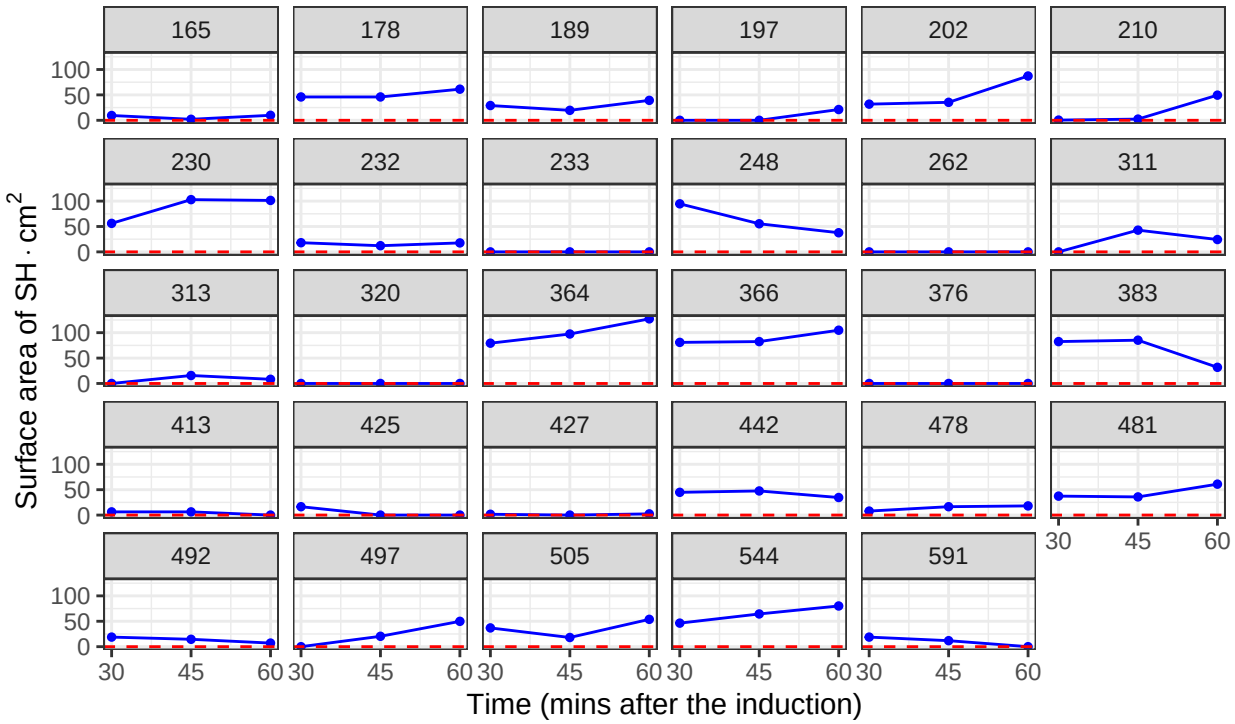

```
## plot raw surface area group ratings over time
sa_over_time <- ggplot(data = raw_sa) + aes(
  x = time_sa,
  y = area,
  group = interaction(time_sa, group),
  fill = group) +
  geom_boxplot(outlier.shape = NA) +
  geom_jitter(
    shape = 21,
    color = "black",
    fill = "black",
    size = 1) +
  labs(title =
"Surface area of SH for each time point after the induction, coloured\nby group (N= 51).",
subtitle = "Black dots indicate individual participant scores. Horizontal lines: 25th, 50th (median)\na
      x = "Time (mins after the induction)",
      y = expression("Surface area of SH" %.% cm ^ {2})) +
  scale_fill_manual(values = c('blue', 'orange'))+
  scale_y_continuous(limits = c(0, 100)) +
  scale_x_continuous(breaks = c(30, 45, 60)) +
  theme_bw()

print(sa_over_time)
```

```
## Warning: Removed 9 rows containing non-finite values (stat_boxplot).
```

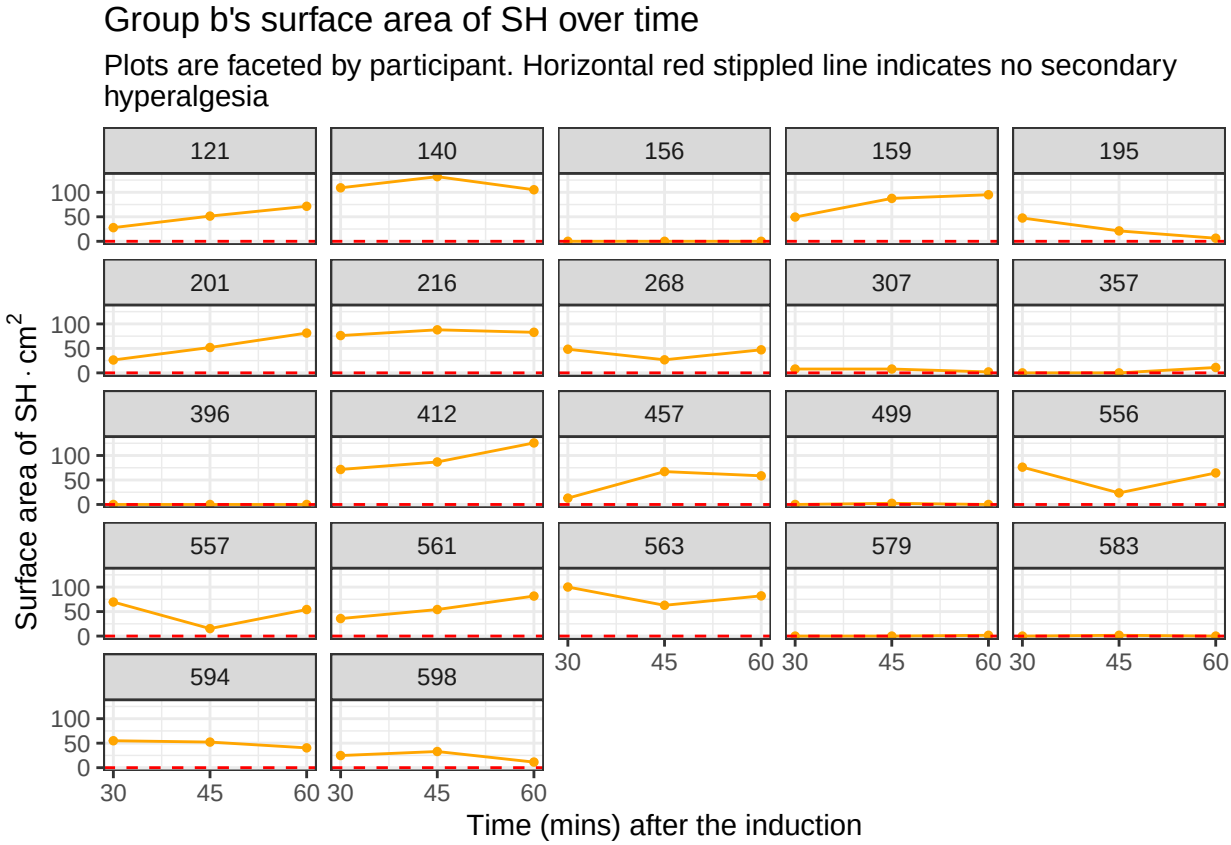
```
## Warning: Removed 27 rows containing missing values (geom_point).
```




```
## plot surface area by individual over time
### plot group a
raw_sa %>%
  filter(group == "a") %>%
  ggplot(.) +
  aes(x = time_sa,
      y = area,
      fill = group) +
  geom_point(size = 1, shape = 21, colour = "blue", fill = "blue") +   geom_line(colour = "blue") +
  facet_wrap(~ pid, nrow = 5) +
  geom_hline(yintercept = 0, linetype = 2, colour = "red") +
  labs(title = "Group a's surface area of SH over time",
       subtitle = "Plots are faceted by participant. Horizontal red stippled line indicates no secondary
       x = "Time (mins after the induction)",
       y = expression("Surface area of SH" %.%
                        cm ^ {2})) +
  scale_x_continuous(breaks = c(30, 45, 60)) +
  theme_bw()
```




```
### plot group b
raw_sa %>%
  filter(group == "b") %>%
  ggplot(.) +
  aes(x = time_sa,
      y = area) +
  geom_point(
    size = 1,
    shape = 21,
    colour = "orange",
    fill = "orange"
  ) +
  geom_line(colour = "orange") +
  facet_wrap( ~ pid, nrow = 5) +
  geom_hline(yintercept = 0,
             linetype = 2,
             colour = "red") +
  labs(
    title = "Group b's surface area of SH over time" ,
    subtitle = "Plots are faceted by participant. Horizontal red stippled line indicates no secondary\nl
    x = "Time (mins) after the induction",
    y = expression("Surface area of SH" %.% cm ^ {2})) +
  scale_x_continuous(breaks = c(30, 45, 60)) +
  theme_bw()
```




```
## calculate AUL using the AUC function
aul_sa_df <- raw_sa %>%
  dplyr::group_by(pid) %>%
  mutate(aul_sa = area_under_curve(time_sa,
                                   area,
                                   method = "trapezoid")) %>% # using the trapezoidal integration rule
  dplyr::select(-time_sa,-area) %>%
  distinct()
```

We have 51 participants with complete surface area data. The histogram plot shows that our surface area data are not normally distributed. In our sample, 76% of participants developed an identifiable surface area of SH (group a: 83%; group b: 68%). The boxplot suggests no noteworthy differences in surface area, over time, between the groups. The surface area AUL has been calculated and carried forward to analysis (see **Primary analysis**).

Magnitude of SH

Here we imported all rating data and filtered by the ratings given for the two pinprick weights (128 mN and 256 mN). We then calculated the mean ratings given for trials using either of the two pinprick weights at each time point. We expressed the follow-up ratings (35, 50, and 65 minutes after the induction of SH) as a change of the baseline ratings. The change from baseline was calculated by subtracting the individual follow-up rating by the baseline rating. We checked the distribution of these data using a histogram and visualised it in plots showing group and individual comparisons. Similarly, the AUL of the magnitude-relative-to-baseline over the three follow-up time points was calculated using the trapezoidal integration approach.

```r
## available magnitude of SH data
available_mag <-
  data_available %>% filter(sh_ratings == "yes") %>% nrow()
glue("There are {available_mag} participants with magnitude data")
```

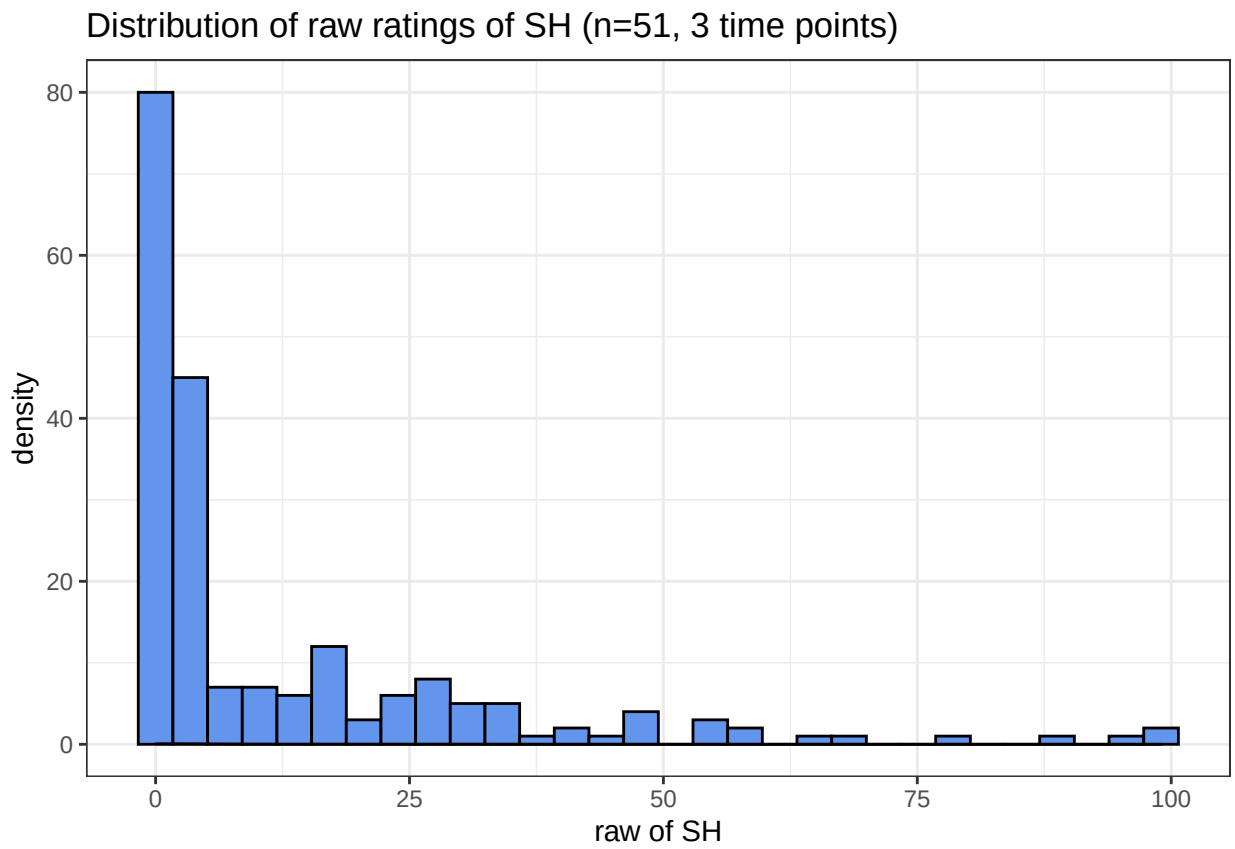
```
## There are 51 participants with magnitude data
```

```r
## import magnitude ratings
raw_mag_ratings_1 <- data.frame(master_data[[4]]) %>%
  rename(pid = blinded_id) %>%
  mutate(pid = as.character(pid)) %>%
  dplyr::select(pid, time_point, modality, rating) %>%
  filter(time_point %in% c("baseline_point", "time_35", "time_50", "time_65")) %>%
  filter(modality %in% c("pp_128", "pp_256")) %>%
  group_by(pid, time_point, modality) %>%
  mutate(rating = mean(rating)) %>% # calculate mean for baseline
  unique() %>%
  ungroup()

raw_mag_ratings <-
  raw_mag_ratings_1  %>% group_by(pid, time_point) %>%
  pivot_wider(names_from = modality,
              values_from = rating) %>%
  ungroup() %>%
  rowwise %>%
  mutate(mag = mean(c_across(contains("pp_")))) %>% # calculate the mean for the two pinprick weights
  dplyr::select(pid,
                time_point, mag) %>%
  separate(col = time_point,
           sep = '_',
           into = c('delete', 'time')) %>%
  dplyr::select(-delete) %>% # delete useless column
  mutate(time = ifelse(time == 'point', '0', time)) %>%
  mutate(time = as.numeric(time)) %>%
  rename(time_mag = time)

raw_mag_ratings %<>%
  right_join(began_procedure) %>% filter(
    grepl("completed", participation_status) |
      grepl("missing post-procedure data", participation_status)) %>%
  filter(!pid == 521) %>% #ineligible
  dplyr::select(pid, group, time_mag, mag, participation_status) %>%
  unique()
```

```
## Joining, by = "pid"
```

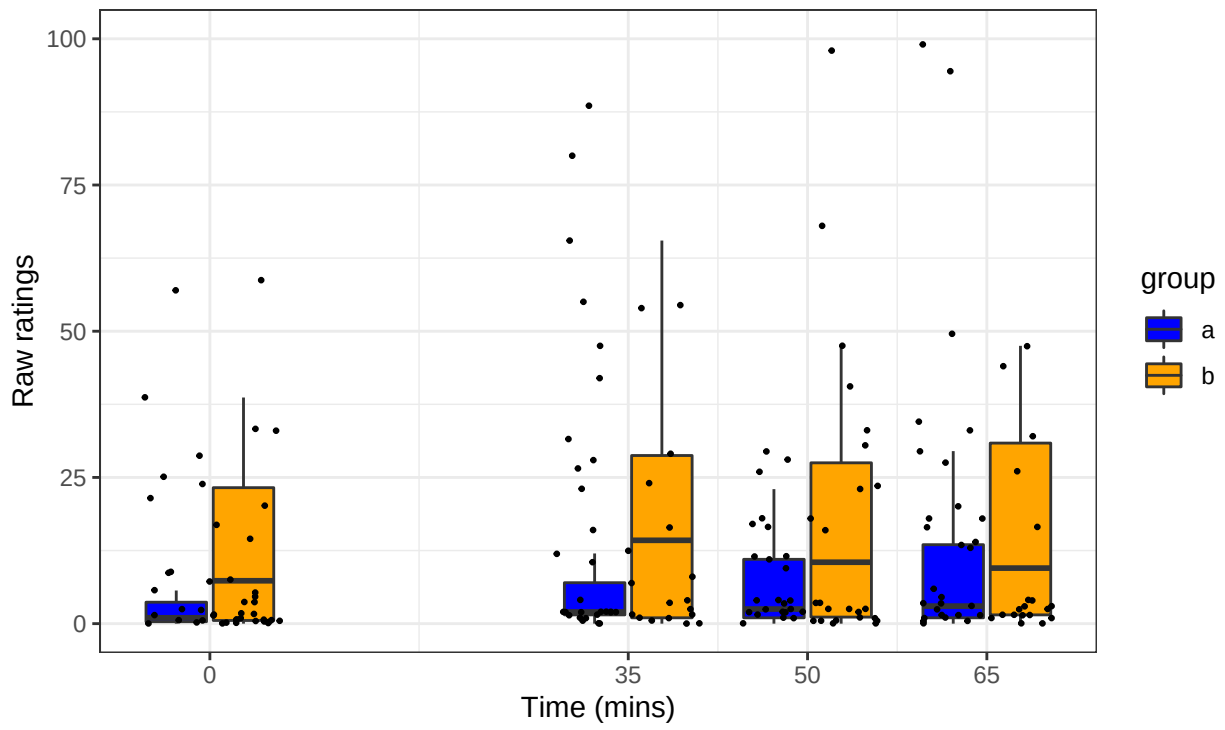
```r
ggplot(raw_mag_ratings) + aes(x = mag) +
  geom_histogram(fill = "#6495ed", col = "black") + geom_density() +
  labs(title =  "Distribution of raw ratings of SH (n=51, 3 time points)",
       x = "raw of SH") + theme_bw()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

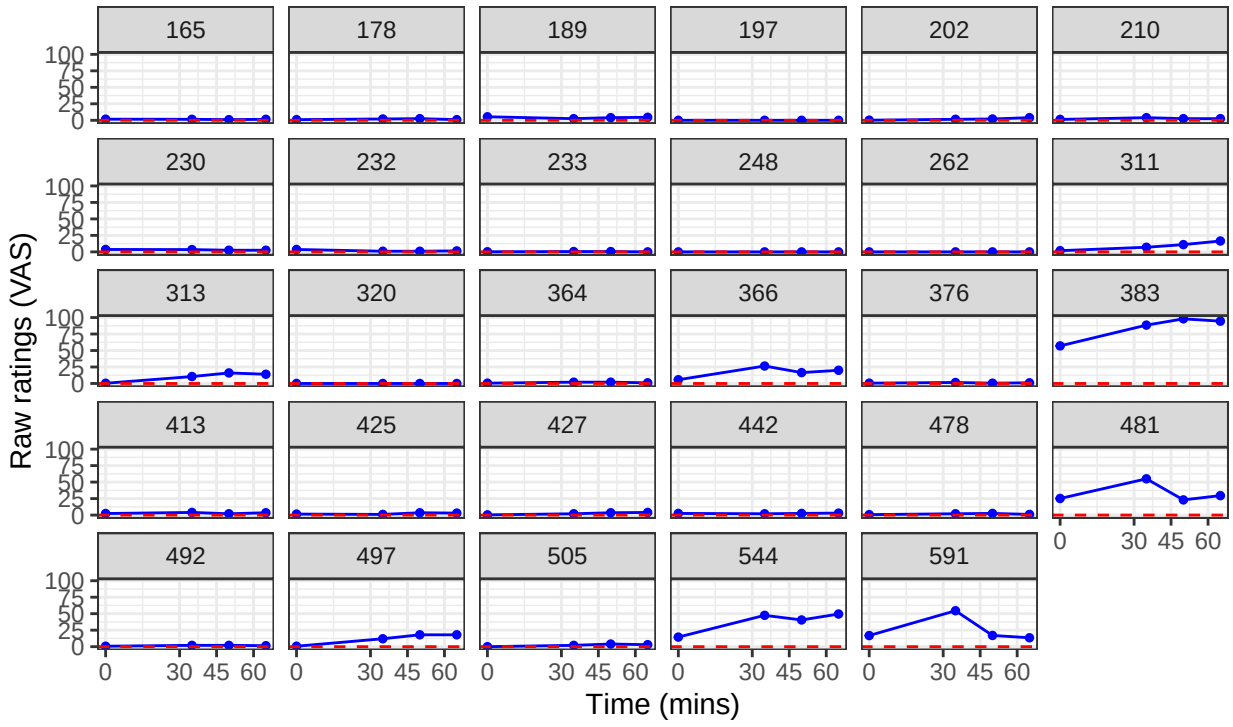



```
## plot magnitude ratings of SH by group over time
ggplot(data = raw_mag_ratings) + aes(
  x = time_mag,
  y = mag,
  group = interaction(time_mag, group),
  fill = group) +
  geom_boxplot(outlier.shape = NA) +
  geom_jitter(
    shape = 21,
    size = 0.5,
    fill = 'black') +
  labs(
    title =
      "Raw ratings of SH over time, coloured by group (N= 51)",
    subtitle = "Black dots indicate individual participant scores. Horizontal lines: 25th, 50th (median)
    x = "Time (mins)",
    y = "Raw ratings") +
  scale_fill_manual(values = c('blue', 'orange'))+
  scale_y_continuous(limits = c(0, 100)) +
  scale_x_continuous(breaks = c(0, 35, 50, 65)) +
  theme_bw()
```

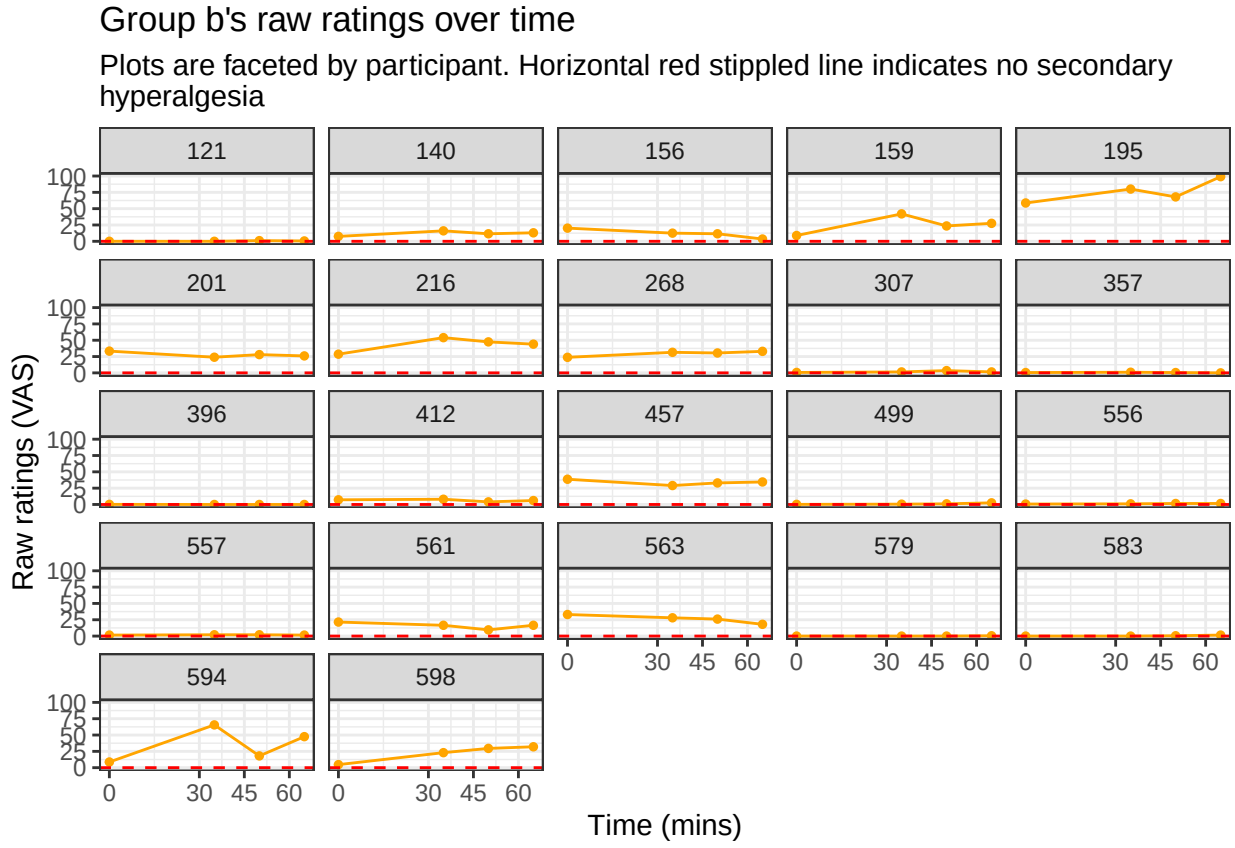
```
## Warning: Removed 17 rows containing missing values (geom_point).
```




```
## plot magnitude ratings by individual over time
## plot group a
raw_mag_ratings %>%
  filter(group == "a") %>%
  ggplot(.) +
  aes(x = time_mag,
      y = mag) +
  geom_point(
    size = 1,
    shape = 21,
    colour = "blue",
    fill = "blue"
  ) +
  geom_line(colour = "blue") +
  geom_hline(yintercept = 0,
             linetype = 2,
             colour = "red") +
  facet_wrap( ~ pid, nrow = 5) +
  labs(
    title = "Group a's raw ratings over time",
    subtitle = "Plots are faceted by participant. Horizontal red stippled line indicates no secondary\nl
    x = "Time (mins)",
    y = "Raw ratings (VAS)"
  ) +
  scale_x_continuous(breaks = c(0, 30, 45, 60)) + theme_bw()
```




```
## plot group b
raw_mag_ratings %>%
  filter(group == "b") %>%
  ggplot(.) +
  aes(x = time_mag,
      y = mag) +
  geom_point(
    size = 1,
    shape = 21,
    colour = "orange",
    fill = "orange"
  ) +
  geom_line(colour = "orange") +
  geom_hline(yintercept = 0,
             linetype = 2,
             colour = "red") +
  facet_wrap( ~ pid, nrow = 5) +
  labs(
    title = "Group b's raw ratings over time" ,
    subtitle = "Plots are faceted by participant. Horizontal red stippled line indicates no secondary\nl
    x = "Time (mins)",
    y = "Raw ratings (VAS)"
  ) +
  scale_x_continuous(breaks = c(0, 30, 45, 60)) + theme_bw()
```




```
## magnitude of SH expressed as change of baseline
mag_change_bl <- raw_mag_ratings_1 %>%
  group_by(pid, modality) %>%
# group and pivot to express follow-up ratings as change from baseline rating
  pivot_wider(names_from = time_point,
              values_from = rating) %>%
  mutate(
    change_35 = ((time_35-baseline_point)),
    change_50 = ((time_50-baseline_point)),
    change_65 = ((time_65-baseline_point))
  ) %>%
  dplyr::select(pid, modality,
                contains("change")) %>%
  group_by(pid, modality) %>%
  pivot_longer(cols = starts_with("change"),
               names_to = "time_point",
               values_to = "sh_mag") %>%
  group_by(pid,
           time_point) %>%
  pivot_wider(names_from = modality,
              values_from = sh_mag) %>%
  ungroup() %>%
  rowwise %>%
  mutate(mag = mean(c_across(contains("pp_")))) %>%
  # find the mean for the two pinprick weights
  dplyr::select(pid,
```

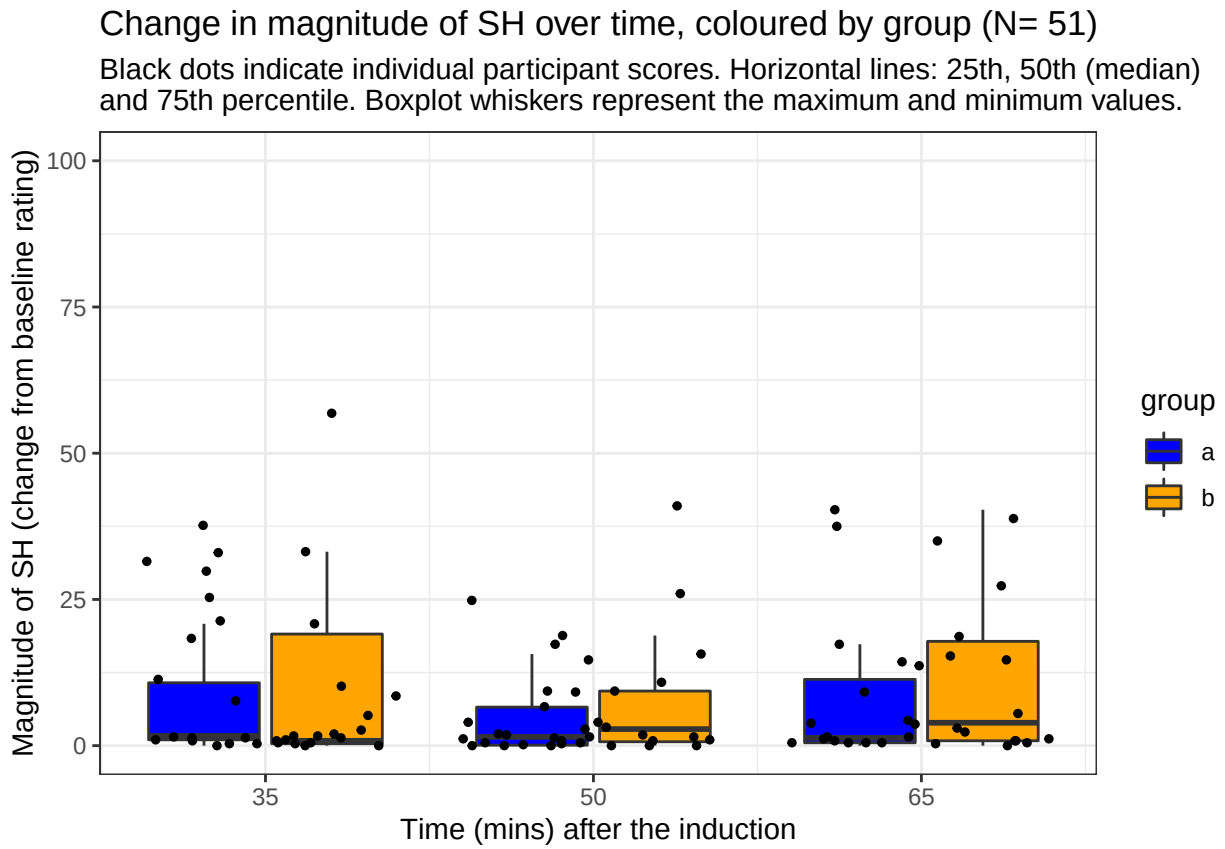
```
         time_point,
         mag) %>%
  separate(col = time_point,
           sep = '_',
           into = c('delete', 'time')) %>%
  dplyr::select(-delete) %>% # delete useless column
  mutate(time = as.numeric(time)) %>%
  rename(time_mag = time)

mag_change_bl %<>% right_join(began_procedure) %>%
  filter(
  grepl("completed", participation_status) |
  grepl("missing post-procedure data", participation_status)) %>%
  dplyr::select(pid, group, time_mag, mag, participation_status)
```

```
## Joining, by = "pid"
```

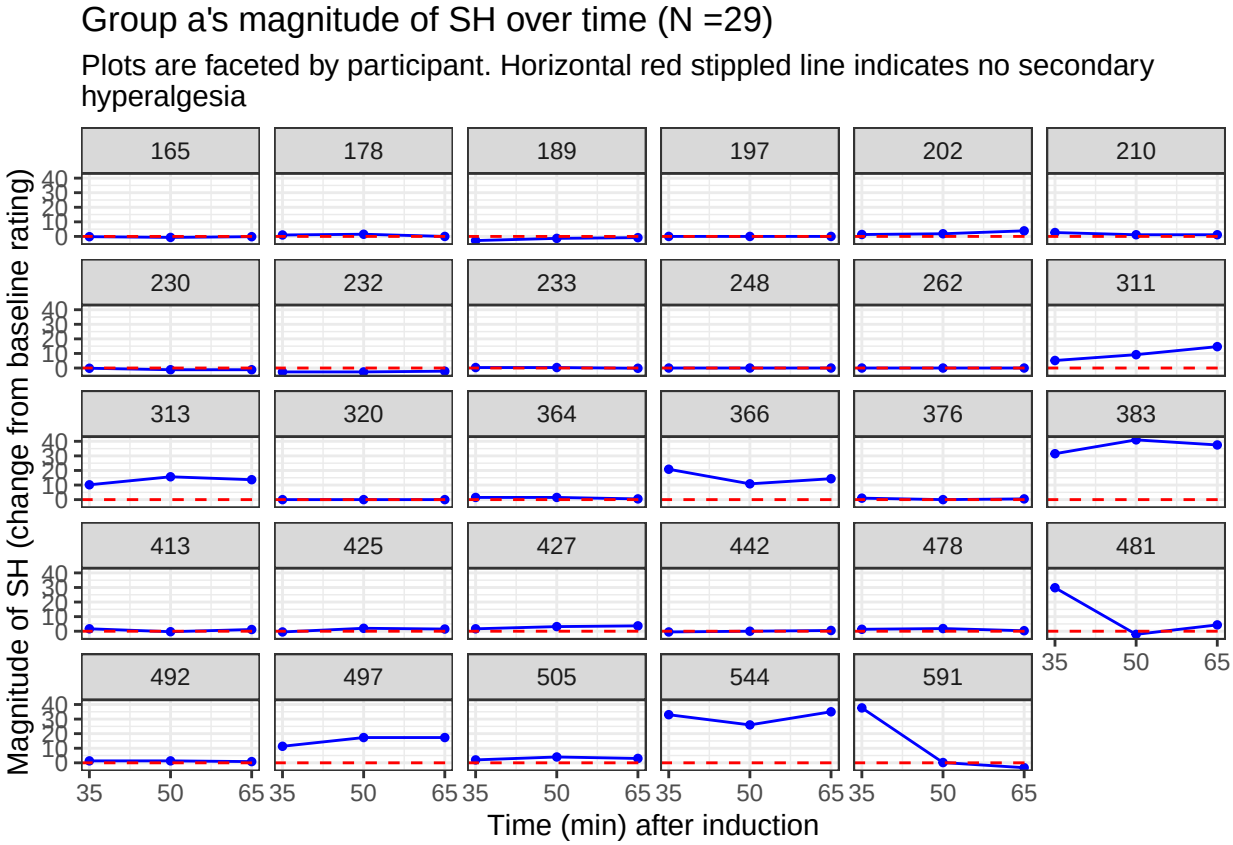
```
## plot the change in magnitude of SH by group over time
mag_change_plot <- ggplot(data = mag_change_bl) +
  aes(
    x = time_mag,
    y = mag,
    group = interaction(time_mag, group),
    fill = group) +
    geom_boxplot(outlier.shape = NA) +
    geom_jitter(shape = 21,
                size = 1,
                fill = 'black') +
  labs(
    title = "Change in magnitude of SH over time, coloured by group (N= 51)",
    subtitle = "Black dots indicate individual participant scores. Horizontal lines: 25th, 50th (median)
    x = "Time (mins) after the induction",
    y = "Magnitude of SH (change from baseline rating)") +
  scale_fill_manual(values = c('blue', 'orange')) +
  scale_y_continuous(limits = c(0, 100)) +
  scale_x_continuous(breaks = c(0, 35, 50, 65)) +
  theme_bw()

print(mag_change_plot)
```

```
## Warning: Removed 40 rows containing non-finite values (stat_boxplot).
```

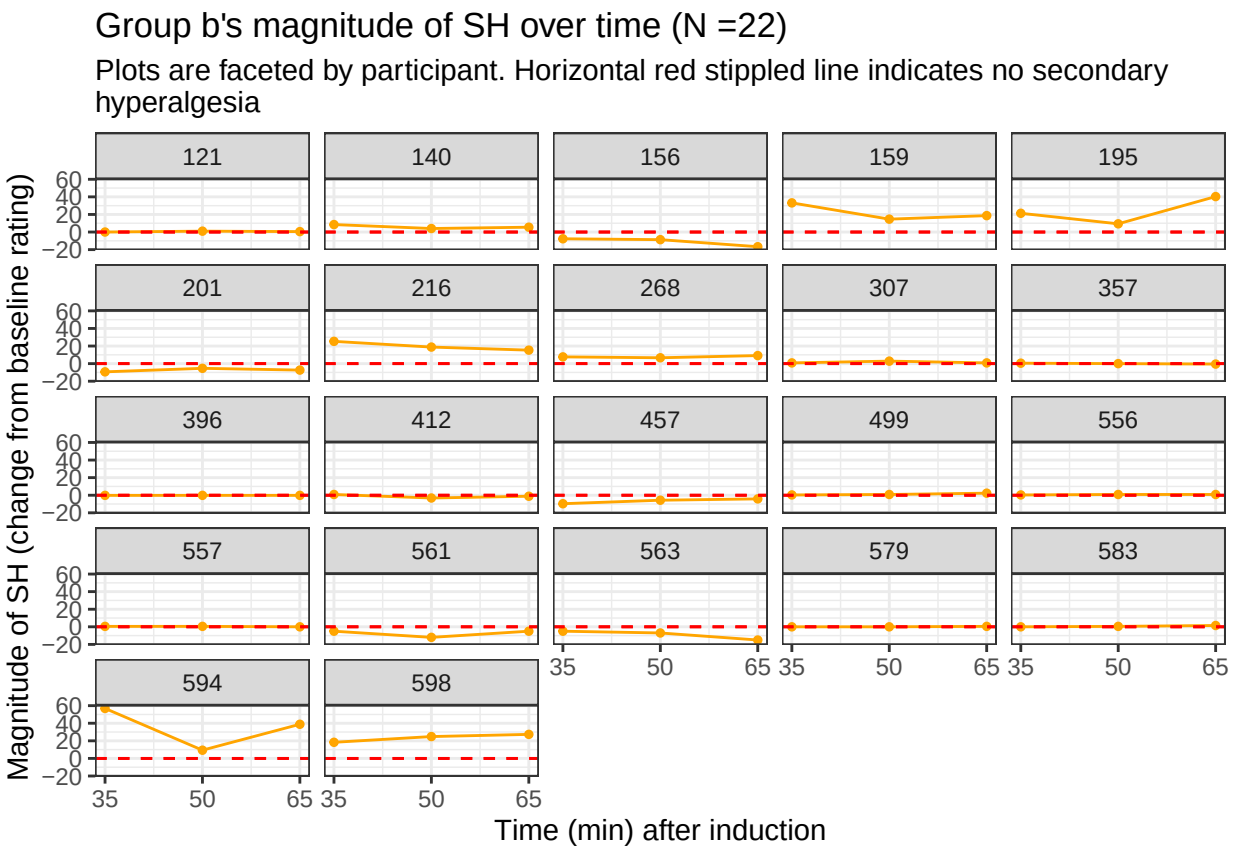
```
## Warning: Removed 50 rows containing missing values (geom_point).
```




```
mag_change_bl %>%
  filter(group == "a") %>%
ggplot(.) +
  aes(x = time_mag,
    y = mag) +
  geom_point(size = 1, shape = 21, fill = "blue", colour = "blue") +
  geom_line(colour = "blue") +
  facet_wrap( ~ pid, nrow = 5) +
  geom_hline(yintercept = 0, linetype = 2, colour = "red") +
  labs(title = "Group a's magnitude of SH over time (N =29)" ,
       subtitle = "Plots are faceted by participant. Horizontal red stippled line indicates no secondary
hyperalgesia",
       x= "Time (min) after induction", y = "Magnitude of SH (change from baseline rating)") +
  scale_x_continuous(breaks = c(35, 50, 65)) + theme_bw()
```




```
mag_change_bl %>%
  filter(group == "b") %>%
  ggplot(.) +
  aes(x = time_mag,
      y = mag) +
  geom_point(
    size = 1,
    shape = 21,
    fill = "orange",
    colour = "orange") +
  geom_line(colour = "orange") +
  geom_hline(yintercept = 0,
             linetype = 2,
             colour = "red") +
  facet_wrap(~ pid, nrow = 5) +
  geom_hline(yintercept = 0,
             linetype = 2,
             colour = "red") +
  labs(
    title = "Group b's magnitude of SH over time (N =22)" ,
    subtitle = "Plots are faceted by participant. Horizontal red stippled line indicates no secondary
hyperalgesia",
    x = "Time (min) after induction",
    y = "Magnitude of SH (change from baseline rating)") +
  scale_x_continuous(breaks = c(35, 50, 65)) +
  theme_bw()
```




```
## calculate area under the line
aul_mag_df <- mag_change_bl %>%
  group_by(pid) %>%
  mutate(aul_mag = area_under_curve(time_mag,
                                    mag,
                                    method = "trapezoid")) %>%
dplyr::select(-time_mag,-mag) %>%
  distinct()


## join outcomes data with participant demographics data
main_outcome_data <- full_join(aul_sa_df, aul_mag_df) %>%
  full_join(began_procedure) %>%
  dplyr::select(
    pid,
    age,
    sex,
    group,
    distress_tot,
    calibrated_current,
    aul_sa,
    aul_mag,
    participation_status
  ) %>%
```

```
  ungroup() %>%
  mutate(pid = as.character(pid)) %>%
  filter(
    grepl("completed", participation_status) |
    grepl("missing post-procedure data", participation_status))
```

```
## Joining, by = c("pid", "group", "participation_status")
## Joining, by = c("pid", "group", "participation_status")
```

```
# filter for participants with completed data sets and those with complete outcome data

rm(aul_sa_df, aul_mag_df, raw_mag_ratings_1)
```

We have 51 participants with complete magnitude data. The histogram plot hows that our magnitude data are not normally distributed. In our sample, 41% of participants developed an identifiable magnitude of SH as indicated by a change in ratings to pinprick stimulation (group a: 26%; group b: 59%). The boxplots shows us that the variance in magnitude ratings or magnitude of SH (chanage from baseline) for group b look much larger than for group a, across all time points. The AUL of the magnitude-relative-to-baseline over the three follow-up time points has been calculated and carried forward to analysis (see **Secondary analysis**).

Induction check

Here we imported ratings given for the 5 HFS trains and checked the distribution of these data using a histogram. These data were used to determine whether the induction was successful. We plotted these data to check for differences in rating of HFS trains between groups.

```
## prepare hfs data
hfs_ratings <- data.frame(master_data[[3]]) %>%
  rename(pid = blinded_id,
         train_rating = rating) %>%
  mutate(pid = as.character(pid)) %>%
  dplyr::select(pid, train_counter, train_rating) %>%
  mutate(train_rating = as.numeric(train_rating)) %>%
  mutate(train_rating = ifelse(train_rating == 999, NA, train_rating)) %>%
  mutate(train_counter = as.numeric(train_counter)) %>%
  left_join(main_outcome_data)
```

```
## Joining, by = "pid"
```

```
hfs_pain <- hfs_ratings %>%
  left_join(began_procedure) %>%
  filter(
    grepl("completed", participation_status) |
    grepl("missing post-procedure data", participation_status))
```

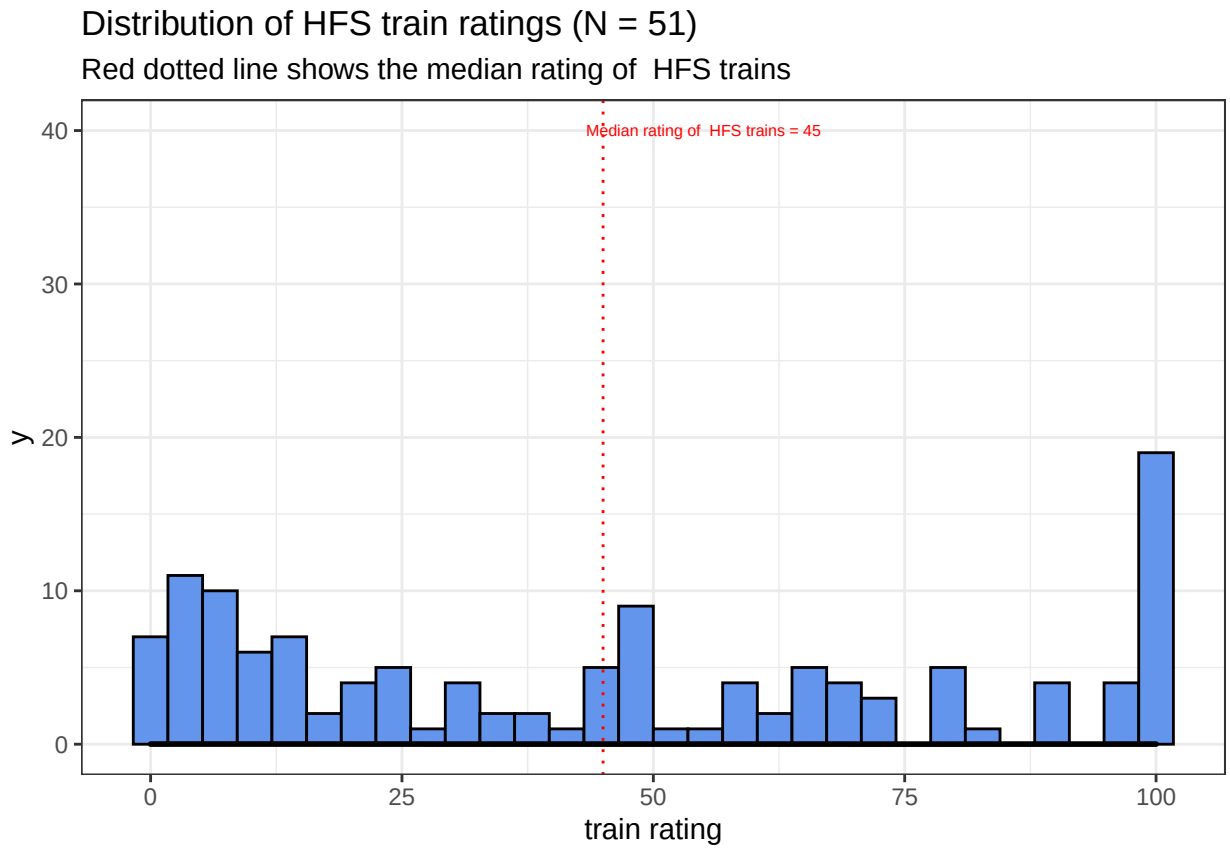
```
## Joining, by = c("pid", "age", "sex", "group", "distress_tot",
## "calibrated_current", "participation_status")
```

```r
## check distribution of hfs data
hfs <-
  hfs_pain %>%
  dplyr::select(pid, group, train_counter, train_rating) %>%
  group_by(pid, train_counter) %>%
  unique()

median_hfs <- median(hfs$train_rating, na.rm = TRUE)# calculate median rating of  HFS trains

ggplot(hfs) + aes(x = train_rating) +
  geom_histogram(fill = "#6495ed", col = "black") +
  geom_density(lwd = 1) +
  geom_vline(
    xintercept = median_hfs,
    col = "red",
    lwd = 0.5,
    linetype = 'dotted') +
  annotate("text", # Add text for median
           x = 55,
           y = 40,
           label = paste("Median rating of  HFS trains = 45"),
           col = "red",
           size = 2) +
  labs(title = "Distribution of HFS train ratings (N = 51)",
       subtitle = "Red dotted line shows the median rating of  HFS trains",
       x = "train rating") +
  theme_bw()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 126 rows containing non-finite values (stat_bin).
```

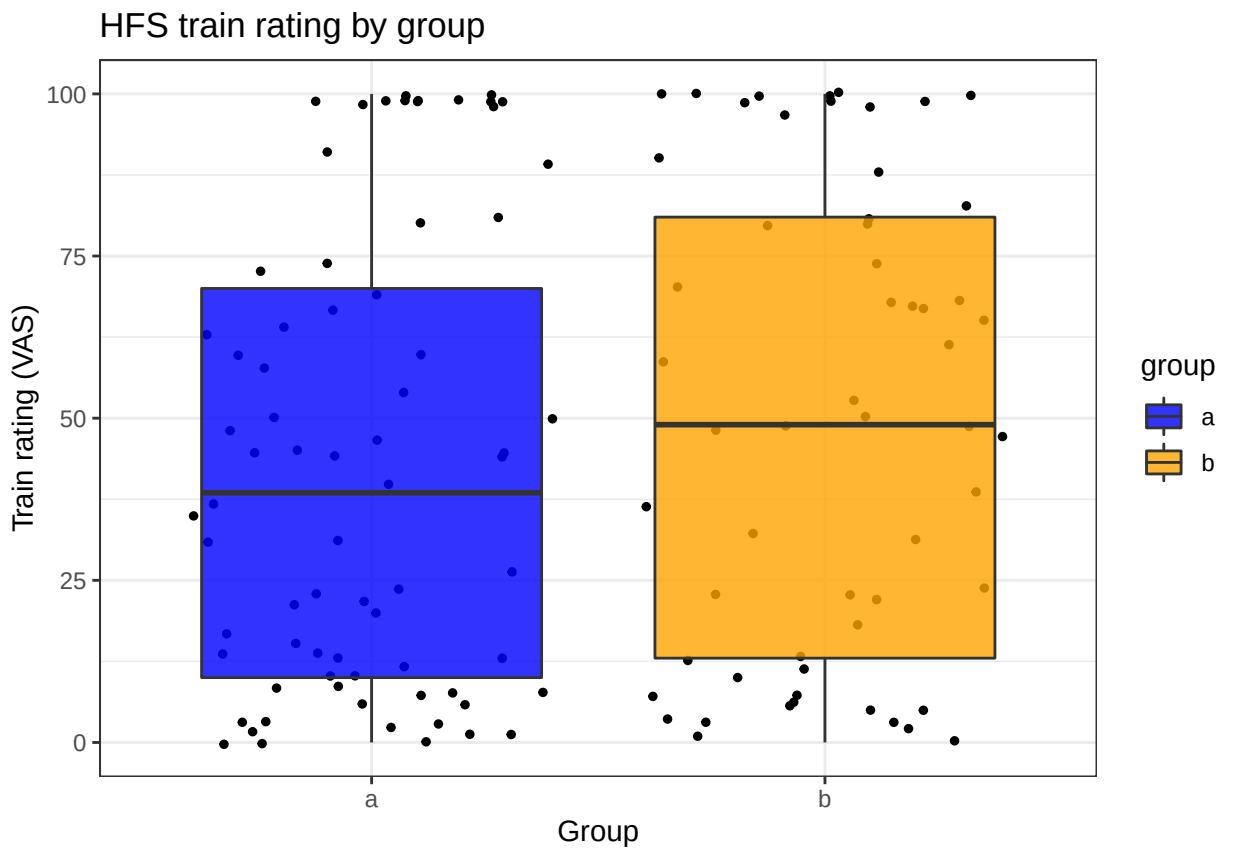
```
## Warning: Removed 126 rows containing non-finite values (stat_density).
```




```
## rating of  HFS trains between groups
ggplot(hfs) +
  aes(x = group,
      y = train_rating,
      fill = group) +
  geom_jitter(size = 1, shape = 21, fill = "black") +
  geom_boxplot(alpha = 0.8) +
  scale_fill_manual(values = c('blue', 'orange')) +
  labs(title = "HFS train rating by group" ,
       x = "Group", y = "Train rating (VAS)") +
  theme_bw()
```

```
## Warning: Removed 126 rows containing non-finite values (stat_boxplot).
```

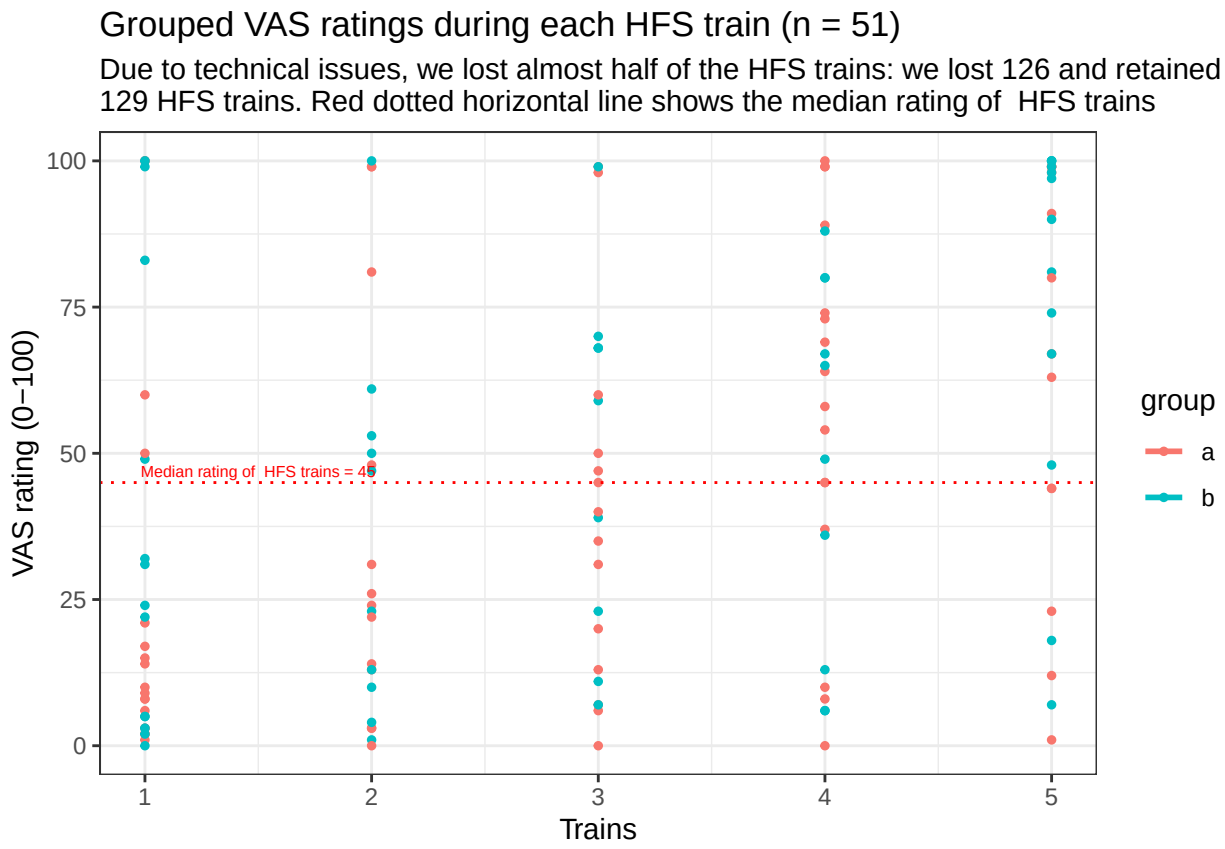
```
## Warning: Removed 126 rows containing missing values (geom_point).
```




```
## manipulation check: HFS train rating
ggplot(data = hfs_pain) +
  aes(
    x = train_counter,
    y = train_rating,
    group = interaction(group, train_counter),
    colour = group) +
  geom_point(size = 1) +
geom_smooth(method = 'lm',se = FALSE) +
  geom_hline(
    yintercept = median_hfs,
    col = "red",
    lwd = 0.5,
    linetype = 'dotted') +
  annotate("text",  # Add text for median
           x = 1.5,
           y = 47,
           label = paste("Median rating of  HFS trains = 45"),
           col = "red",
           size = 2) +
  labs(title = "Grouped VAS ratings during each HFS train (n = 51)",
       subtitle = "Due to technical issues, we lost almost half of the HFS trains: we lost 126 and reta
       x = "Trains",
       y = "VAS rating (0-100)") +
  theme_bw()
```

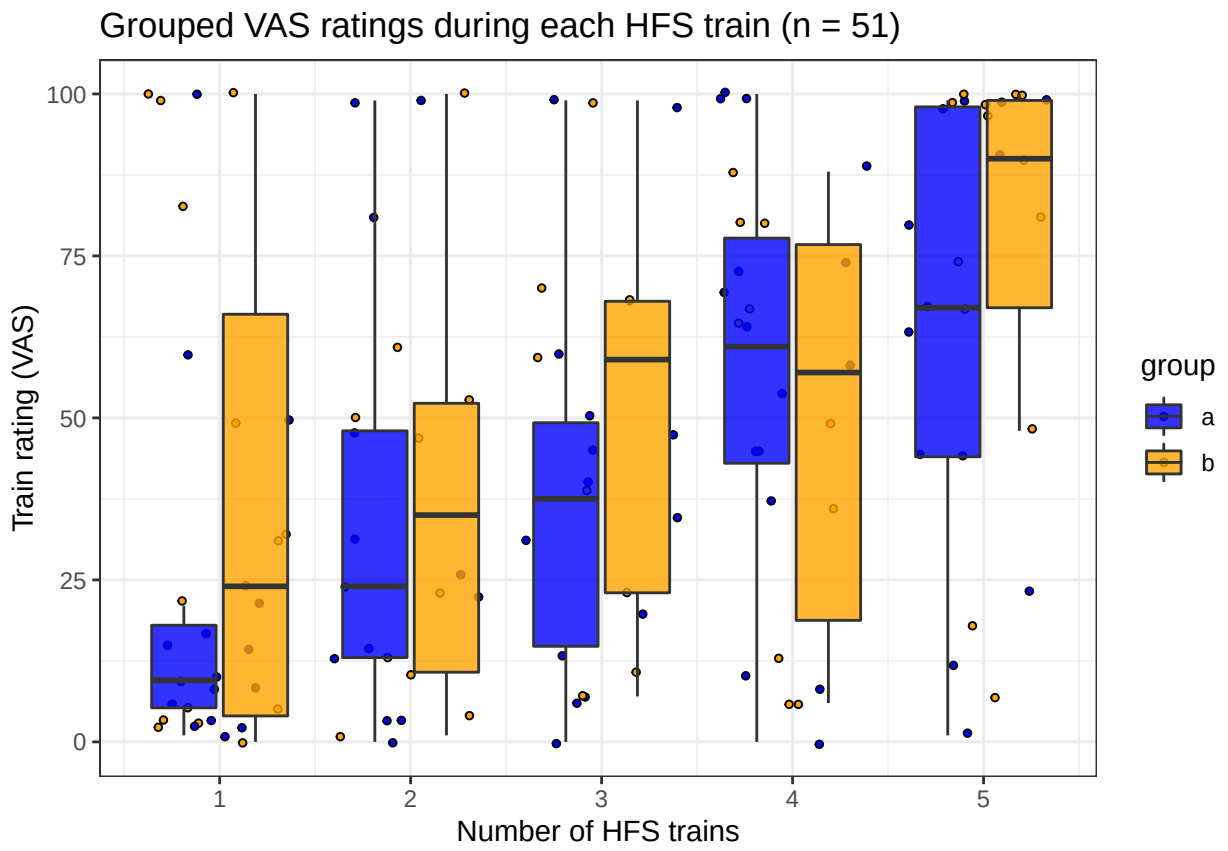
```
## `geom_smooth()` using formula 'y ~ x'
```

```
## Warning: Removed 126 rows containing non-finite values (stat_smooth).
## Removed 126 rows containing missing values (geom_point).
```




```
## HFS train rating, faceted by group
ggplot(hfs_pain) +
  aes(
    x = train_counter,
    y = train_rating,
    group = interaction(group, train_counter),
    fill = group) +
  geom_jitter(size = 1, shape = 21) +
  geom_boxplot(outlier.shape = NA, alpha =0.8) +
  scale_fill_manual(values = c('blue', 'orange')) +
  labs(title = "Grouped VAS ratings during each HFS train (n = 51)" ,
       x = "Number of HFS trains", y = "Train rating (VAS)") +
  theme_bw()
```

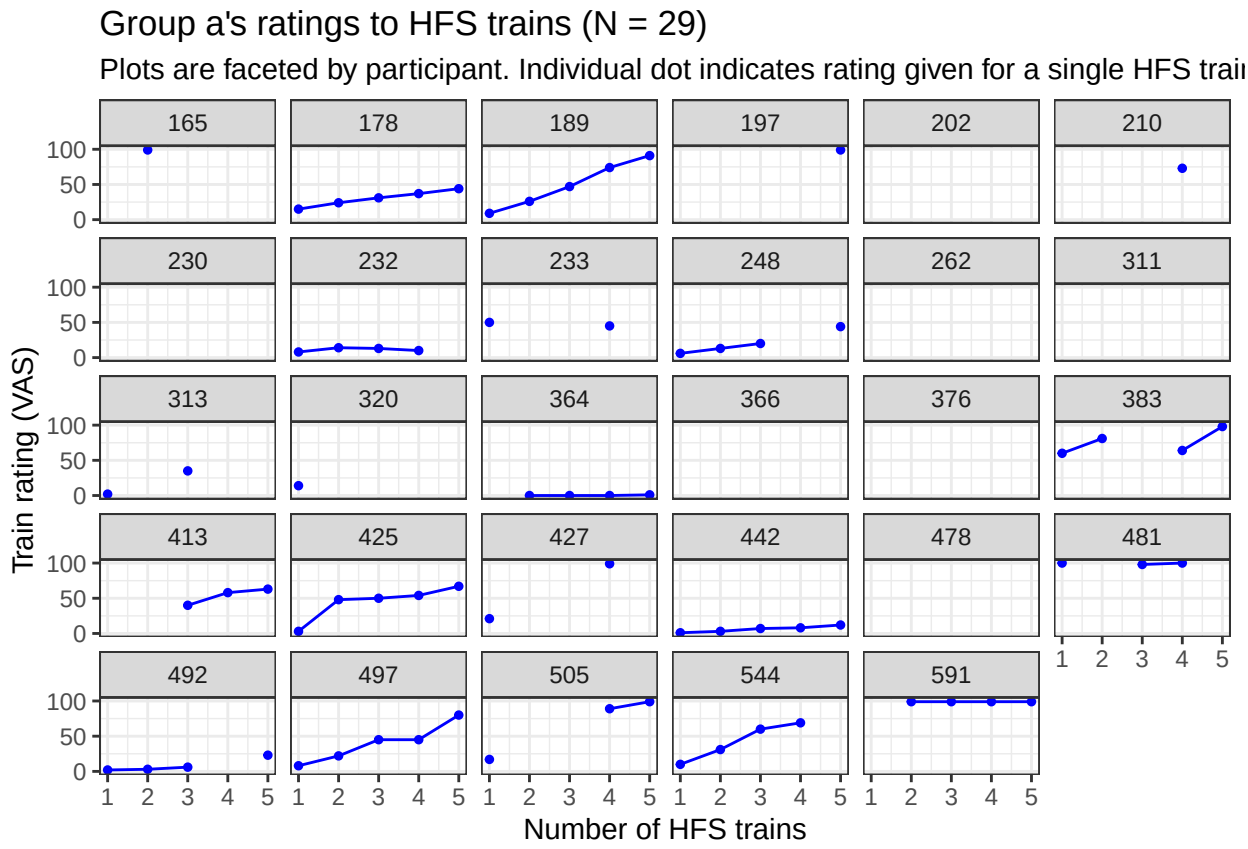
```
## Warning: Removed 126 rows containing non-finite values (stat_boxplot).
## Removed 126 rows containing missing values (geom_point).
```




```
## plot HFS VAS ratings for each participant; please note missing values on plot
### group a
hfs %>% filter(group == 'a') %>%
  ggplot(.) +
  aes(x = train_counter,
      y = train_rating) +
  geom_point(
    size = 1,
    shape = 21,
    fill = 'blue',
    colour = 'blue') +
  geom_line(colour = 'blue') +
  facet_wrap(~ pid, nrow = 5) +
  labs(
    title = "Group a's ratings to HFS trains (N = 29)" ,
    subtitle = 'Plots are faceted by participant. Individual dot indicates rating given for a single HFS
    x = "Number of HFS trains",
    y = "Train rating (VAS)") +
  scale_y_continuous(breaks = c(0, 50, 100)) +
  theme_bw()
```

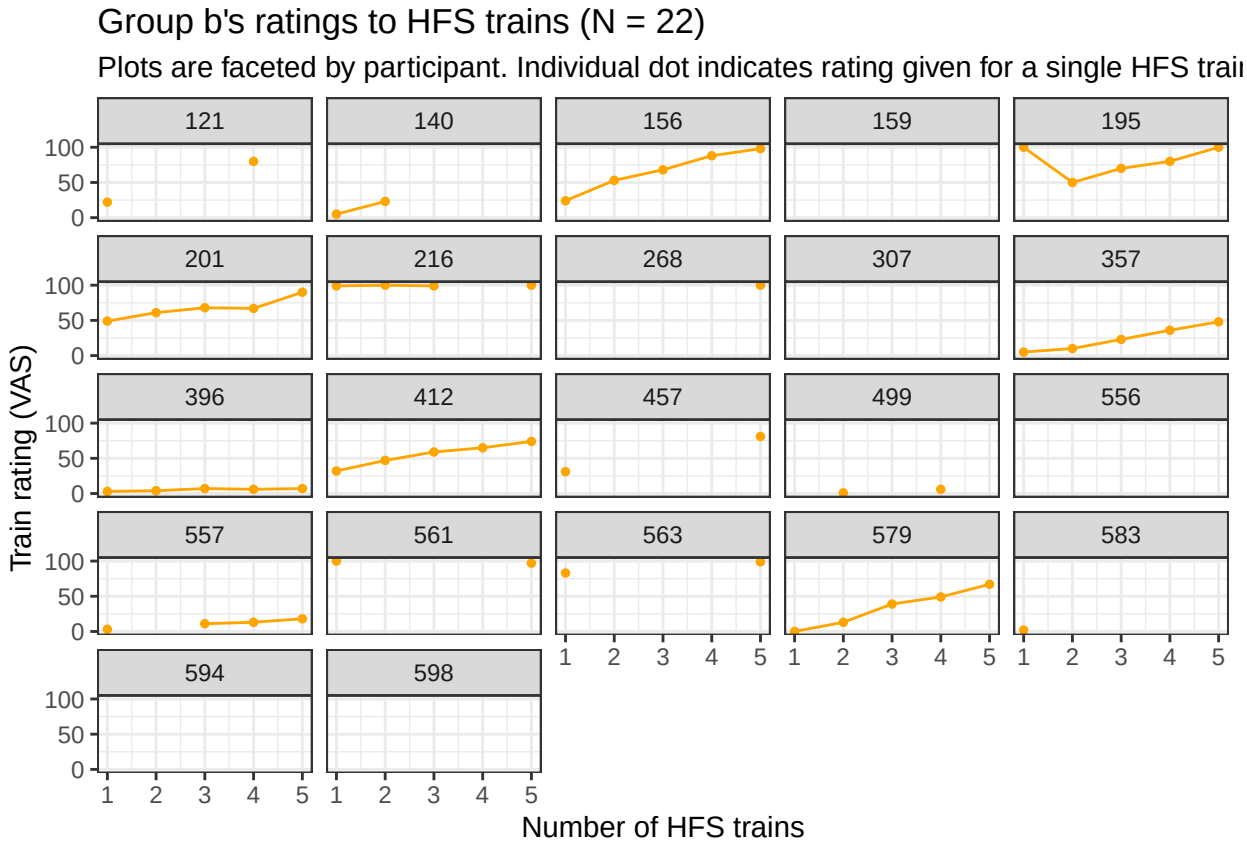
```
## Warning: Removed 73 rows containing missing values (geom_point).
```

```
## Warning: Removed 1 row(s) containing missing values (geom_path).
```




```
### group b
hfs %>% filter(group == 'b') %>%
  ggplot(.) +
  aes(x = train_counter,
      y = train_rating) +
  geom_point(
    size = 1,
    shape = 21,
    fill = 'orange',
    colour = 'orange') +
  geom_line(colour = 'orange') +
  facet_wrap(~ pid, nrow = 5) +
  labs(
    title = "Group b's ratings to HFS trains (N = 22)" ,
    subtitle = 'Plots are faceted by participant. Individual dot indicates rating given for a single HF
    x = "Number of HFS trains",
    y = "Train rating (VAS)") +
  scale_y_continuous(breaks = c(0, 50, 100)) +
  theme_bw()
```

```
## Warning: Removed 53 rows containing missing values (geom_point).

## Warning: Removed 14 row(s) containing missing values (geom_path).

## geom_path: Each group consists of only one observation. Do you need to adjust
## the group aesthetic?
```




```
# all 52 participants received 5 trains of hfs.
total_trains <- hfs %>% filter(!is.na(train_counter)) %>% nrow()
glue("We anticipated a total of {total_trains} HFS trains")
```

```
## We anticipated a total of 255 HFS trains
```

```
# we anticipated collecting a total of 255 trains, however,
# due to technical issues some hfs data were lost.

## calculate number of missing trains
missing_trains <- hfs %>% filter(is.na(train_rating)) %>% nrow()
glue("From the {total_trains} trains we anticipated, we lost {missing_trains} HFS trains")
```

```
## From the 255 trains we anticipated, we lost 126 HFS trains
```

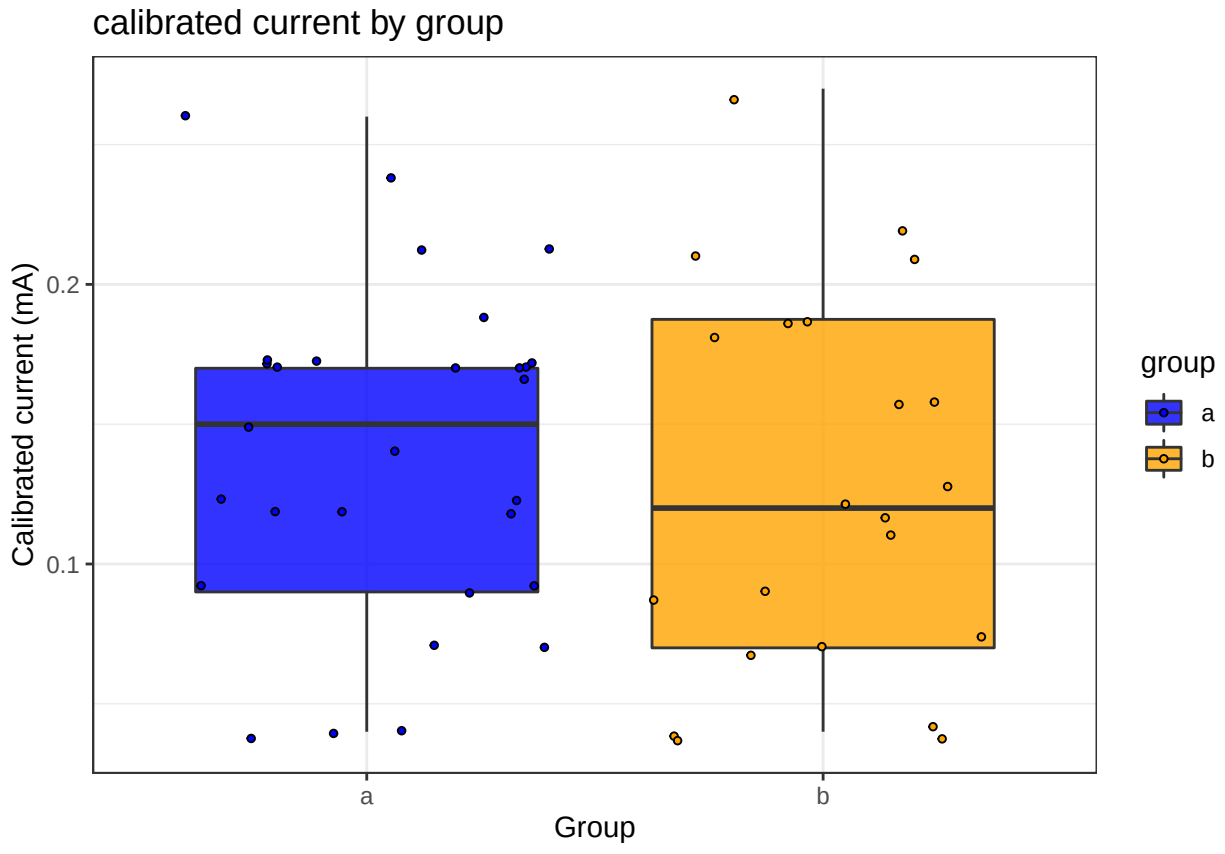
```
rm(hfs_ratings, hfs)
```

Here we imported HFS data and checked the distribution of these data. The histogram plot showed that our data are not normally distributed. Due to technical issues, almost half of the HFS train ratings were missing: 126 ratings were missing; 129 ratings were available. Based on our plots, there is no difference in rating of HFS trains between groups. However, in both groups, we saw a difference by train number: later trains were rated higher on the scale than earlier trains. Ratings to HFS trains ranged between 0 and 100 on the VAS, and the median rating of HFS trains was 45.

Plots

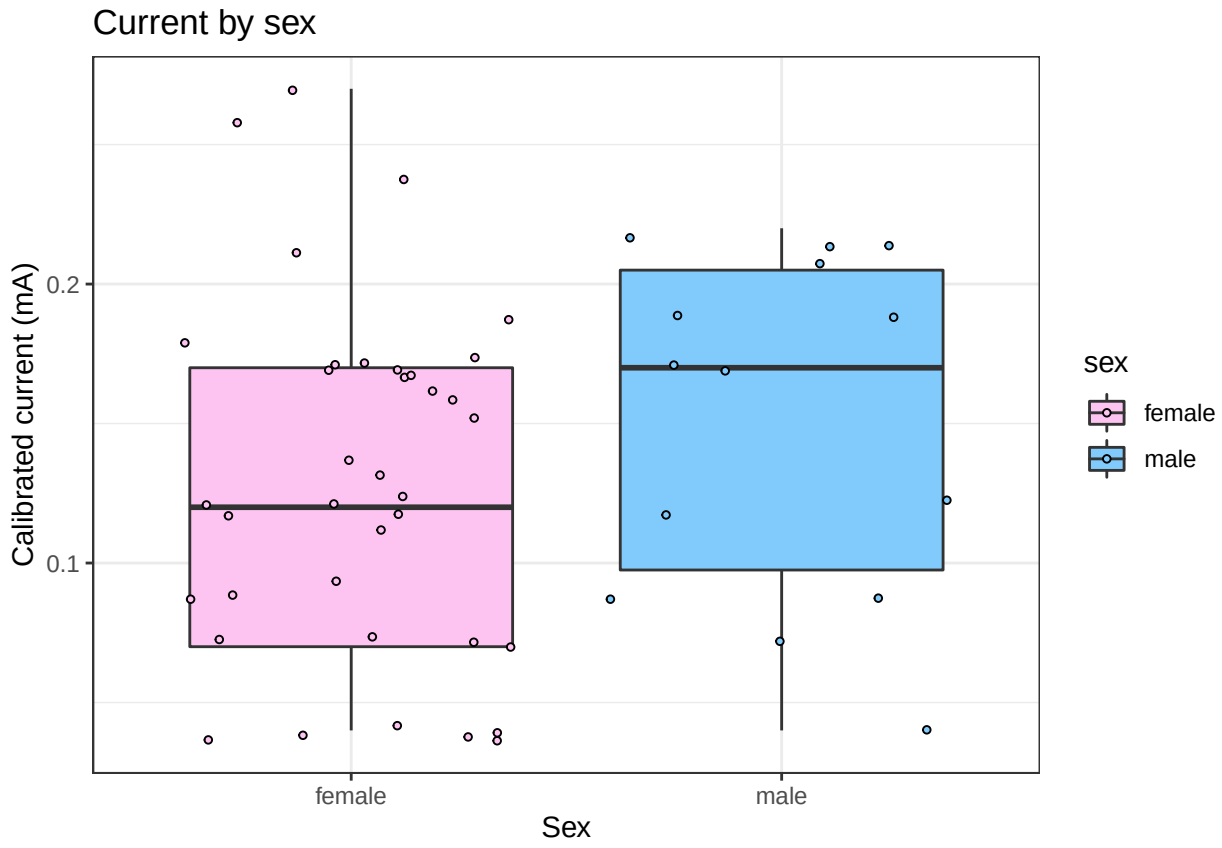
Here we plotted calibrated current, rating of HFS trains, and PSD severity to identify differences between (i) groups, and across (ii) sex and (iii) age.

```
## calibrated current
## plot current by group
main_outcome_data %>% filter(grepl("completed", participation_status) |
  grepl("missing post-procedure data", participation_status)) %>%
    dplyr::select(pid, group, sex, age, calibrated_current) %>%
    unique() %>%
  ggplot(data = .) +
  aes(x = group,
      y = calibrated_current,
      fill = group) +
  geom_boxplot(alpha = 0.8) +
  geom_jitter(shape = 21,
    size = 1) +
  scale_fill_manual(values = c('blue', 'orange')) +
  labs(x = "Group", y = "Calibrated current (mA)",
       title = "calibrated current by group") +
  theme_bw()
```




```
## plot current by  sex
main_outcome_data %>% filter(grepl("completed", participation_status) |
```

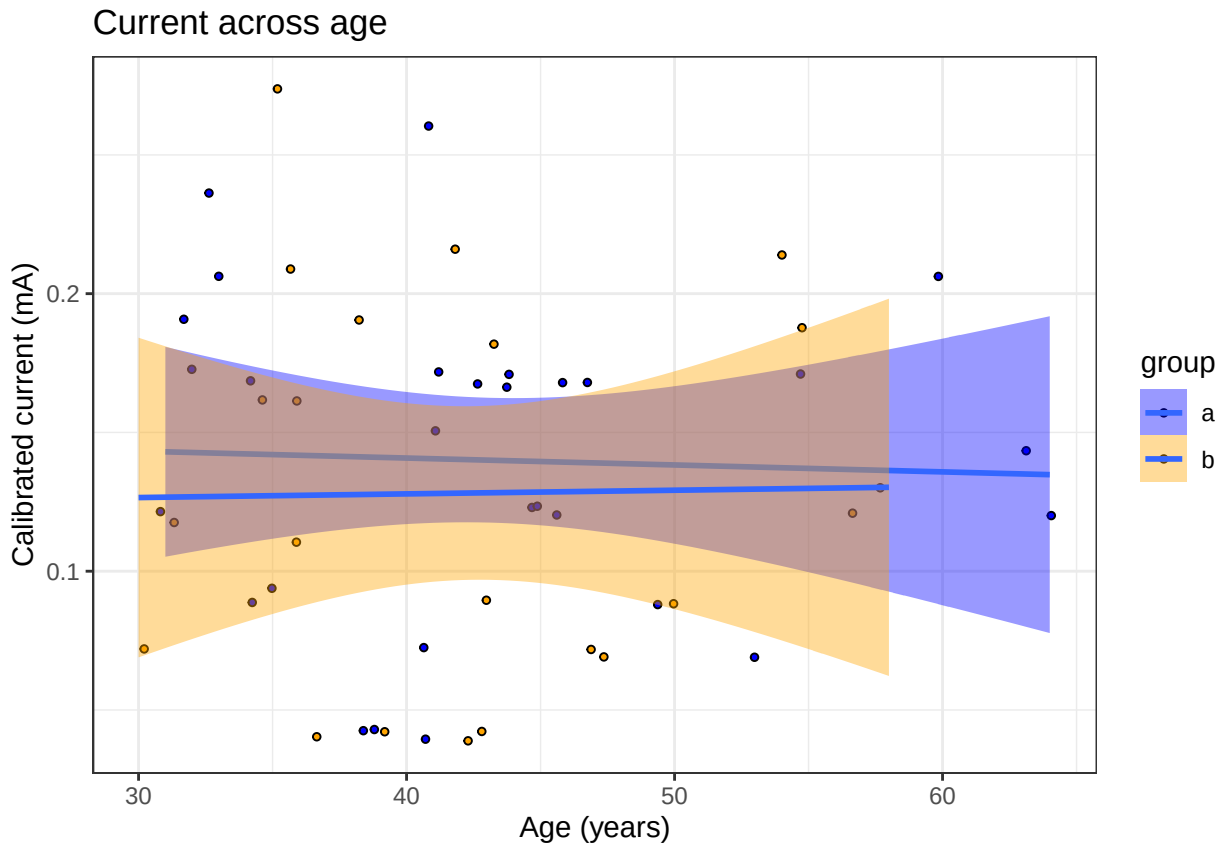
```
  grepl("missing post-procedure data", participation_status)) %>%
    dplyr::select(pid, group, sex, age, calibrated_current) %>%
    unique() %>%
ggplot(data = .) +
  aes(x = sex,
      y = calibrated_current,
      fill = sex) +
  geom_boxplot() +
  scale_fill_manual(values = c("female" = "#FDC4F1", "male" = "#80CBFB")) +
  geom_jitter(
    shape = 21,
    size = 1) +
  labs(x = "Sex", y = "Calibrated current (mA)",
       title = "Current by sex") +
  theme_bw()
```




```
## plot current across ages
main_outcome_data %>% filter(grepl("completed", participation_status) |
  grepl("missing post-procedure data", participation_status)) %>%
    dplyr::select(pid, group, sex, age, calibrated_current) %>%
    unique() %>%
ggplot(data = .) +
  aes(x = age,
      y = calibrated_current,
      fill = group) +
```

```
  geom_jitter(shape = 21,
    size = 1) +
  scale_fill_manual(values = c('blue', 'orange')) +
  geom_smooth(method = "lm") +
  labs(x = "Age (years)", y = "Calibrated current (mA)",
       title = "Current across age") +
  theme_bw()
```

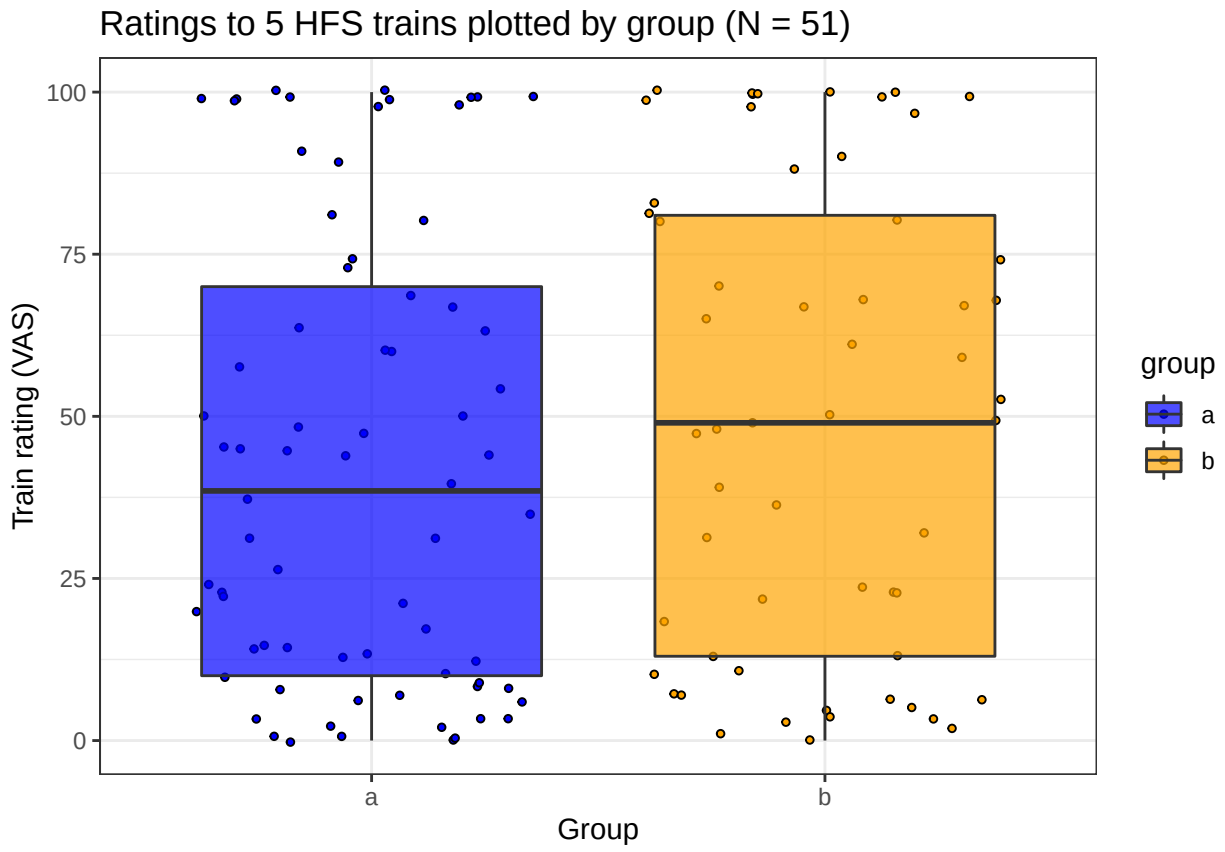
```
## `geom_smooth()` using formula 'y ~ x'
```




```
## rating of  HFS trains
## rating of  HFS trains by group
ggplot(hfs_pain) +
  aes(x = group,
    y = train_rating,
    fill = group) +
  geom_jitter(size = 1, shape = 21) +
  geom_boxplot(alpha = 0.7) +
  scale_fill_manual(values = c('blue', 'orange')) +
  labs(title = "Ratings to 5 HFS trains plotted by group (N = 51)" ,
       x= "Group", y = "Train rating (VAS)") +
  theme_bw()
```

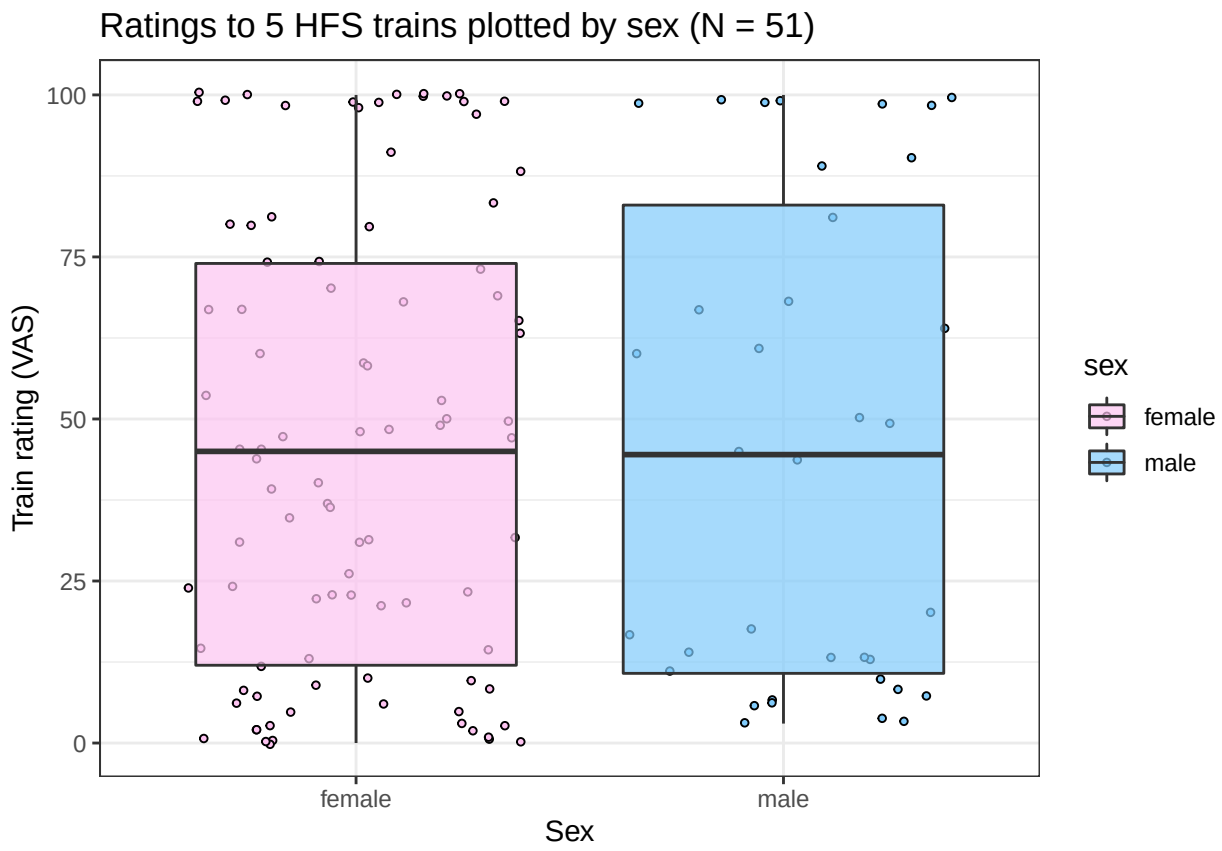
```
## Warning: Removed 126 rows containing non-finite values (stat_boxplot).
```

```
## Warning: Removed 126 rows containing missing values (geom_point).
```



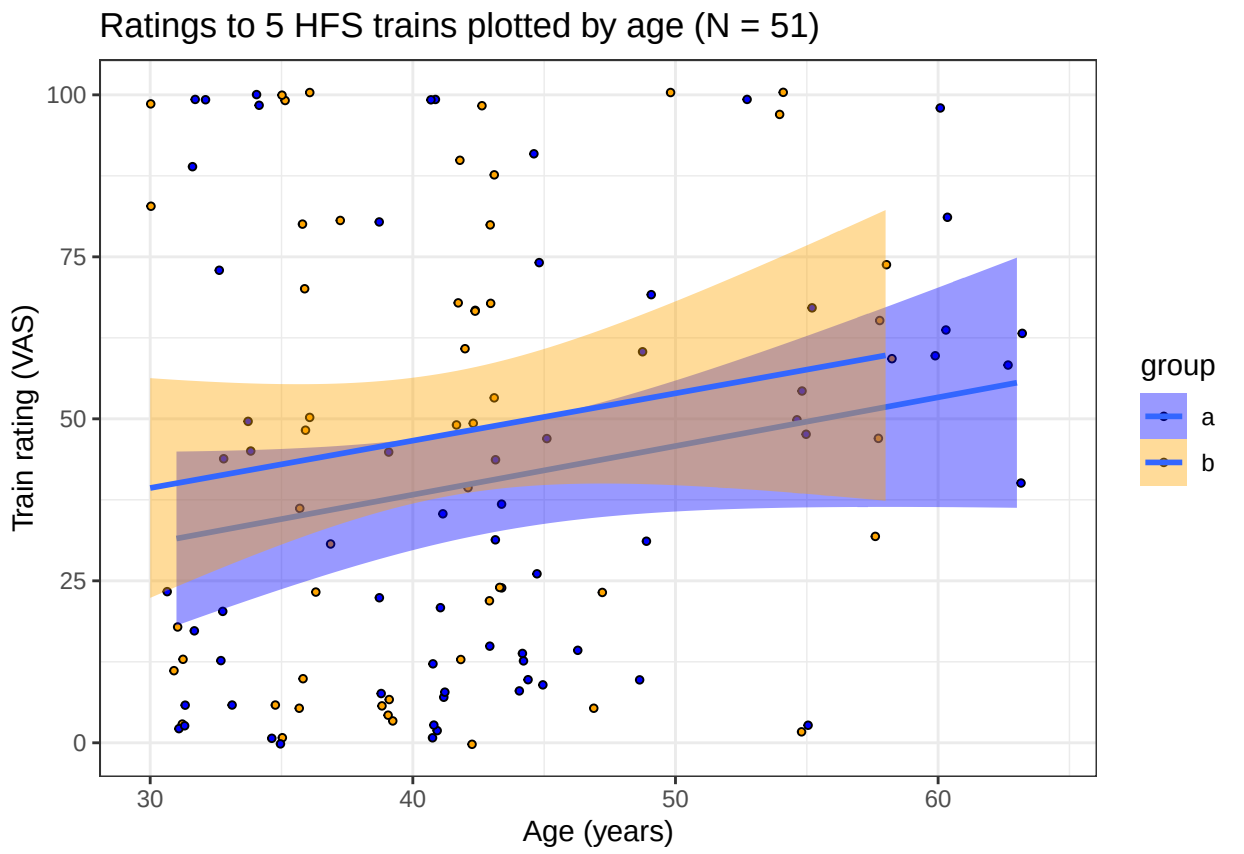

```
## rating of  HFS trains by  sex
ggplot(hfs_pain) +
  aes(x = sex,
    y = train_rating,
    fill = sex) +
  scale_fill_manual(values = c("female" = "#FDC4F1", "male" = "#80CBFB"))+
  geom_jitter(size = 1, shape = 21) +
  geom_boxplot(alpha = 0.7) +
  labs(title = "Ratings to 5 HFS trains plotted by sex (N = 51)" ,
       x= "Sex", y = "Train rating (VAS)") +
  theme_bw()
```

```
## Warning: Removed 126 rows containing non-finite values (stat_boxplot).
## Removed 126 rows containing missing values (geom_point).
```




```
## rating of  HFS trains across ages
hfs_pain %>%
  dplyr::select(pid, age, sex, group, train_rating) %>%
  unique() %>%
ggplot(.) +
  aes(x = age,
    y = train_rating,
    fill = group) +
  geom_jitter(size = 1, shape = 21) +
  scale_fill_manual(values = c('blue', 'orange')) +
  labs(title = "Ratings to 5 HFS trains plotted by age (N = 51)" ,
       x= "Age (years)", y = "Train rating (VAS)") +
  theme_bw() +
    geom_smooth(method=lm) # plot suggests that age predicted HFS ratings
```

```
## `geom_smooth()` using formula 'y ~ x'

## Warning: Removed 39 rows containing non-finite values (stat_smooth).

## Warning: Removed 39 rows containing missing values (geom_point).
```




```
# Did age predict ratings to HFS trains?
model_hfs_age <- lm(train_rating ~ age, data = hfs_pain)
summary(model_hfs_age) # age predicted ratings to HFS trains p = 0.02
```

Call: lm(formula = train_rating ~ age, data = hfs_pain)

Residuals: Min 1Q Median 3Q Max -53.359 -33.015 -5.244 33.556 62.214

Coefficients: Estimate Std. Error t value Pr(>|t|)
(Intercept) 14.4989 15.5105 0.935 0.3517
age 0.7429 0.3600 2.064 0.0411 * — Signif. codes: 0 ‘’ ***0.001*** ’’ *0.01* ’’ 0.05 ‘.’ 0.1 ’ ’ 1

Residual standard error: 35.01 on 127 degrees of freedom (126 observations deleted due to missingness) Multiple R-squared: 0.03244, Adjusted R-squared: 0.02483 F-statistic: 4.259 on 1 and 127 DF, p-value: 0.04109

```
# Did age and group predict ratings to HFS trains?
model_hfs_age_multiple <- lm(train_rating ~ age+group, data = hfs_pain)
summary(model_hfs_age_multiple) # group did not predict ratings to HFS trains p = 0.17
```

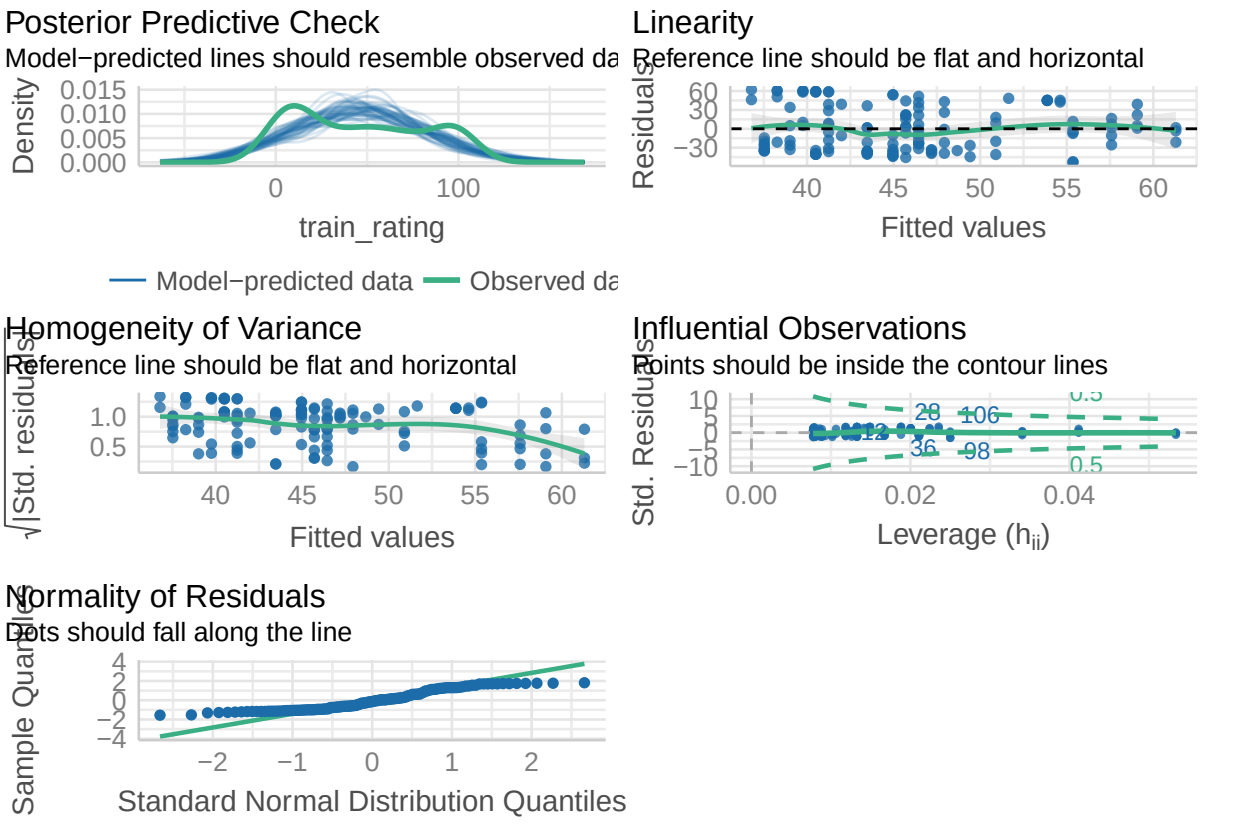
Call: lm(formula = train_rating ~ age + group, data = hfs_pain)

Residuals: Min 1Q Median 3Q Max -58.877 -31.095 -4.284 29.674 65.103

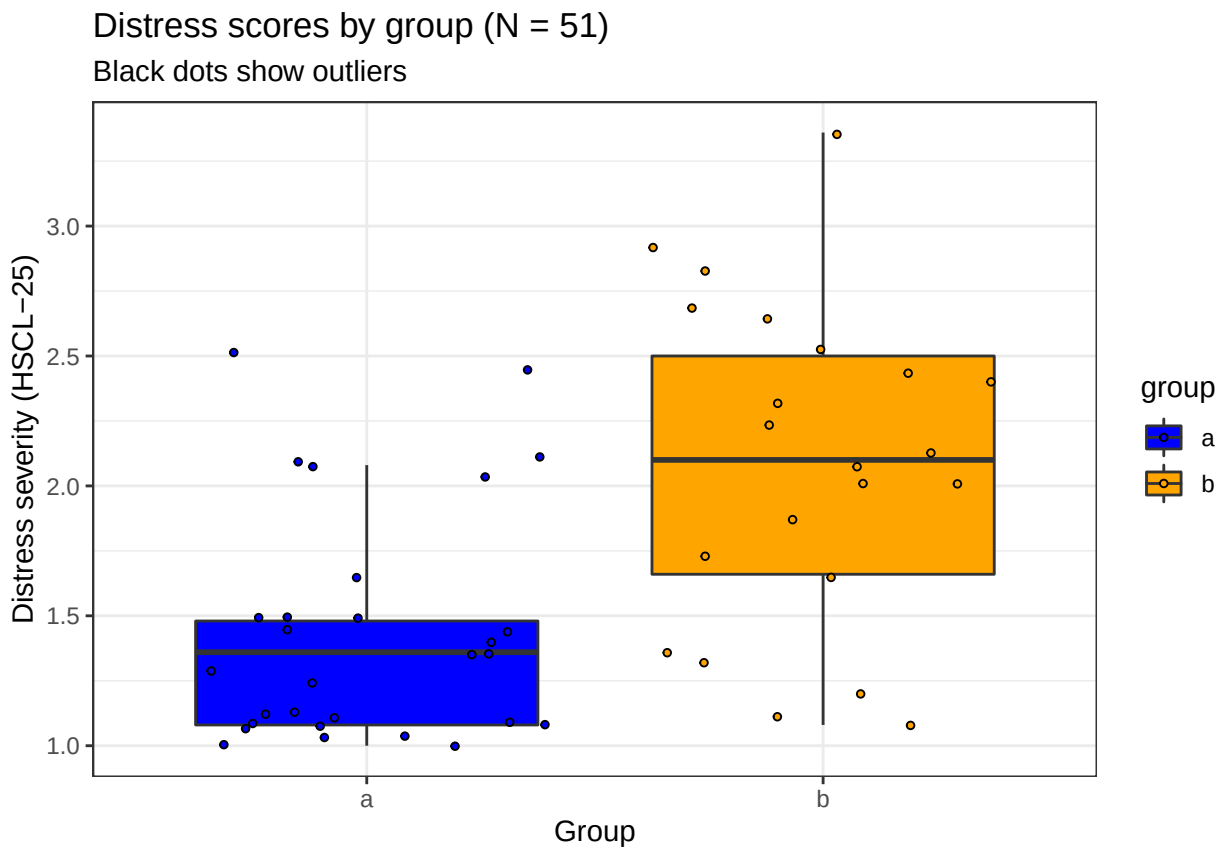
Coefficients: Estimate Std. Error t value Pr(>|t|)
(Intercept) 8.2251 16.1229 0.510 0.6108
age 0.8023 0.3614 2.220 0.0282 * groupb 8.5281 6.2312 1.369 0.1736
— Signif. codes: 0 ‘’ ***0.001*** ’’ *0.01* ’’ 0.05 ‘.’ 0.1 ’ ’ 1

Residual standard error: 34.89 on 126 degrees of freedom (126 observations deleted due to missingness) Multiple R-squared: 0.04662, Adjusted R-squared: 0.03148 F-statistic: 3.08 on 2 and 126 DF, p-value: 0.04941

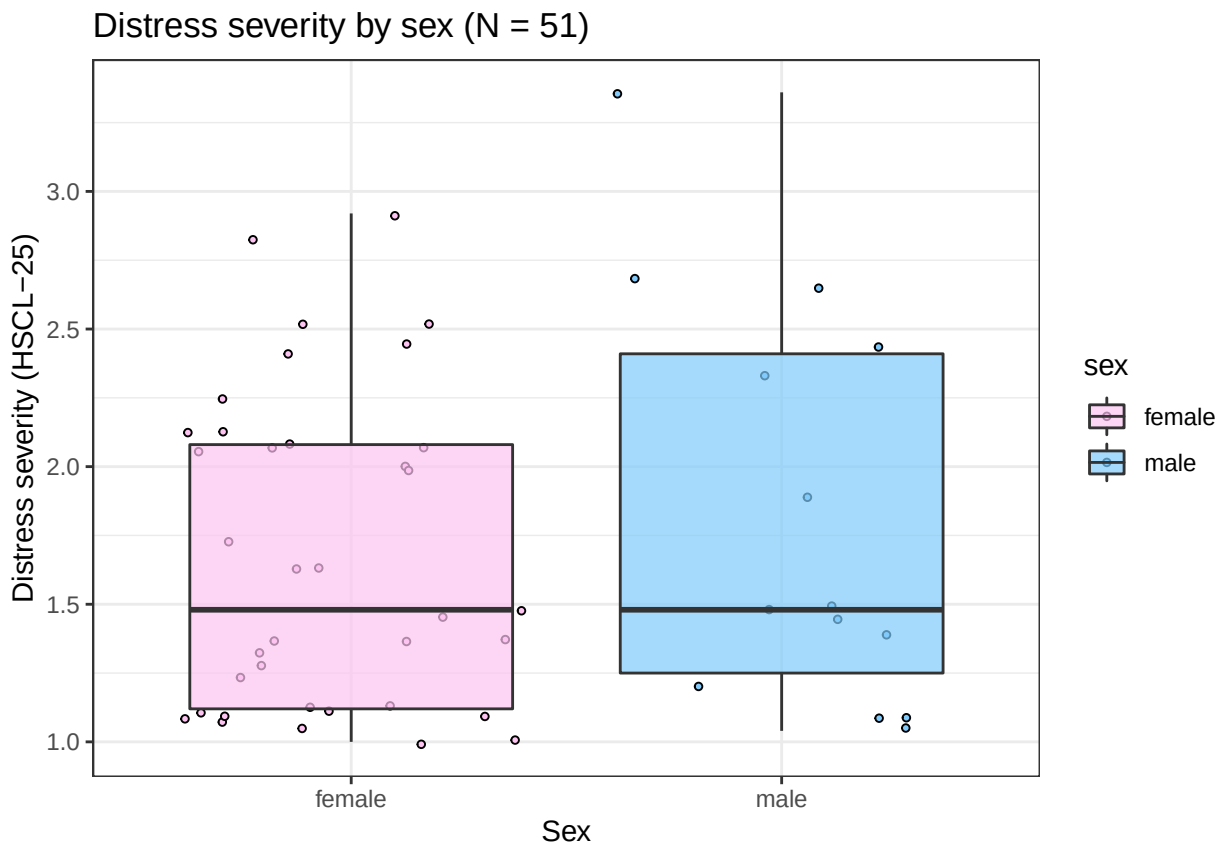
```
check_model(model_hfs_age) # assumption of the posterior predictive check and normality of residuals ma
```




```
## PSD severity
## Plot PSD severity by group
ggplot(analyse) +
aes(x = group,
y = distress_tot,
fill = group) +
geom_boxplot(outlier.shape = NA) +
geom_jitter(size = 1, shape = 21) +
  scale_fill_manual(values = c('blue', 'orange')) +
  labs(title = "Distress scores by group (N = 51)", subtitle = "Black dots show outliers",
x= "Group", y = "Distress severity (HSCL-25)") +
theme_bw() # clear difference between groups
```

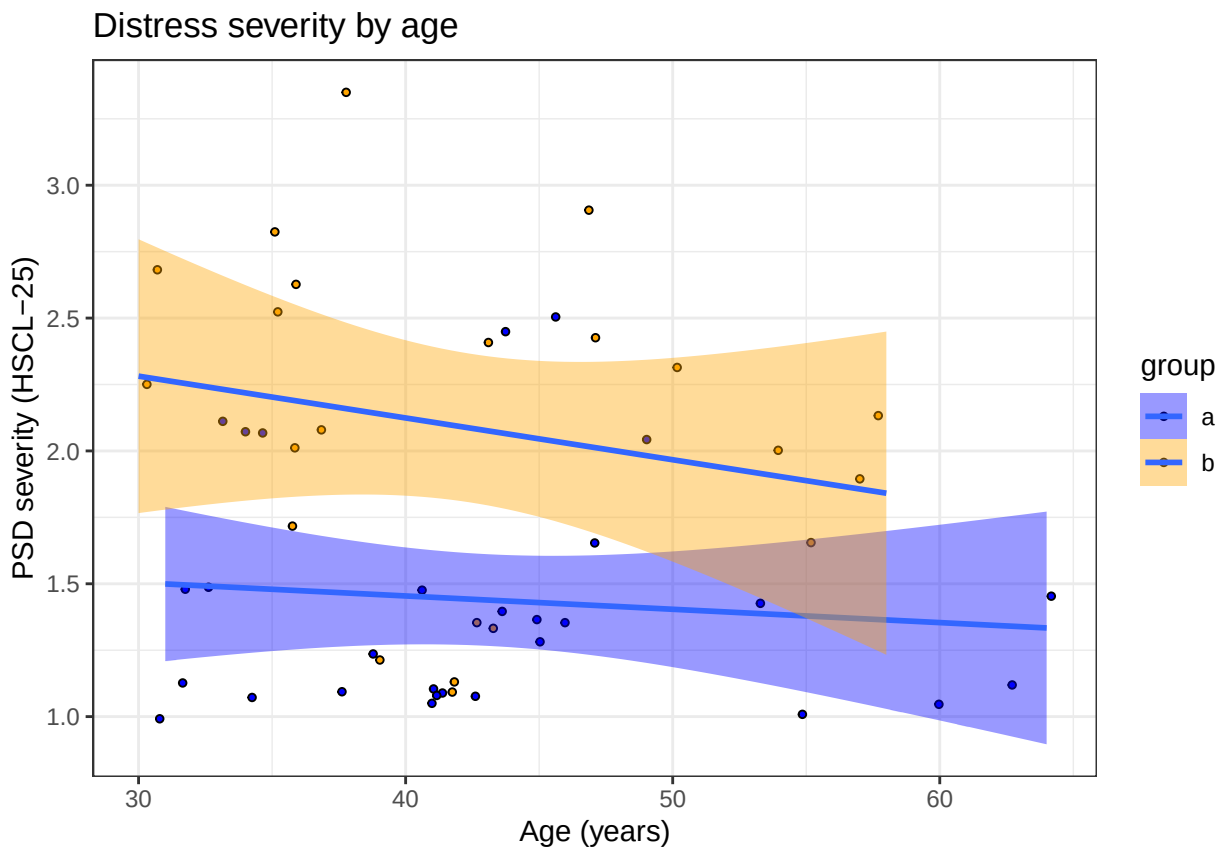



```
## Plot PSD severity by sex
ggplot(analyse) +
  aes(x = sex,
    y = distress_tot,
    fill = sex) +
  geom_jitter(size = 1, shape = 21) +
  geom_boxplot(alpha = 0.7) +
  scale_fill_manual(values = c("female" = "#FDC4F1", "male" = "#80CBFB")) +
  labs(title = "Distress severity by sex (N = 51)" ,
       x= "Sex", y = "Distress severity (HSCL-25)") +
  theme_bw()
```




```
## Plot PSD severity across ages
ggplot(analyse) +
  aes(x = age,
    y = distress_tot,
    fill = group) +
  geom_jitter(size = 1, shape = 21) +
  scale_fill_manual(values = c('blue', 'orange')) +
  labs(title = "Distress severity by age" ,
       x= "Age (years)", y = "PSD severity (HSCL-25)") +
  theme_bw() +
    geom_smooth(method=lm)
```

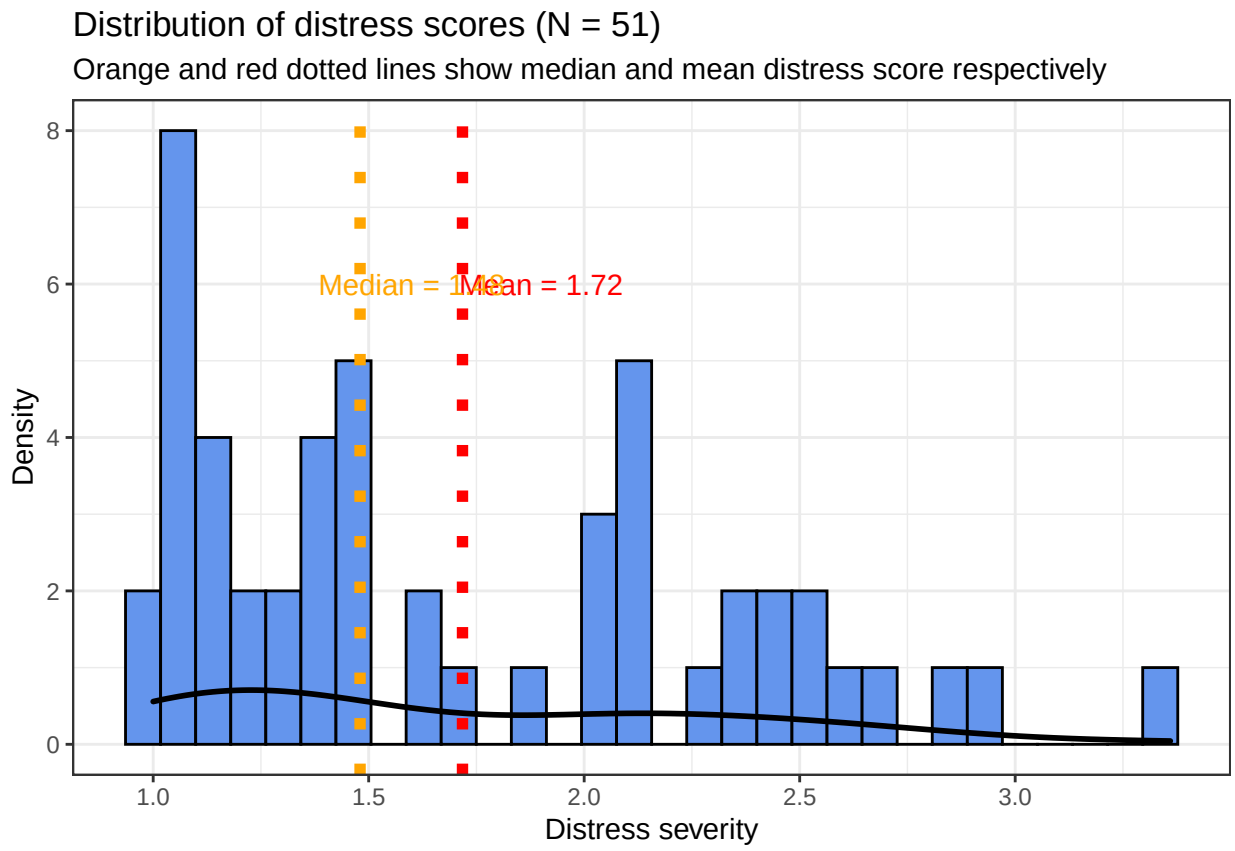
```
## `geom_smooth()` using formula 'y ~ x'
```




```
## we saw a difference in the PSD severity between groups; now
## check PSD data distribution
mean_psd <- mean(analyse$distress_tot) # mean psd
median_psd <- median(analyse$distress_tot) # median psd

ggplot(analyse) + aes(x = distress_tot) +
geom_histogram(fill = "#6495ed", col = "black") + geom_density(lwd =1) +
geom_vline(
    xintercept = mean_psd,
    col = "red",
    lwd = 2,
    linetype = 'dotted') +
  geom_vline(
    xintercept = median_psd,
    col = "orange",
    lwd = 2,
    linetype = 'dotted') +
  annotate("text",  # Add text for mean
           x = 1.9,
           y = 6,
           label = paste("Mean = 1.72"),
           col = "red",
           size = 4) +
  annotate("text",   # Add text for mean
           x = 1.6,
           y = 6,
```

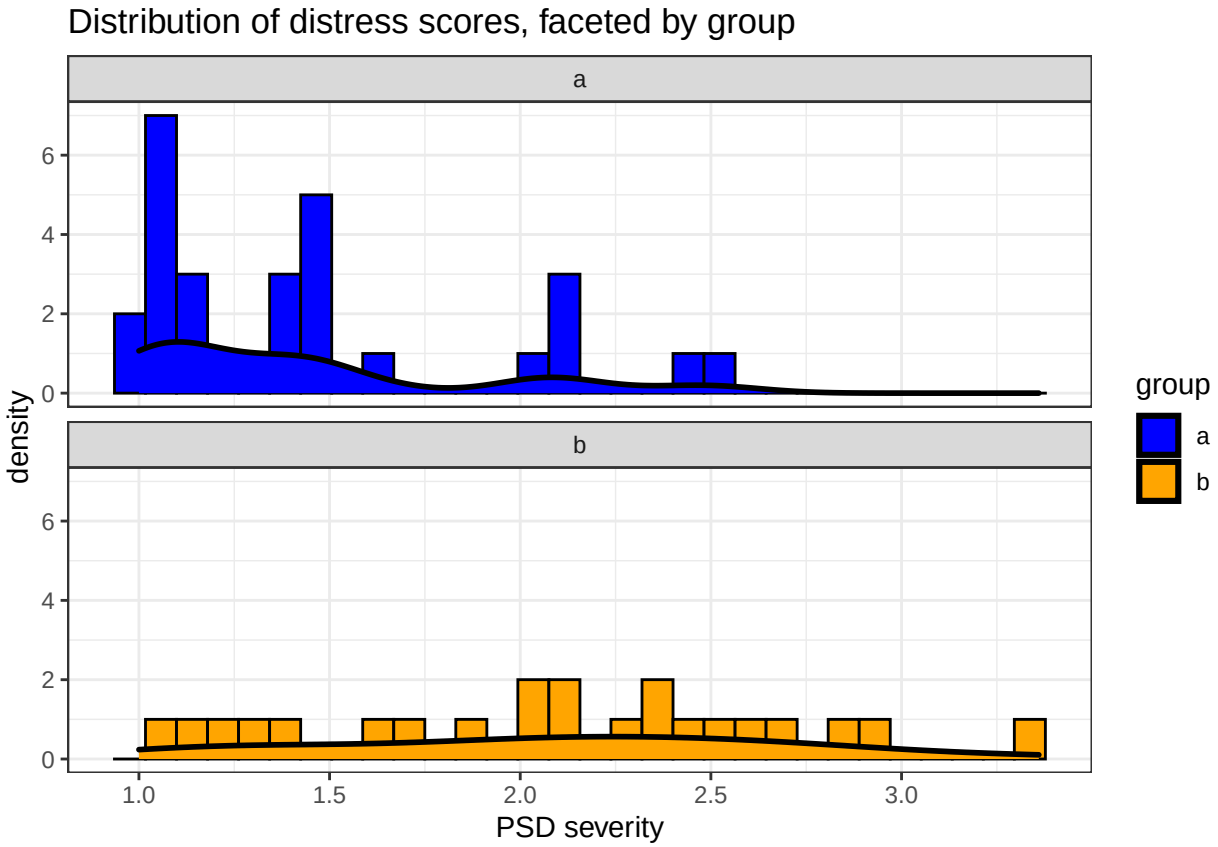
```
           label = paste("Median = 1.48"),
           col = "orange",
           size = 4) +
labs(title = "Distribution of distress scores (N = 51)",
     subtitle = "Orange and red dotted lines show median and mean distress score respectively",
x = "Distress severity", y = 'Density') + theme_bw()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
## PSD distribution by group
ggplot(analyse) + aes(x = distress_tot, fill = group) +
geom_histogram(colour = "black") + geom_density(lwd =1) +
  facet_wrap(~ group, nrow = 2) +
  scale_fill_manual(values = c('blue', 'orange')) +
labs(title = "Distribution of distress scores, faceted by group", x = "PSD severity") +
theme_bw()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
## calculate confidence interval of psd scores by groups
MeanCI(analyse$distress_tot) # CI by total participants option 1: DescTools package or option 2 below
```

```
mean   lwr.ci   upr.ci
```

1.717647 1.543805 1.891489

```
groupwiseMean(distress_tot ~ group, data = analyse) %>%
kbl(caption = "Distress severity scores by group", booktabs = T) %>%
kable_styling(latex_options = c("striped", "HOLD_position"))
```

Table 4: Distress severity scores by group

group	n	Mean	Conf.level	Trad.lower	Trad.upper
a	29	1.44	0.95	1.27	1.61
b	22	2.09	0.95	1.81	2.36

Summaries on calibrated current, rating of HFS trains, and PSD severity

- Plots suggested no difference between current (i) between groups, (ii) across sex and (iii) age.

- Plots suggested no difference between ratings given to HFS trains (i) between groups and (ii) across sex. However, our plots suggest a strong relationship difference between ratings are given to HFS trains and (iii) age. This relationship was parallel between groups and stronger in group b than in group a. Formal testing confirmed that 'age' predicted ratings given to HFS trains (p =0.02); however, this relationship was not different between groups (p = 017). Even so, model assumptions of the posterior predictive check and normality of residuals may be a problem for this model (see plots above).
- Plots suggested a relationship between PSD severity between groups, and this relationship seems stronger in group b than in a. No relationship was observed between PSD and sex. Our plots suggest a weak negative relationship between PSD and age, which was weaker in group a than in b. For PSD severity, we are 95% confident that the population mean (1.717) falls between 1.54 and 1.89. The mean difference between groups in PSD severity score is tabulated in **Table 4** abobe.

Blinding check

Here, we wanted to know whether the participants and Researcher 2 remained blinded. We calculated and reported the percentage of participants who correctly guessed the aim of the study and the number of participants for whom Researcher 2 accurately guessed group membership. Due to a technical error, the 'neutral' option was missing from the assessor blinding assessment. Therefore, blind assessment was performed on a four-point Linkert scale of "not at all confident", "not confident", "confident", and "extremely confident". Given the two-alternative forced choice design, 50% accuracy in Researcher 2's guesses about participants' group membership would represent chance accuracy, not blinding failure (55). However, no recommendations could be found to guide an acceptable threshold for the variance around this 50% value, outside which the data should be interpreted as indicating broken blinding. Therefore, instead of using a purely descriptive method to handle these data, we used two statistical methods to assess the blinding of Researcher 2 to group membership and planned to conduct sensitivity analyses in the event of broken blinding.

James' blinding index (BI) and the Chi-square goodness of fit test were used to assess blinding to group membership. The James' BI is a variation of the kappa coefficient, which is sensitive to the degree of 'disagreement' and not 'agreement' and prioritises 'do not know' responses. It is designed for use in studies that use a 3-alternative choice design including a 'don't know' option. However, our study lacked that option because we wanted to force Researcher 2 to give her best guess of each participant's group membership, rather than allowing a convenient 'don't know' response. Therefore, in our analysis, the number of 'don't know' responses were assigned a weight of 0. Accurate guesses were assigned a weight of 0, and inaccurate guesses were assigned a weight of 0.5, as previously described (Bang, Ni & Davis, 2004). If the BI is equal to 0, all guesses by Researcher 2 were accurate (Bang, Ni & Davis, 2004). If the BI is equal to 0.5, half of the guesses made by Researcher 2 were accurate, while half were inaccurate, which would mean random guessing (Bang, Ni & Davis, 2004). If the upper limit of the two-sided confidence interval is less than 0.5, then Researcher 2 was unblinded. Otherwise, we concluded that there was insufficient evidence to validate the blinding of Researcher 2 (Bang, Ni & Davis, 2004).

Given the limited match between our two-alternative forced choice design and the 3-alternative design anticipated by James' Blinding index, we supplemented that analysis with a Chi-square goodness of fit test. We used the Chi-square goodness of fit test to compare the observed distribution to the expected distribution by random chance (at 50% for each group) given by Researcher 2's guesses to participants group membership (Tezel et al., 2021). Together, these two tests would provide information on the breaking of blinding at the group level, but they shed no light on the potential influence on the study results if blinding were broken for a small number of participants. To test for this possible influence, we took a very conservative approach and identified participants for whom Researcher 2 has accurately guessed the group and also reported confidence of 4 (i.e. confident) or 5 (i.e. extremely confident), and conducted sensitivity analyses to examine the influence of cases in which blinding was possibly broken on the findings of the primary and secondary analyses. To achieve this, we repeated the primary and secondary analyses with data from these participants omitted.

```r
## were participants blinded to the aim of the study?
blinded <- analyse %>%  dplyr::select(pid, participants_blinded)
# note one participant does not have blinding data (missing post procedure data)

total_blinded <- sum(blinded$participants_blinded == 'y', na.rm = TRUE)
percentage_blinded <- round((total_blinded / analysed.n) * 100, 0)
glue("{total_blinded} of {analysed.n} ({percentage_blinded}%) participants were blinded; one participan
```

```
## 50 of 51 (98%) participants were blinded; one participant's blinding data were lost due to unanticipa
```

```r
## was Researcher 2 blinded to the pain status of each participant?
assessor_blinding <- analyse %>%
  dplyr::select(pid,
         group,
         participants_blinded,
         assessor_group_guess,
         assessor_confidence) %>%
  unique()

assessor_blinding %<>%
mutate(accuracy = if_else(group == assessor_group_guess, 'accurate','inaccurate'))

## James' BI has four possible combinations to blinding outcomes,
## filter by each condition and calculate the value for each.

# condition 1: actual group (a) vs guessed group (a)
accurate_group_a <- assessor_blinding %>%
  dplyr::select(pid, group, assessor_group_guess, accuracy) %>%
  filter(group == "a") %>%
  filter(accuracy == "accurate")
paste0("Researcher 2 guessed correctly for ",nrow(accurate_group_a)," participants in the a group.")
```

```
## [1] "Researcher 2 guessed correctly for 15 participants in the a group."
```

```r
# condition 2: actual group (a) vs guessed group (b)
inaccurate_group_a <- assessor_blinding %>%
  dplyr::select(pid, group, assessor_group_guess, accuracy) %>%
  filter(group == "a") %>%
  filter(accuracy == "inaccurate")
paste0("Researcher 2 guessed inaccurately for ", nrow(inaccurate_group_a)," participants in the a group
```

```
## [1] "Researcher 2 guessed inaccurately for 14 participants in the a group."
```

```r
# condition 3: actual group (b) vs guessed group (b)
accurate_group_b <- assessor_blinding %>%
  dplyr::select(pid, group, assessor_group_guess, accuracy) %>%
  filter(group == "b") %>%
  filter(accuracy == "accurate")
paste0("Researcher 2 guessed accurately for ", nrow(accurate_group_b)," participants in the b group.")
```

```
## [1] "Researcher 2 guessed accurately for 8 participants in the b group."
```

```r
# condition 4: actual group (b) vs guessed group (pain)
inaccurate_group_b <- assessor_blinding %>%
  dplyr::select(pid, group, assessor_group_guess, accuracy) %>%
  filter(group == "b") %>%
  filter(accuracy == "inaccurate")
paste0("Researcher 2 guessed inaccurately for ",nrow(inaccurate_group_b)," participants in the b group.
```

```
## [1] "Researcher 2 guessed inaccurately for 13 participants in the b group."
```

```r
# generate table based on the above conditions
table_blind_con <- matrix(
  data = c(
    nrow(accurate_group_a),
    nrow(inaccurate_group_a),
    nrow(inaccurate_group_b),
    nrow(accurate_group_b),
    0, # DK
    0),# DK
  nrow = 3,
  ncol = 2,
  byrow = TRUE,
  dimnames = list(c('a', 'b', 'DK'), # DK = Don't know group = 0
                  c('a', 'b')))

table_blind_con %>%
  kbl(caption = "Researcher 2's blinding outcome to group membership", booktabs = T) %>%
  kable_styling(latex_options = c("striped", "HOLD_position"))
```

Table 5: Researcher 2's blinding outcome to group membership

	a	b
a	15	14
b	13	8
DK	0	0

```r
BI(table_blind_con)$JamesBI
```

```
##         Estimate Std. Error 95% LCL (2-Sided) 95% UCL (2-Sided)
## Overall 0.550571 0.06996293         0.4134461         0.6876958
```

```r
# BI= 0.55 and Upper confidence level = 0.69

## Chi-square goodness of fit
# Requires separate testing of each group (i.e., a and b)

# a group
chisq.test(table_blind_con)
```

```
##
```

```
##  Pearson's Chi-squared test
##
## data:  table_blind_con
## X-squared = NaN, df = 2, p-value = NA
```

```
chisq.test(x = c(nrow(accurate_group_a), nrow(inaccurate_group_a)), # counts in guessed 'a' group
           p = c(0.5, 0.5)) # Probabilities assuming random guessing.
```

```
##
##  Chi-squared test for given probabilities
##
## data:  c(nrow(accurate_group_a), nrow(inaccurate_group_a))
## X-squared = 0.034483, df = 1, p-value = 0.8527
```

```
# X2 = 0.03, Probability lies approximately at 0.5; null hypothesis is upheld?
#https://www.bmj.com/sites/default/files/attachments/resources/2011/08/appendix-table-c.pdf

# b group
chisq.test(x = c(nrow(accurate_group_b), nrow(inaccurate_group_b)),  # counts in guessed 'b' group
           p = c(0.5, 0.5))
```

```
##
##  Chi-squared test for given probabilities
##
## data:  c(nrow(accurate_group_b), nrow(inaccurate_group_b))
## X-squared = 1.1905, df = 1, p-value = 0.2752
```

```
# X2 = 1.19, Probability lies between 0.10 and 0.5; null hypothesis is upheld.

# two groups, therefore df = 1. With df = 1, the X2 value must exceed critical value
# to reach statistical significance. If the X2 value is less than this, p will be > 0.05,
# indicating that the distribution does not significantly deviate from the distribution
# that would be expected given chance guessing (which would be ~50% accuracy).

## determine the number of accurate responses

began_procedure <- began_procedure  %>% mutate(pid = as.character( pid))
assessor_blinding %<>%
right_join(began_procedure) %>% filter(
grepl("completed", participation_status) |
grepl("missing post-procedure data",participation_status)
) %>%
  dplyr::select(pid, group,participants_blinded, assessor_group_guess, assessor_confidence, accuracy, pa
unique()
```

```
## Joining, by = c("pid", "group", "participants_blinded", "assessor_group_guess",
## "assessor_confidence")
```

```
accurate.guess <-
  sum(assessor_blinding$accuracy == 'accurate', na.rm = TRUE)
guesses_made <- assessor_blinding %>%
  filter(accuracy != "NA") %>%
```

```
  nrow()
paste0( "Researcher 2 accurately guessed the group membership of ",accurate.guess," out of ", guesses_ma

## [1] "Researcher 2 accurately guessed the group membership of 23 out of 50 participants"

# there is a 50% chance that Researcher 2 can accurately guess the participant's group membership;
# however, this does not necessarily demonstrate blinding failure.
# therefore, we will quantify the accurate guesses for which the researcher
# reported a confidence interval of 4 (i.e. confident) and 5 (i.e. extremely confident)

## filter by accurate guesses and tabulate the guesses for which
## Researcher 2 was "confident" and "extremely confident"
assessor_blinding %>%
  dplyr::select(pid, assessor_group_guess, assessor_confidence, accuracy) %>%
  filter(accuracy == "accurate") %>%
  filter(assessor_confidence >= "4") %>%
  kbl(caption = "Participants for whom the Researcher 2 accurately and confidently guessed group members
  kable_styling(latex_options = c("striped","HOLD_position"))
```

Table 6: Participants for whom the Researcher 2 accurately and confidently guessed group membership

pid	assessor_group_guess	assessor_confidence	accuracy
320	a	4	accurate
497	a	4	accurate
544	a	4	accurate
556	b	4	accurate
561	b	4	accurate

```
# Researcher 2 accurately guessed group membership for 5 participants (3 in group a and 2 in group b),

## filter by inaccurate guesses and tabulate the guesses for which
## the researcher was "confident"  (4) and "extremely confident" (5)
assessor_blinding %>%
  dplyr::select(pid, assessor_group_guess, assessor_confidence, accuracy) %>%
  filter(accuracy == "inaccurate") %>%
  filter(assessor_confidence >= "4") %>%
  kbl(caption = "Participants for whom the Researcher 2 inaccurately but confidently guessed group membe
  kable_styling(latex_options = c("striped", "HOLD_position"))
```

Table 7: Participants for whom the Researcher 2 inaccurately but confidently guessed group membership

pid	assessor_group_guess	assessor_confidence	accuracy
140	a	4	inaccurate
156	a	4	inaccurate
159	a	4	inaccurate
197	b	4	inaccurate
268	a	4	inaccurate
311	b	4	inaccurate
383	b	4	inaccurate
396	a	4	inaccurate
425	b	4	inaccurate
442	b	4	inaccurate
598	a	4	inaccurate

```
# Researcher 2 inaccurately but confidently guessed group membership for 11 participants (6 in group a)
```

Blinding outcome summary

Due to an unexpected power failure, blinding check data were not saved for one participant. Therefore, blinding data were available for 50 participants, all of whom (100%) were judged to be blinded to the aim of the study. Researcher 2 accurately guessed the group membership of 23 out of 50 participants. The James' BI was 0.55 (upper limit: 0.69) which confirms the blinding of Researcher 2 to participant group membership and that the researcher accurately guessed group membership for almost half of the participants meaning that guessing was random. This finding is also supported by the Chi-squared goodness of fit test conducted for both groups (group a: 0.03; group b: X2= 1.19). Researcher 2 accurately and confidently guessed group membership for five participants (three in group a). Researcher 2 inaccurately but confidently guessed group membership for 11 participants (six in group a).

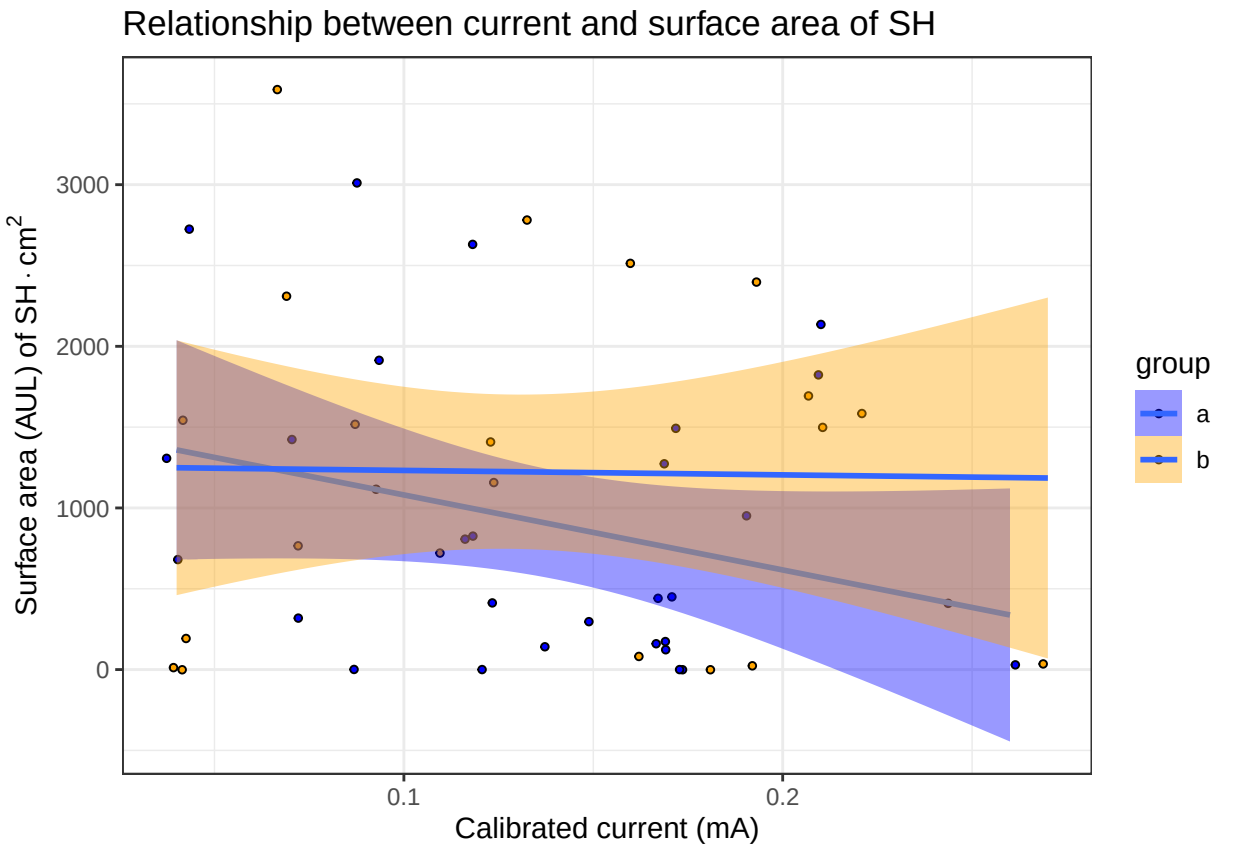
Confounding variables

We anticipated that three variables could confound our results. First, the current used for the SH induction was calibrated for each participant, and could influence the magnitude or area of SH. Second, the painfulness of the HFS induction typically varies between individuals, and could influence the magnitude or area of SH. For these two potential confounding variables, we tested for an association between the potential confounding variable and each SH outcome and planned to include the potential confounding variable as a covariate in the analyses of the SH outcomes if a significant association was found. Relationships between the confounders and each study outcome were formally tested using Kendall's test due to the distribution of our data and the small sample size. Third, the delay between the assessment of self-reported PSD and the induction and assessment of SH (which affected only a few participants) could result in the PSD values poorly representing the participant's state at the time of the SH induction and assessment. Therefore, we wanted to assess the influence of days between PSD and SH assessments on the relationship between PSD and the SH outcomes. For this, we assessed whether including the days between assessments as a covariate improved model fit in either the primary of secondary analysis.

```
## plot and test for a correlation between the surface area
## of SH and confounding variables (1. current, 2. HFS
## pain, 3. delay between PSD assessments and HFS
## induction)
```

```
## plot surface area (AUL) and current
main_outcome_data %>%
  filter(grepl("completed", participation_status) | grepl("missing post-procedure data",
    participation_status)) %>%
  dplyr::select(pid, group, sex, age, calibrated_current, aul_sa) %>%
  unique() %>%
  ggplot(data = .) + aes(x = calibrated_current, y = aul_sa,
  fill = group) + geom_jitter(shape = 21, size = 1) + scale_fill_manual(values = c("blue",
  "orange")) + labs(title = "Relationship between current and surface area of SH ",
  x = "Calibrated current (mA)", y = expression("Surface area (AUL) of SH" %.%
    cm^{
      2
    })) + theme_bw() + geom_smooth(method = lm)
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```
# no relationship between current and surface area in group
# b but there might be a negative relationship in group a

# formally check correlation using kendall's test.
cor_sa_current <- main_outcome_data %>%
  filter(grepl("completed", participation_status) | grepl("missing post-procedure data",
    participation_status)) %>%
  dplyr::select(pid, group, sex, age, calibrated_current, aul_sa) %>%
```

```
  unique()

cor.test(x = cor_sa_current$calibrated_current, y = cor_sa_current$aul_sa,
  method = "kendall")  # no relationship between current and surface area (p =0.36)
```

```
##
##  Kendall's rank correlation tau
##
## data:  cor_sa_current$calibrated_current and cor_sa_current$aul_sa
## z = -0.90922, p-value = 0.3632
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##         tau
## -0.09153818
```

```
# plot suggests weak negative relationship between current
# and surface area for group a lets formally test this to
# check group a
cor_sa_current_a <- main_outcome_data %>%
  filter(grepl("completed", participation_status) | grepl("missing post-procedure data",
    participation_status)) %>%
  filter(group == "a") %>%
  dplyr::select(pid, group, sex, age, calibrated_current, aul_sa) %>%
  unique()

cor.test(x = cor_sa_current_a$calibrated_current, y = cor_sa_current_a$aul_sa,
  method = "kendall")  # tau = -0.16, weak relationship which was not significant (p = 0.2)
```

```
## Warning in cor.test.default(x = cor_sa_current_a$calibrated_current, y =
## cor_sa_current_a$aul_sa, : Cannot compute exact p-value with ties
```

```
##
##  Kendall's rank correlation tau
##
## data:  cor_sa_current_a$calibrated_current and cor_sa_current_a$aul_sa
## z = -1.1704, p-value = 0.2418
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##        tau
## -0.1625655
```

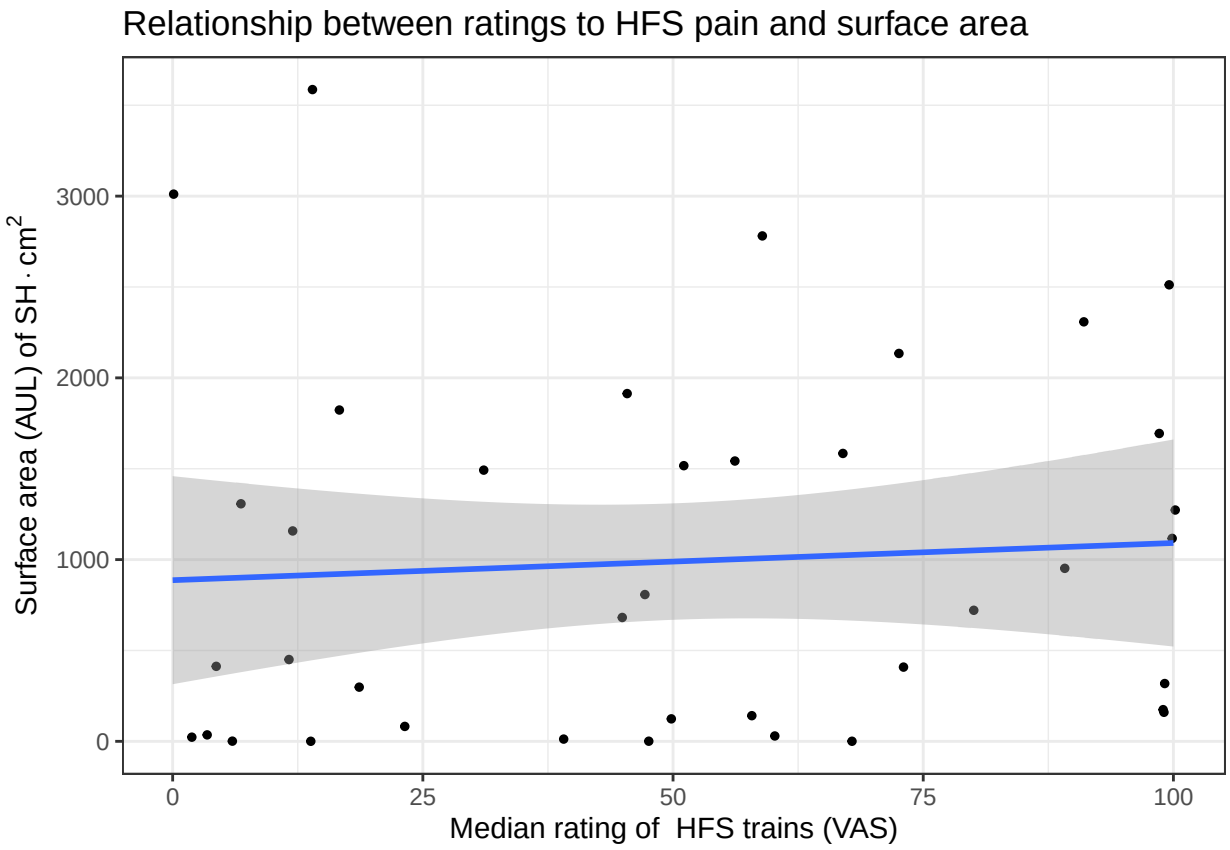
```
## plot painfulness of the HFS induction and surface area
## (AUL)
hfs_pain %>%
  dplyr::select(pid, aul_sa, train_counter, train_rating) %>%
  unique() %>%
  group_by(pid, aul_sa) %>%
  summarise(median_rating = median(train_rating, na.rm = TRUE)) %>%
  ggplot(data = .) + aes(x = median_rating, y = aul_sa) + geom_jitter(shape = 21,
  size = 1, fill = "black") + labs(title = "Relationship between ratings to HFS pain and surface area "
  x = "Median rating of  HFS trains (VAS)", y = expression("Surface area (AUL) of SH" %.%
    cm^{
```

```
    2
})) + theme_bw() + geom_smooth(method = lm)
```

```
## `summarise()` has grouped output by 'pid'. You can override using the `.groups`
## argument.
## `geom_smooth()` using formula 'y ~ x'
```

```
## Warning: Removed 12 rows containing non-finite values (stat_smooth).
```

```
## Warning: Removed 12 rows containing missing values (geom_point).
```




```
# no relationship betwteen ratings to HFS pain and surface
# area SH

## formally check correlation using Kendall's tests
cor_sa_hfs <- hfs_pain %>%
  dplyr::select(pid, aul_sa, train_counter, train_rating) %>%
  unique() %>%
  group_by(pid, aul_sa) %>%
  summarise(median_rating = median(train_rating, na.rm = TRUE))
```

```
## `summarise()` has grouped output by 'pid'. You can override using the `.groups`
## argument.
```

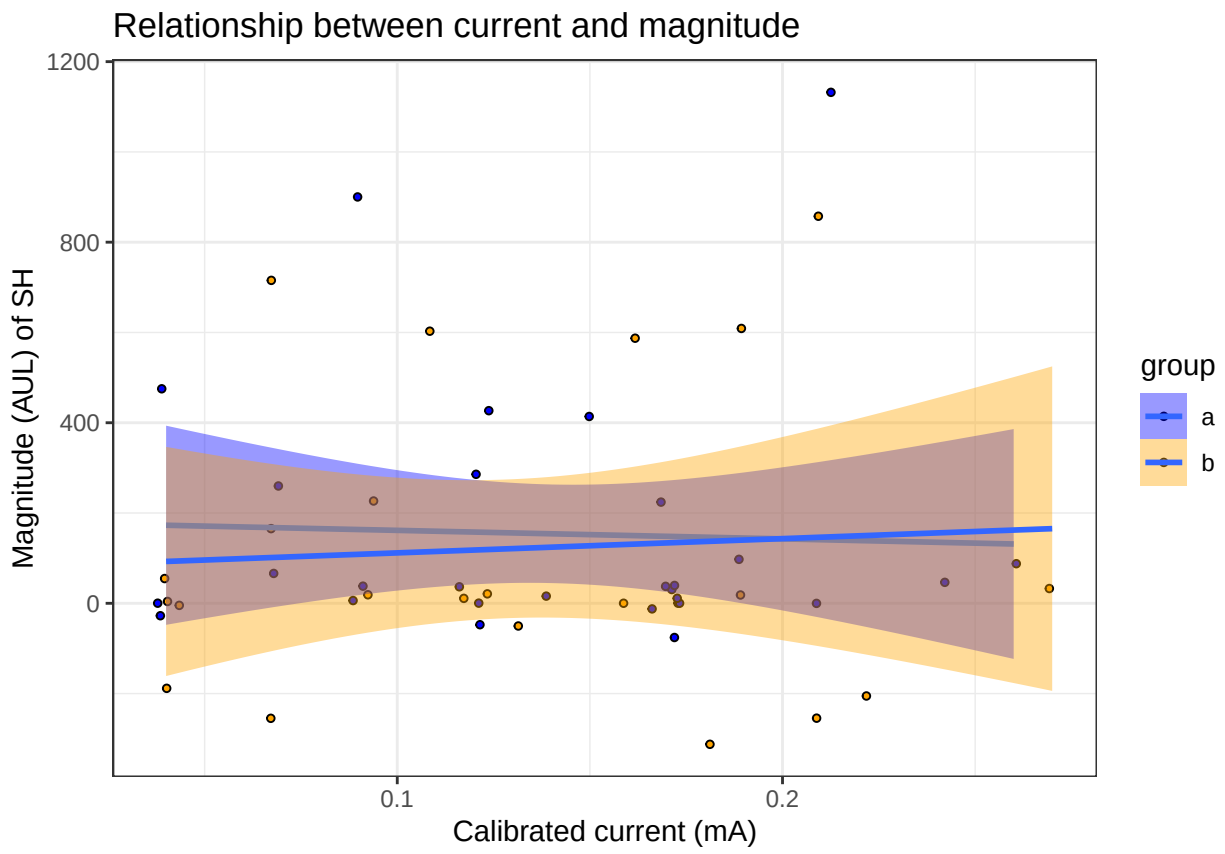
```r
cor.test(x = cor_sa_hfs$median_rating, y = cor_sa_hfs$aul_sa,
  method = "kendall")
```

```
## Warning in cor.test.default(x = cor_sa_hfs$median_rating, y =
## cor_sa_hfs$aul_sa, : Cannot compute exact p-value with ties
```

```
##
##  Kendall's rank correlation tau
##
## data:  cor_sa_hfs$median_rating and cor_sa_hfs$aul_sa
## z = 1.1867, p-value = 0.2353
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##       tau
## 0.1332427
```

```r
# tau= 0.13; positive weak relationship which was not
# significant (p =0.23)


## plot and test for a correlation between the magnitude of
## SH and confounding variables (1. current, and 2. ratings
## to HFS pain )

## plot magnitude (AUL) and current
main_outcome_data %>%
  filter(grepl("completed", participation_status) | grepl("missing post-procedure data",
    participation_status)) %>%
  dplyr::select(pid, group, sex, age, calibrated_current, aul_mag) %>%
  unique() %>%
  ggplot(data = .) + aes(x = calibrated_current, y = aul_mag,
  fill = group) + scale_fill_manual(values = c("blue", "orange")) +
  geom_jitter(shape = 21, size = 1) + labs(title = "Relationship between current and magnitude",
  x = "Calibrated current (mA)", y = "Magnitude (AUL) of SH") +
  theme_bw() + geom_smooth(method = lm)
```

```
## `geom_smooth()` using formula 'y ~ x'
```



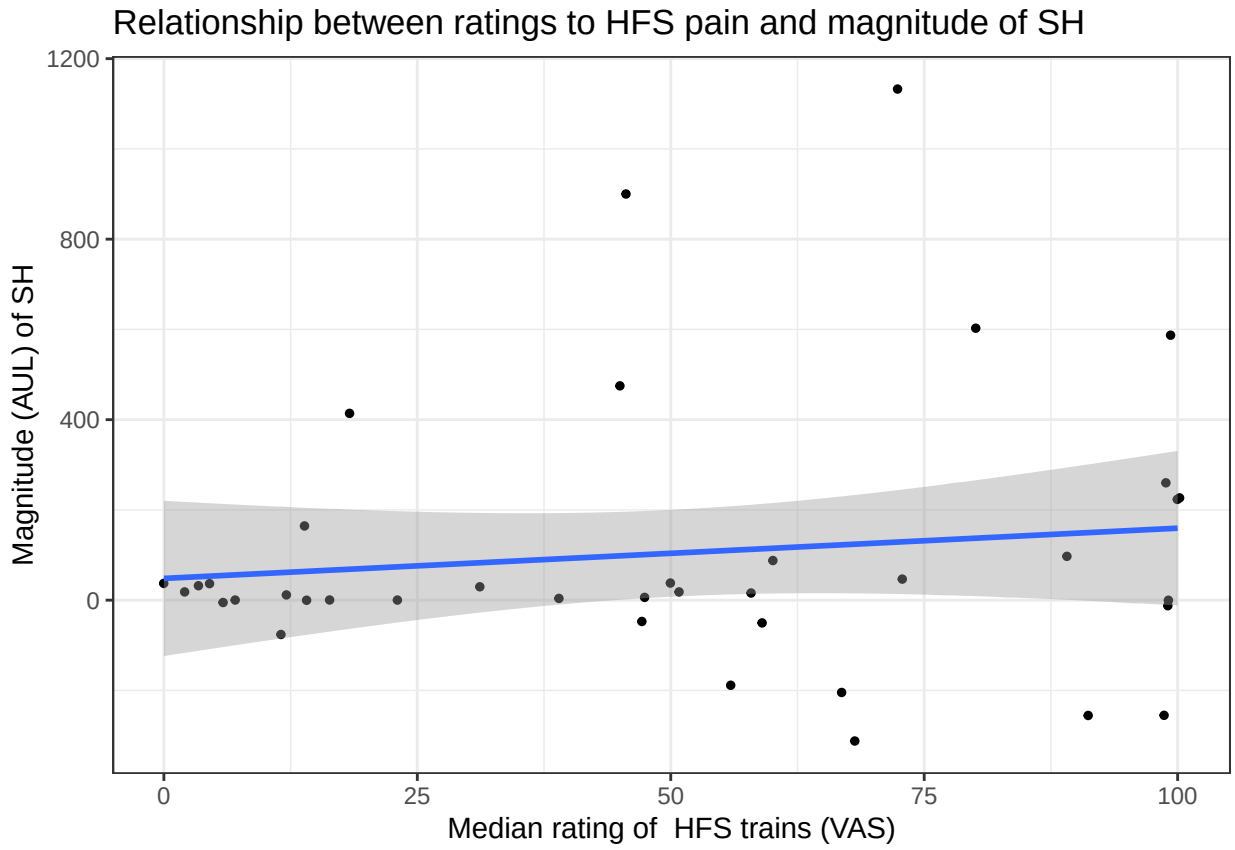

```
# no relationship between current and magnitude

# formally check correlation, using Kendall's tests
cor_mag_current <- main_outcome_data %>%
  filter(grepl("completed", participation_status) | grepl("missing post-procedure data",
    participation_status)) %>%
  dplyr::select(pid, group, sex, age, calibrated_current, aul_mag) %>%
  unique()

cor.test(x = cor_mag_current$calibrated_current, y = cor_mag_current$aul_mag,
  method = "kendall")
```

```
##
##  Kendall's rank correlation tau
##
## data:  cor_mag_current$calibrated_current and cor_mag_current$aul_mag
## z = -0.040961, p-value = 0.9673
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##          tau
## -0.004129902
```

```
# -0.004129902 very weak negative relationship which was
# not significant (p=0.97)
```

```
## plot magnitude (AUL) and painfulness of the HFS
## induction.
hfs_pain %>%
  dplyr::select(pid, aul_mag, train_counter, train_rating) %>%
  unique() %>%
  group_by(pid, aul_mag) %>%
  summarise(median_rating = median(train_rating, na.rm = TRUE)) %>%
  ggplot(data = .) + aes(x = median_rating, y = aul_mag) +
  geom_jitter(shape = 21, size = 1, fill = "black") + labs(title = "Relationship between ratings to HFS
  x = "Median rating of  HFS trains (VAS)", y = "Magnitude (AUL) of SH") +
  theme_bw() + geom_smooth(method = lm)
```

```
## `summarise()` has grouped output by 'pid'. You can override using the `.groups`
## argument.
## `geom_smooth()` using formula 'y ~ x'

## Warning: Removed 12 rows containing non-finite values (stat_smooth).
## Removed 12 rows containing missing values (geom_point).
```




```
# no relationship between hfs painfulness and magnitude

# formally check correlation using Kendall's tests
cor_mag_hfs <- hfs_pain %>%
  dplyr::select(pid, aul_mag, train_counter, train_rating) %>%
```

```
  unique() %>%
  group_by(pid, aul_mag) %>%
  summarise(median_rating = median(train_rating, na.rm = TRUE))
```

```
## `summarise()` has grouped output by 'pid'. You can override using the `.groups`
## argument.
```

```
cor.test(x = cor_mag_hfs$median_rating, y = cor_mag_hfs$aul_mag,
  method = "kendall")
```

```
## Warning in cor.test.default(x = cor_mag_hfs$median_rating, y =
## cor_mag_hfs$aul_mag, : Cannot compute exact p-value with ties
```

```
##
##  Kendall's rank correlation tau
##
## data:  cor_mag_hfs$median_rating and cor_mag_hfs$aul_mag
## z = 0.52113, p-value = 0.6023
## alternative hypothesis: true tau is not equal to 0
## sample estimates:
##        tau
## 0.05874405
```

```
# 0.05874405 very weak relationship which was not
# significant (p = 0.60)
```

Confounding variables summary We anticipated three possible confounding variables: (i) calibrated current and (ii) painfulness of the HFS induction. The correlation tests showed no statistically significant relationship between the surface area of experimentally induced SH and any confounding variable (in (i) tau= -0.09, p= 0.36; and (ii) tau = 0.13, p =0.24. Similarly, the correlation test showed no relationship between the magnitude of experimentally induced SH and either confounding variable (in (i) tau= -0.004, p = 0.97; and (ii) tau = 0.06, p= 0.60. Therefore, neither of these covariates was included in the analysis of the surface area and magnitude of SH. (iii) was carried forward to primary and secondary analyses as a potential covariate.

Primary and secondary analyses

We will use three models to formally test the relationships between dependent variables (SH outcomes) and independent variables: (i) PSD, (ii) group (i.e., pain status), and (iii) delay in days between PSD assessment and SH induction and assessment (possible confounder). For each outcome, we will choose one model (where applicable) that fits our data, and we will use that model to test for the influence of (iii) because we anticipate that (iii) would influence relationships between PSD and SH outcomes. Finally, we will provide the estimated marginal means and assessments of model fit for the chosen model.

The primary analysis will use data on surface area (at 30, 45 and 60 minutes after induction), whereas the secondary analysis will use data on magnitude (at 35, 50 and 65 minutes after induction) of SH. We will use the AUL of each SH outcome data in scatter-and-line plots to visualise the relationship between PSD and SH outcomes, first for the whole sample, and then by group. To establish whether the experimentally induced SH is associated with the severity of PSD, we will (1) test whether the PSD (independent variable) predicts each outcome (dependent variable) using the data from both groups. If this model indicates that

each outcome is predicted by PSD, we will use the plots to guide decisions to potentially (2) test for an effect of group membership on each outcome by including group membership as a second independent variable, and (3) test for a difference in this relationship between PSD and group membership, by including an interaction between PSD and group in the model. We expect to interpret steps 2 and 3 as exploratory, since we are underpowered, especially to statistically detect an interaction effect. Below we tabulate descriptive statistics for secondary hyperalgesia outcomes (AUL)

```
outcomes_aul_table <- main_outcome_data %>% right_join(analyse) %>%
  dplyr::select(pid, group, aul_mag, aul_sa, participation_status) %>% unique()
```

```
## Joining, by = c("pid", "age", "sex", "group", "distress_tot",
## "calibrated_current", "participation_status")
```

```
controls_outcomes <-
  arsenal::tableby.control(
    numeric.stats = c("N", "mean", "sd", "median", "range"),
    cat_stats = c("countpct"),
    stats.labels = c(
      list(
        N = "number of participants",
        mean = "mean",
        standard_deviation = "sd",
        median = "median",
        range = "range")))
table_three <- tableby(group ~ aul_sa + aul_mag,
                       data = outcomes_aul_table,
                       control = controls_outcomes)
summary(table_three, text = TRUE) %>%
kbl(caption = "Descriptive statistics for secondary hyperalgesia outcomes (AUL)",
    booktabs = T) %>%
kable_styling(latex_options = c("striped", "HOLD_position"))
```

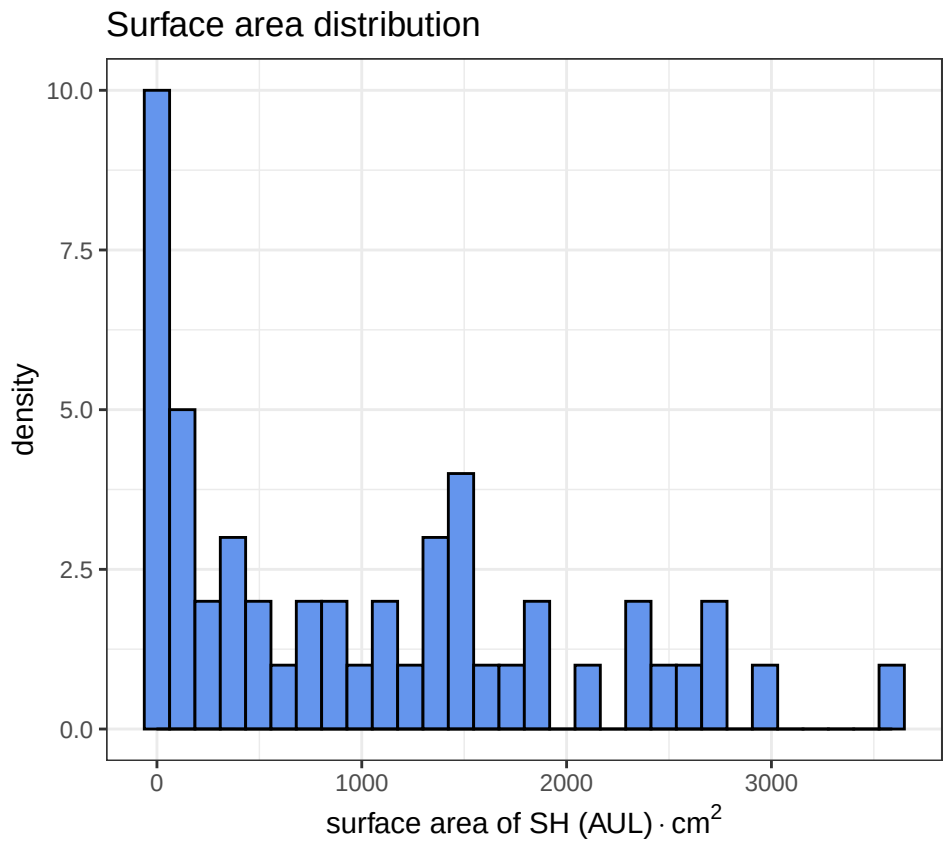
Table 8: Descriptive statistics for secondary hyperalgesia outcomes (AUL)

	a (N=29)	b (N=22)	Total (N=51)	p value
aul_sa				0.235
- number of participants	29	22	51	
- mean	894.948	1224.418	1037.072	
- SD	906.715	1047.365	973.798	
- median	450.622	1282.654	806.997	
- range	0.000 - 3010.039	0.000 - 3587.307	0.000 - 3587.307	
aul_mag				0.702
- number of participants	29	22	51	
- mean	154.009	120.568	139.583	
- SD	281.739	337.860	304.423	
- median	37.500	18.750	30.000	
- range	-76.250 - 1132.500	-312.500 - 857.500	-312.500 - 1132.500	

Surface area of SH

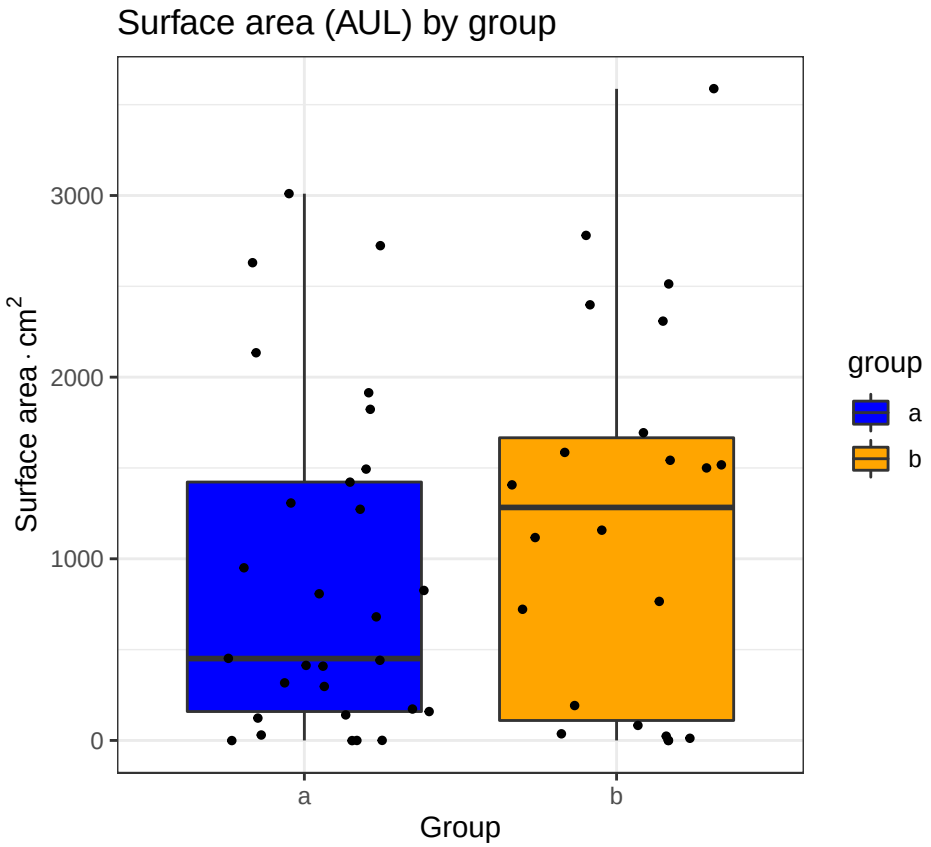
```
## prepare data frame to run model
main_outcome_data %<>%
  dplyr::select(pid, group, distress_tot, aul_sa, aul_mag) %>%
  unique() %>%
  mutate(group = factor(group))

## check distribution of surface area data (AUL) using
## histogram
ggplot(main_outcome_data) + aes(x = aul_sa) + geom_histogram(fill = "#6495ed",
  col = "black") + geom_density() + labs(title = "Surface area distribution",
  x = expression("surface area of SH (AUL)" %.% cm^{
    2
  })) + theme_bw()  # data clearly not normally distributed
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
# plot relationship between SA and PSD
sa_psd <- ggplot(main_outcome_data) + aes(x = distress_tot, y = aul_sa) +
  geom_point() + geom_smooth(method = "lm") + labs(title = "Relationship between surface area of SH and
  x = "PSD severity", y = expression("surface area of SH (AUL)" %.%
    cm^{
      2
```

```
    })) + theme_bw()

# plot relationship between SA and PSD by group
sa_psd_group <- ggplot(data = main_outcome_data) + aes(x = distress_tot,
  y = aul_sa, group = group, colour = group) + geom_point() +
  geom_smooth(method = "lm") + scale_colour_manual(values = c("blue",
  "orange")) + labs(title = "Relationship between surface area of SH and PSD severity, \nby group",
  x = "PSD severity", y = expression("surface area of SH (AUL)" %.%
    cm^{
      2
    })) + theme_bw()


# plot surface area aul between groups
sa_aul_groups <- ggplot(data = main_outcome_data) + aes(x = group,
  y = aul_sa, fill = group) + geom_boxplot() + geom_jitter(shape = 21,
  color = "black", fill = "black", size = 1) + scale_fill_manual(values = c("blue",
  "orange")) + labs(title = "Surface area (AUL) by group",
  x = "Group", y = expression("Surface area" %.% cm^{
    2
  })) + theme_bw()

print(sa_aul_groups)
```



```r
## test for main effect first, plot model and and examine
## marginal effects
sa_model_main <- lm(aul_sa ~ distress_tot, data = main_outcome_data)
summary(sa_model_main)
```

```
##
## Call:
## lm(formula = aul_sa ~ distress_tot, data = main_outcome_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1469.3  -698.4  -238.1   658.9  2161.1
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)     111.7      385.7   0.290   0.7732
## distress_tot    538.7      211.5   2.547   0.0141 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 924.4 on 49 degrees of freedom
## Multiple R-squared:  0.1169, Adjusted R-squared:  0.0989
## F-statistic: 6.488 on 1 and 49 DF,  p-value: 0.01405
```

```r
df1 <- ggpredict(sa_model_main, terms = "distress_tot")
print(df1)
```

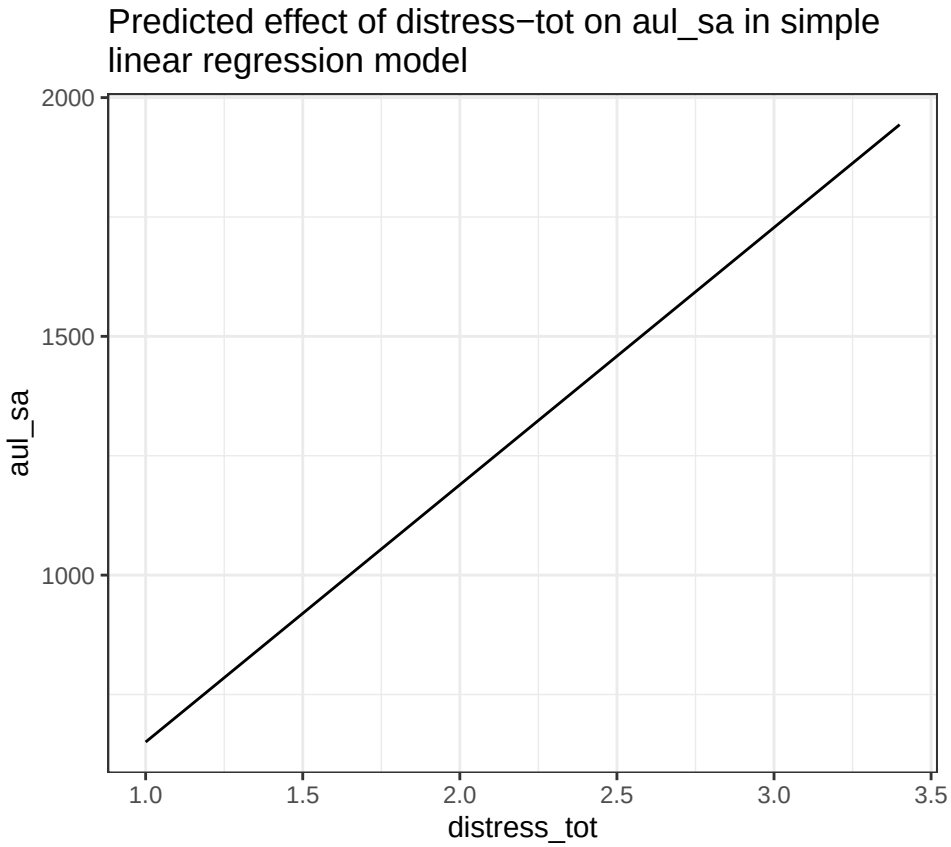
```
## # Predicted values of aul_sa
##
## distress_tot | Predicted |             95% CI
## ------------------------------------------------
##         1.00 |    650.46 | [ 259.48, 1041.44]
##         1.20 |    758.20 | [ 425.93, 1090.48]
##         1.60 |    973.69 | [ 715.35, 1232.04]
##         2.00 |   1189.18 | [ 909.79, 1468.58]
##         2.20 |   1296.93 | [ 973.90, 1619.95]
##         2.40 |   1404.67 | [1024.71, 1784.63]
##         2.80 |   1620.16 | [1104.73, 2135.59]
##         3.40 |   1943.39 | [1201.28, 2685.50]
```

```r
# df1_table <-
df1 %>%
  kbl(col.names = c("PSD", "Predicted surface area", "standard error",
    "CI(lower bound)", "CI(upper bound)", names(df1)[-1:-5]),
    caption = "Estimated marginal means: predicted effect of PSD on the surface area of SH in Model 1",
    booktabs = T) %>%
  kable_styling(latex_options = c("striped", "HOLD_position"))
```

Table 9: Estimated marginal means: predicted effect of PSD on the surface area of SH in Model 1

PSD	Predicted surface area	standard error	CI(lower bound)	CI(upper bound)	group
1.0	650.4603	199.4820	259.4827	1041.438	1
1.2	758.2046	169.5329	425.9262	1090.483	1
1.4	865.9489	145.8367	580.1141	1151.784	1
1.6	973.6932	131.8103	715.3498	1232.037	1
1.8	1081.4375	130.6070	825.4525	1337.423	1
2.0	1189.1818	142.5520	909.7851	1468.579	1
2.2	1296.9261	164.8108	973.9030	1619.949	1
2.4	1404.6704	193.8627	1024.7065	1784.634	1
2.6	1512.4147	227.1157	1067.2761	1957.553	1
2.8	1620.1590	262.9810	1104.7258	2135.592	1
3.0	1727.9033	300.5247	1138.8858	2316.921	1
3.2	1835.6476	339.1899	1170.8476	2500.448	1
3.4	1943.3919	378.6333	1201.2844	2685.500	1

```
# print(df1_table) remove_column(df1_table, c(3,6)) #
# remove unwanted columns; NOTE: Cannot print in latex

ggplot(df1, aes(x, predicted)) + geom_line() + labs(x = get_x_title(df1),
  y = get_y_title(df1), title = get_title(df1)) + theme_bw() +
  labs(title = "Predicted effect of distress-tot on aul_sa in simple\nlinear regression model")
```

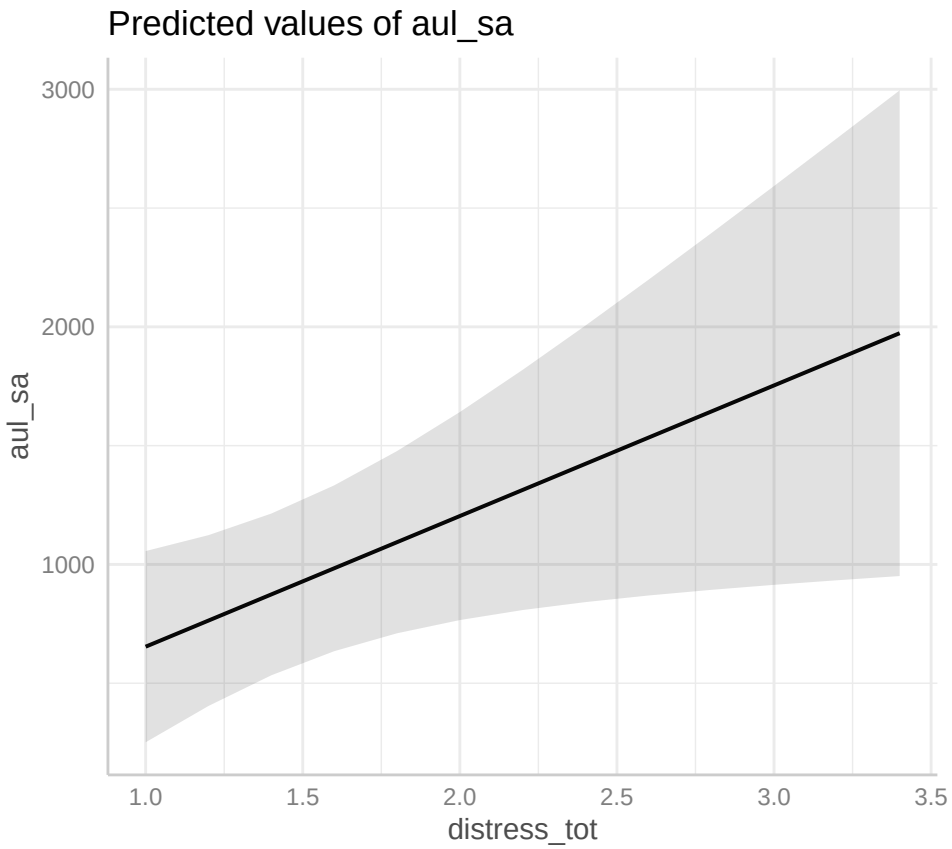


```r
# in the above we can see the the slope for the regression
# line is consistent for each value of our predictor
# (distress_tot); however, the value for the outcome is not
# the same for each value of the predictor

## Test for group effect and examine marginal effects
sa_model_multiple <- lm(aul_sa ~ distress_tot + group, data = main_outcome_data)
summary(sa_model_multiple)
```

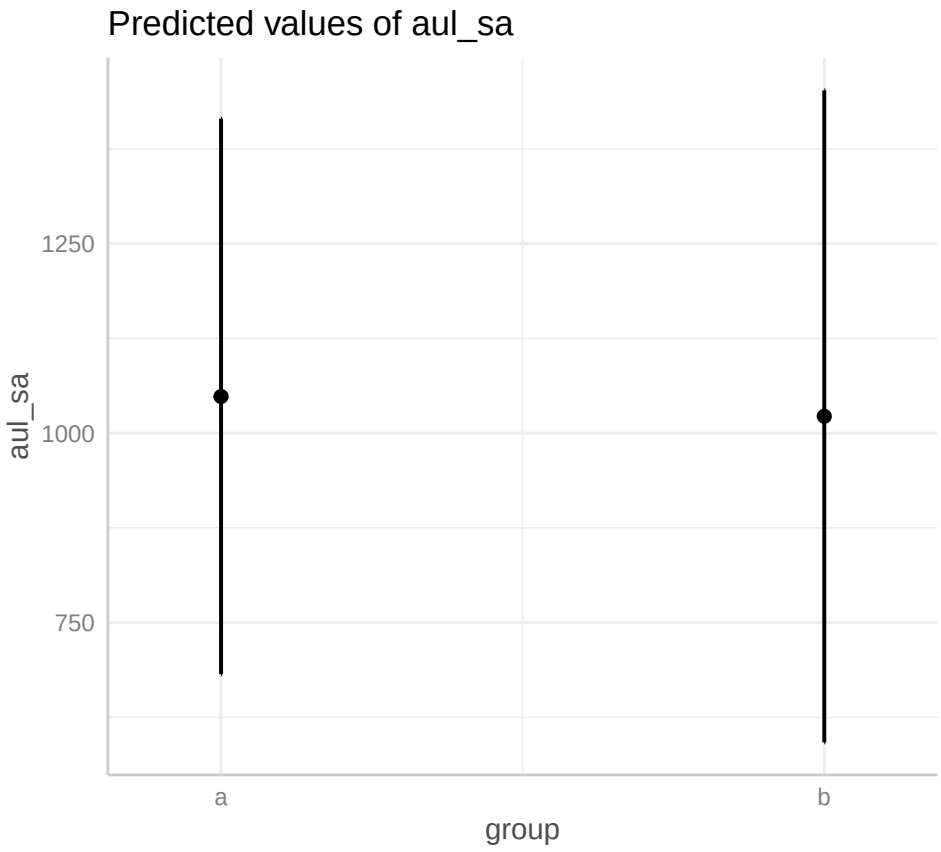
```
##
## Call:
## lm(formula = aul_sa ~ distress_tot + group, data = main_outcome_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1489.5  -689.8  -237.1   649.4  2168.0
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    103.99     400.29   0.260   0.7961
## distress_tot   549.80     250.78   2.192   0.0332 *
## groupb         -26.16     309.89  -0.084   0.9331
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 933.9 on 48 degrees of freedom
## Multiple R-squared:  0.1171, Adjusted R-squared:  0.08027
## F-statistic: 3.182 on 2 and 48 DF,  p-value: 0.0504
```

```r
df2 <- ggpredict(sa_model_multiple)
plot(df2)
```

```
## $distress_tot
```


Predicted values of aul_sa
3000
2000
1000
aul_sa
1.0
1.5
2.0
2.5
3.0
3.5
distress_tot

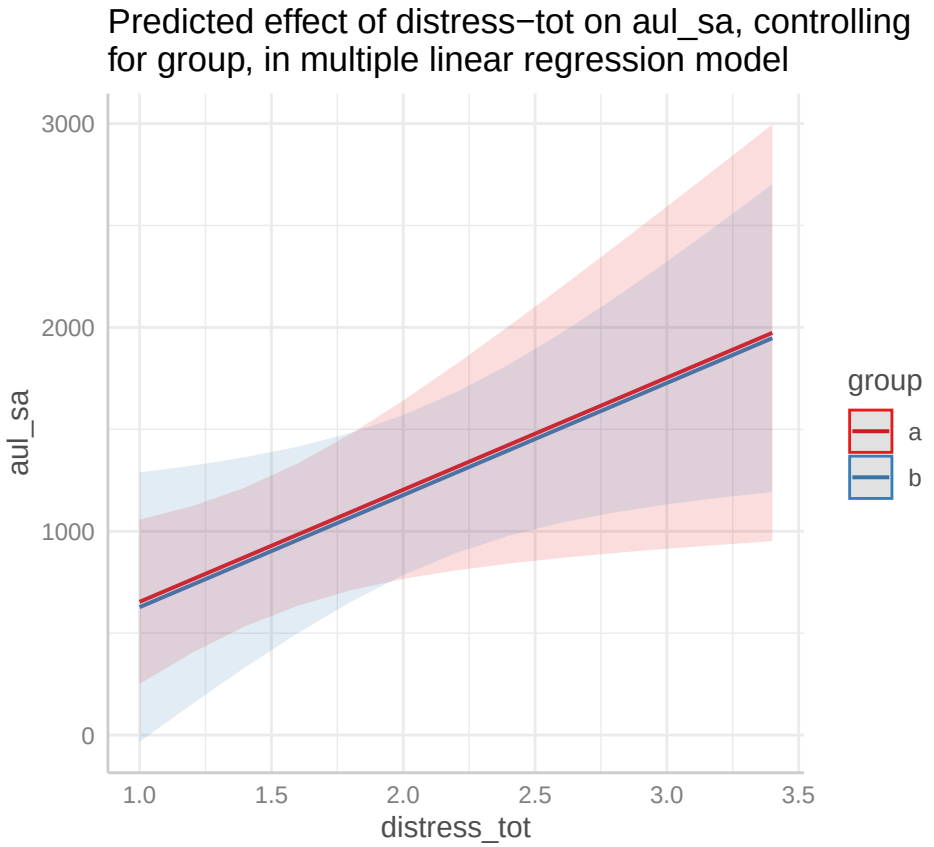
```
##
## $group
```




```
ggpredict(sa_model_multiple, terms = c("distress_tot", "group"))
```

```
## # Predicted values of aul_sa
##
## # group = a
##
## distress_tot | Predicted |            95% CI
## ------------------------------------------
##         1.00 |    653.79 | [251.29, 1056.30]
##         1.40 |    873.71 | [533.29, 1214.14]
##         1.80 |   1093.64 | [710.12, 1477.15]
##         2.20 |   1313.56 | [808.01, 1819.11]
##         2.60 |   1533.48 | [869.11, 2197.85]
##         3.40 |   1973.32 | [951.10, 2995.54]
##
## # group = b
##
## distress_tot | Predicted |             95% CI
## -------------------------------------------
##         1.00 |    627.63 | [ -33.38, 1288.64]
##         1.40 |    847.55 | [ 331.99, 1363.11]
##         1.80 |   1067.47 | [ 652.77, 1482.17]
##         2.20 |   1287.40 | [ 893.11, 1681.68]
##         2.60 |   1507.32 | [1042.29, 1972.35]
##         3.40 |   1947.16 | [1192.33, 2701.99]
```

```
df3 <- ggpredict(sa_model_multiple, terms = c("distress_tot",
  "group"))
plot(df3) + labs(title = "Predicted effect of distress-tot on aul_sa, controlling\nfor group, in multipl
```



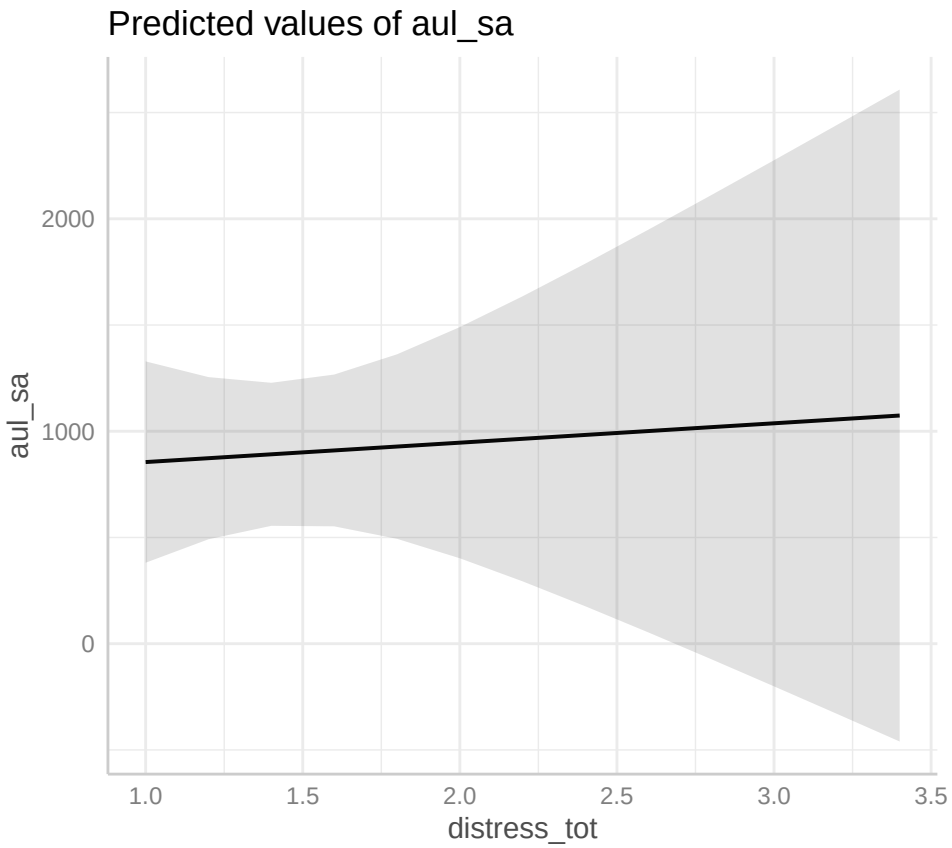

```
## Test for interaction effect and examine marginal effects
sa_model_interact <- lm(aul_sa ~ distress_tot * group, data = main_outcome_data)
summary(sa_model_interact)
```

```
##
## Call:
## lm(formula = aul_sa ~ distress_tot * group, data = main_outcome_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1562.8  -646.3  -202.2   553.3  2058.0
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)           763.67     585.85   1.304    0.199
## distress_tot           91.25     389.48   0.234    0.816
## groupb              -1332.60     909.95  -1.464    0.150
## distress_tot:groupb   768.68     504.27   1.524    0.134
##
## Residual standard error: 921.3 on 47 degrees of freedom
## Multiple R-squared:  0.1586, Adjusted R-squared:  0.1049
```

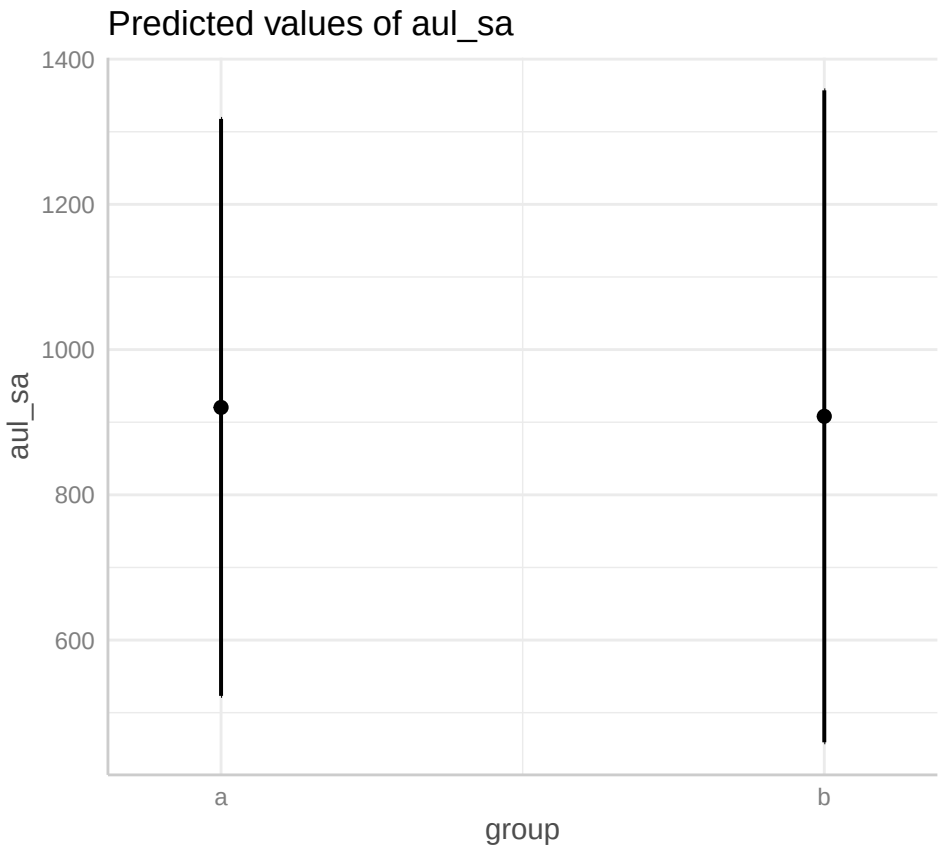
```
## F-statistic: 2.954 on 3 and 47 DF,  p-value: 0.04199
```

```
df4 <- ggpredict(sa_model_interact)
plot(df4)
```

```
## $distress_tot
```



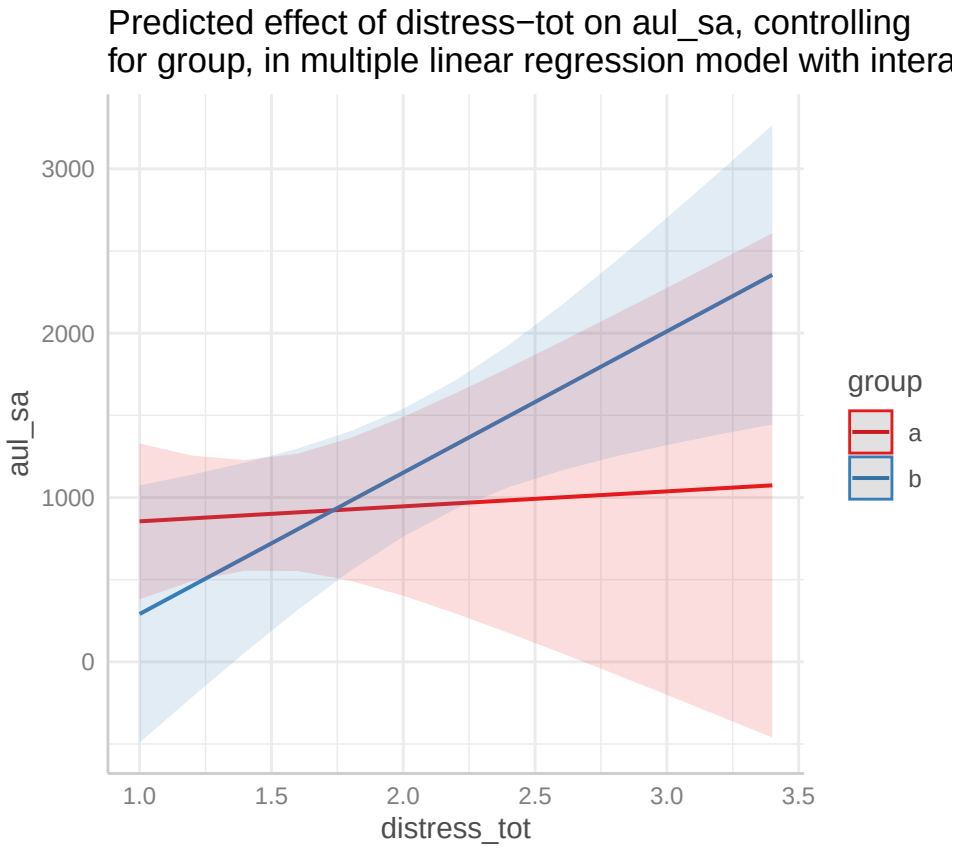

```
##
## $group
```




```
ggpredict(sa_model_interact, terms = c("distress_tot", "group"))
```

```
## # Predicted values of aul_sa
##
## # group = a
##
## distress_tot | Predicted |             95% CI
## ------------------------------------------------
##         1.00 |    854.92 | [ 381.06, 1328.78]
##         1.40 |    891.42 | [ 554.82, 1228.02]
##         1.80 |    927.92 | [ 493.72, 1362.13]
##         2.20 |    964.43 | [ 293.43, 1635.43]
##         2.60 |   1000.93 | [  53.07, 1948.78]
##         3.40 |   1073.93 | [-460.42, 2608.28]
##
## # group = b
##
## distress_tot | Predicted |             95% CI
## ------------------------------------------------
##         1.00 |    291.00 | [-491.66, 1073.66]
##         1.40 |    634.97 | [  57.59, 1212.36]
##         1.80 |    978.95 | [ 554.31, 1403.59]
##         2.20 |   1322.92 | [ 931.29, 1714.55]
##         2.60 |   1666.89 | [1164.35, 2169.43]
##         3.40 |   2354.84 | [1444.21, 3265.46]
```

```
df5 <- ggpredict(sa_model_interact, terms = c("distress_tot",
  "group"))
plot(df5) + labs(title = "Predicted effect of distress-tot on aul_sa, controlling\nfor group, in multipl
```




```
# Compare predictive value of models
anova(sa_model_main, sa_model_multiple)  # carry forward simpler model
```

```
## Analysis of Variance Table
##
## Model 1: aul_sa ~ distress_tot
## Model 2: aul_sa ~ distress_tot + group
##   Res.Df      RSS Df Sum of Sq      F Pr(>F)
## 1     49 41870267
## 2     48 41864051  1    6215.9 0.0071 0.9331
```

```
anova(sa_model_main, sa_model_interact)
```

```
## Analysis of Variance Table
##
## Model 1: aul_sa ~ distress_tot
## Model 2: aul_sa ~ distress_tot * group
##   Res.Df      RSS Df Sum of Sq      F Pr(>F)
## 1     49 41870267
## 2     47 39891886  2   1978381 1.1654 0.3206
```

```r
# No statistically significant difference in predictive
# value of models.  Carry forward the simplest model
# (sa_model_main) and check for improved fit when including
# covariate of days between assessments.  First calculate
# days between assessments

## plot surface area and delay between PSD assessments and
## HFS induction prepare hfs and distress dates

hfs_dates <- data.frame(master_data[[1]]) %>%
  rename(hfs_date = date) %>%
  rename(pid = blinded_id) %>%
  dplyr::select(pid, hfs_date) %>%
  unique()

distress_dates <- distress_data %>%
  dplyr::select(pid, date_and_time_hop) %>%
  rename(distress_date = date_and_time_hop) %>%
  mutate(pid = as.numeric(pid))

distress_hfs_dates <- distress_dates %>%
  right_join(hfs_dates)
```

```
## Joining, by = "pid"
```

```r
days <- main_outcome_data %>%
  mutate(pid = as.numeric(pid)) %>%
  left_join(distress_hfs_dates) %>%
  mutate(day_diff = (hfs_date - distress_date)) %>%
  mutate(day_diff = as.numeric(day_diff))
```

```
## Joining, by = "pid"
```

```r
# calculate the delay (in days) between PSD assessment and
# hfs induction

sa_model_main_cov <- lm(aul_sa ~ distress_tot + day_diff, data = days)
summary(sa_model_main_cov)  # No effect
```

```
##
## Call:
## lm(formula = aul_sa ~ distress_tot + day_diff, data = days)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1378.6  -711.2  -200.1   565.8  2034.5
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -20.22     404.25  -0.050   0.9603
## distress_tot   527.15     211.45   2.493   0.0162 *
## day_diff        35.20      32.82   1.072   0.2889
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 923 on 48 degrees of freedom
## Multiple R-squared:  0.1376, Adjusted R-squared:  0.1017
## F-statistic: 3.829 on 2 and 48 DF,  p-value: 0.02865
```

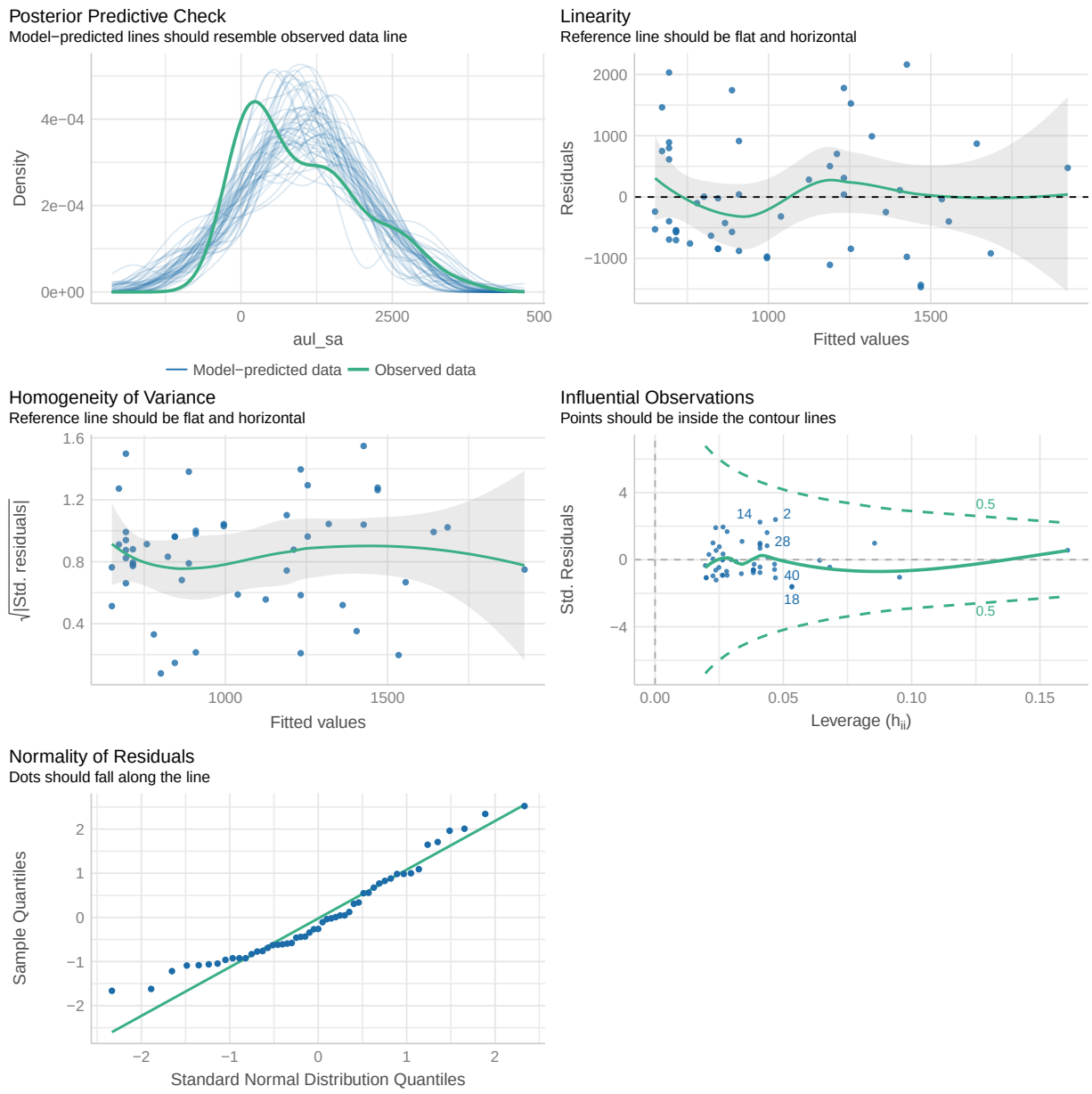
```
## Compare model fit.
anova(sa_model_main, sa_model_main_cov)
```

```
## Analysis of Variance Table
##
## Model 1: aul_sa ~ distress_tot
## Model 2: aul_sa ~ distress_tot + day_diff
##   Res.Df      RSS Df Sum of Sq      F Pr(>F)
## 1     49 41870267
## 2     48 40890448  1    979819 1.1502 0.2889
```

```
# If p-value is not significant, therefore sa_model_main is
# the final model
```

Surface area model assumptions

```
## plot model assumptions; main effect only
check_model(sa_model_main)
```




```
check_heteroscedasticity(sa_model_main)
```

```
## OK: Error variance appears to be homoscedastic (p = 0.457).
```

```
check_outliers(sa_model_main)
```

```
## OK: No outliers detected.
## - Based on the following method and threshold: cook (0.7).
## - For variable: (Whole model)
```

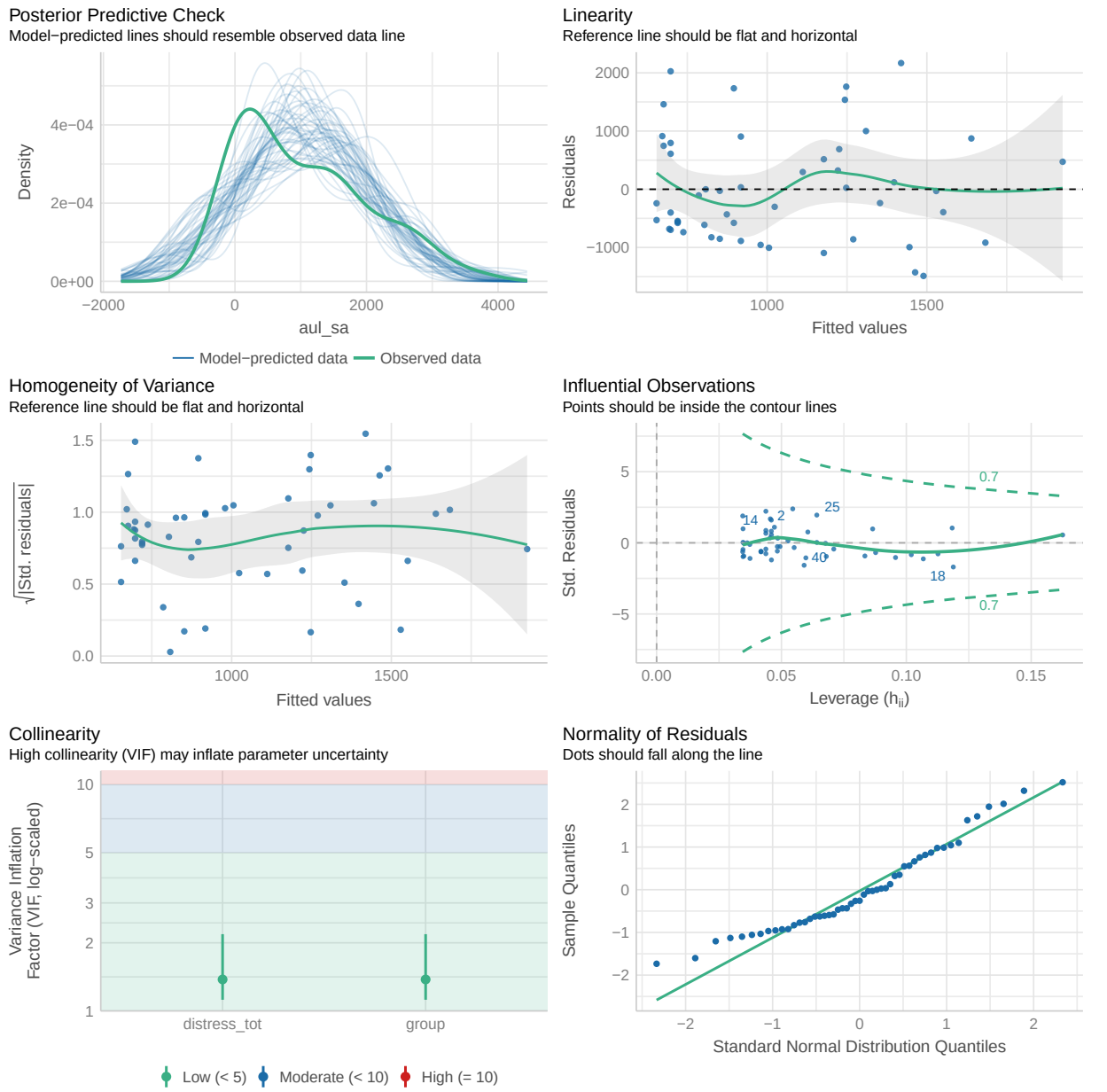
```
check_normality(sa_model_main)
```

```
## Warning: Non-normality of residuals detected (p = 0.017).
```

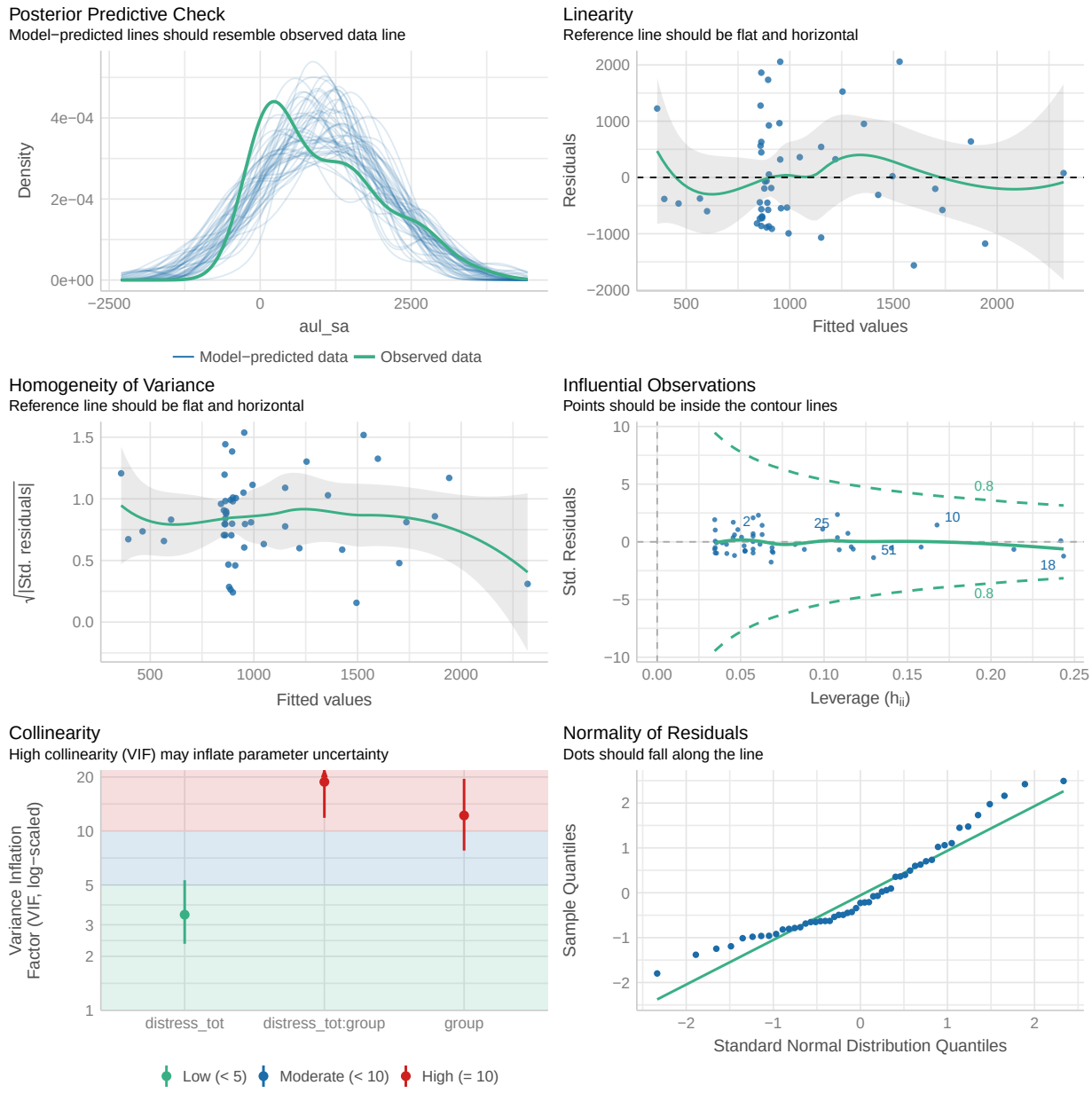
```
check_homogeneity(sa_model_main, method = "fligner")
```

```
## OK: There is not clear evidence for different variances across groups (Fligner-Killeen Test, p = 0.08
```

```
# not normally distributed groups; used fligner method

# plot to assess assumptions; model with two main effects
check_model(sa_model_multiple)
```



```
# plot to assess assumptions; model with interaction term
check_model(sa_model_interact)
```




```
# all three models look acceptable; collinearity expected
# in interaction model
```

Surface area model summary:

Formal modelling of the relationship between surface area and PSD (sa_model_main) confirmed that the surface area was indeed predicted by PSD ($F(1,49) = 6.49$, $p=0.01$). sa_model_main predicted that a one-unit increase in PSD is associated with a 538.7 cm2 increase in the surface area of SH, where PSD could range between one and three. Our plots suggest that group b showed a stronger positive association between surface area and PSD than group a. Formal modelling of the relationship between surface area, PSD and group (sa_model_multiple) confirms that the surface area was still predicted by PSD ($p = 0.03$)

but not group (p = 0.93), (F(2,48) = 3.18, p=0.05). The inclusion of 'group' as an independent variable in sa_model_multiple did not improve the model fit (ANOVA comparing models sa_model_main and sa_model_multiple Pr(>F) = 0.93).

Given the suggestion of a between-group difference in the relationship between PSD and surface area suggested by the plots above, we formally tested for an interaction effect between PSD and group (sa_model_interact). sa_model_interact confirms that the surface area of SH was not predicted by PSD (p=0.82), group (p=0.15) or the interaction between PSD and group (p=0.13), (F(3,47) = 2.95, p=0.04). We had a statistically significant relationship between PSD and surface area in the sa_model_main and sa_model_multiple models but not in the model that includes the interaction term (sa_model_interact). The loss of statistical power could explain the apparent mismatch between the plot (which suggests an interaction effect, with a strong relationship in b than a) and the interaction model. Including the interaction effect between PSD and group did not improve the model fit (ANOVA comparing sa_model_multiple and sa_model_interact p=0.13). Therefore, sa_model_main was deemed the final model.

Finally, we wanted to test if including the delay in days between PSD assessment and SH induction as a covariate improved the fit of sa_model_main in the current analysis (sa_model_main_cov). sa_model_main_cov confirms that the surface area was still predicted by PSD (p = 0.01) but not by the delay between PSD and SH induction (p = 0.29), (F(2,48) = 3.83, p=0.03). The inclusion of the 'delay between PSD and SH induction' as an independent variable in sa_model_main_cov did not improve the model fit (ANOVA comparing sa_model_main and sa_model_main_cov Pr(>F) = 0.29). Therefore, sa_model_main is still deemed the final model. The estimated marginal means for the predicted effect of PSD on the surface area of SH when using sa_model_main are presented above.

sa_model_main model assumptions:

- Posterior predictive check assumption: The model-fitted data closely resembled the observed data; however, the observed data formed a small blob to the left. This similarity is acceptable for the current sample size, and the assumption of the posterior predictive check was deemed to have been met.
- Linearity assumption: The scatter plot shows that the fitted values did not consistently lie flat on the reference line (when residuals equal 0). Still, the degree of deviation from the line was deemed acceptable for this sample, and the assumption of linearity was deemed to have been met.

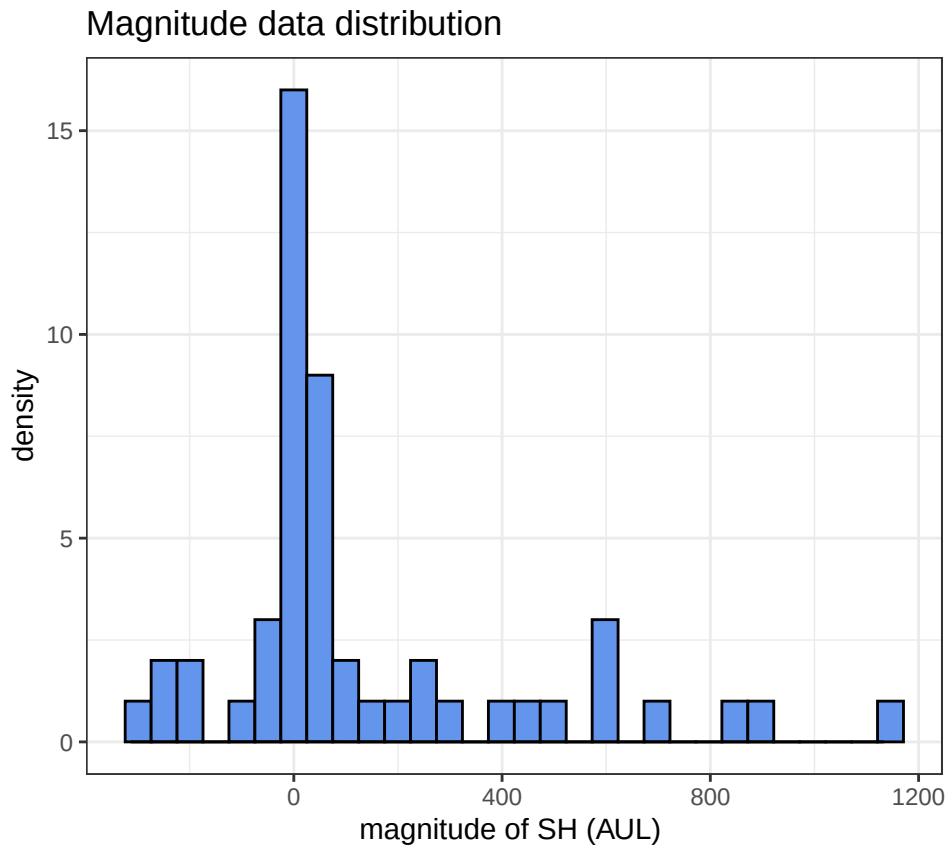
- Homogeneity of variance assumption: The scatter plot shows that the fitted values did not consistently lie flat on the reference line. Again, the degree of deviation from the line was acceptable for this sample, and the dots on the scatterplot did not form a discernible pattern. This suggests that the assumption of homogeneity of variance may have been met. Formal testing confirmed that this assumption was met (p = 0.46).
- Influential observations assumption: We had no outliers or observations with high leverage. Observations fell within the contour lines. Formal testing confirmed that this assumption was met (Cook's distance; 0.7).
- The normality of residuals assumption: The plot shows that the dots did not consistently resemble the linear diagonal reference line. The plot also suggests that the data points had a somewhat equal distribution of positive and negative residuals, suggesting that the normality of residuals assumption may still need to be met. Formal testing confirmed the non-normality of residuals; therefore, this assumption was not met (p = 0.02).

Magnitude

```
# check the distribution of the magnitude data using a
# histogram
ggplot(main_outcome_data) + aes(x = aul_mag) + geom_histogram(fill = "#6495ed",
```

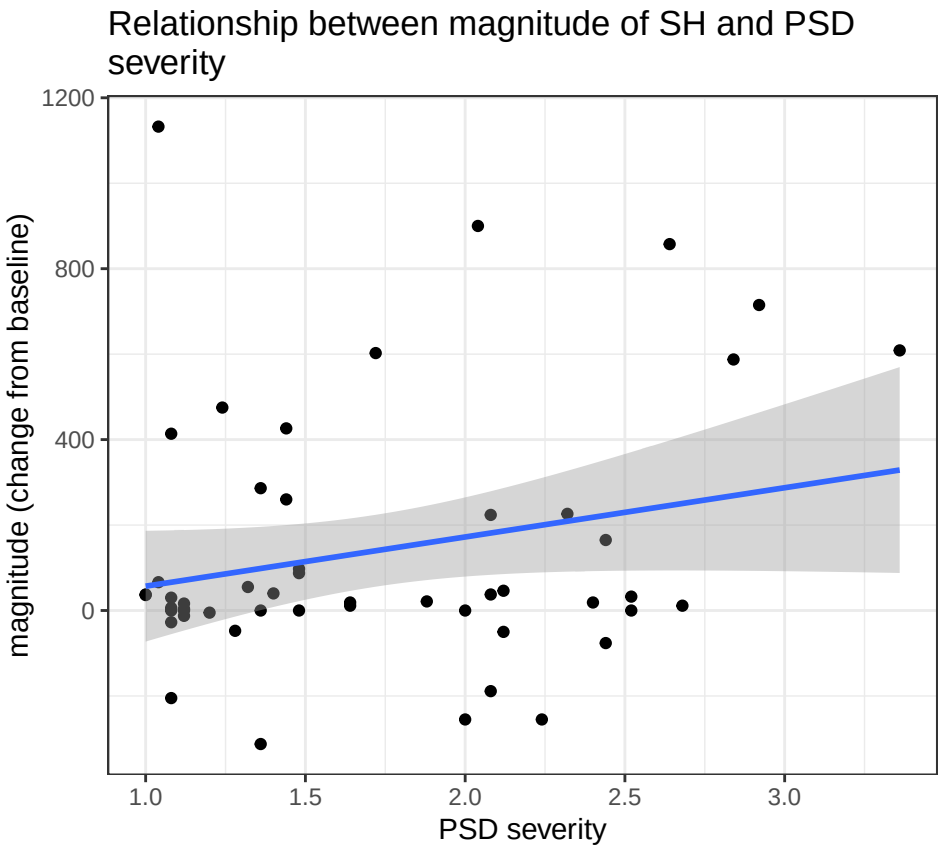
```
  col = "black") + geom_density() + labs(title = "Magnitude data distribution",
  x = "magnitude of SH (AUL) ") + theme_bw()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



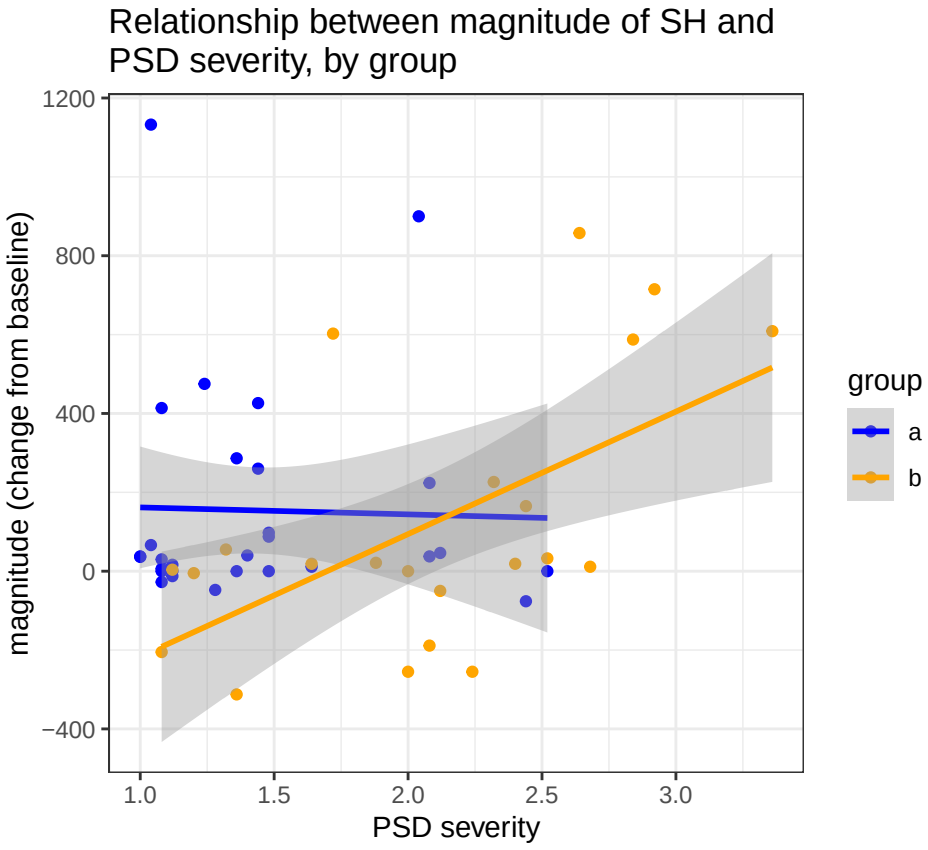

```
# plot relationship between magnitude and PSD
mag_psd <- ggplot(main_outcome_data) + aes(x = distress_tot,
  y = aul_mag) + geom_point() + geom_smooth(method = "lm") +
  labs(title = "Relationship between magnitude of SH and PSD\nseverity",
    x = "PSD severity", y = "magnitude (change from baseline)") +
  theme_bw()
print(mag_psd)
```

```
## `geom_smooth()` using formula 'y ~ x'
```

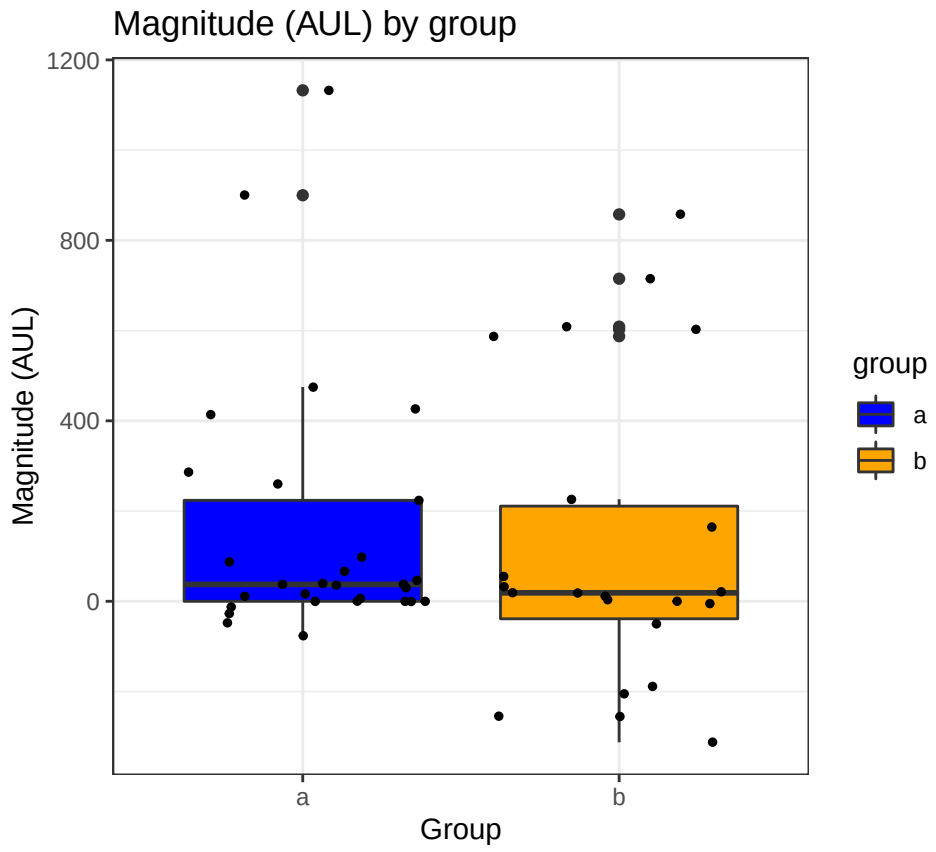



```
# plot relationship between magnitude and PSD, by group
mag_psd_group <- ggplot(main_outcome_data) + aes(x = distress_tot,
  y = aul_mag, group = group, colour = group) + geom_point() +
  geom_smooth(method = "lm") + scale_colour_manual(values = c("blue",
  "orange")) + labs(title = "Relationship between magnitude of SH and \nPSD severity, by group",
  x = "PSD severity", y = "magnitude (change from baseline)") +
  theme_bw()

print(mag_psd_group)
```

```
## `geom_smooth()` using formula 'y ~ x'
```



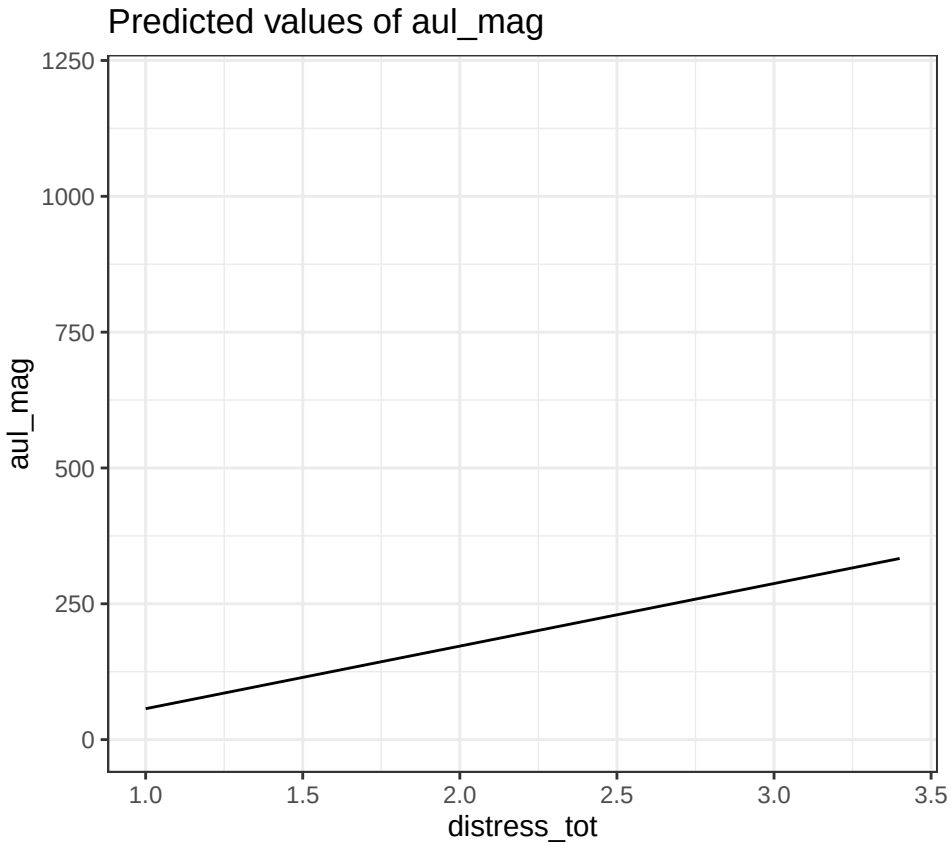

```
# plot magnitude aul between groups
mag_aul <- ggplot(data = main_outcome_data) + aes(x = group,
  y = aul_mag, fill = group) + geom_boxplot() + geom_jitter(shape = 21,
  color = "black", fill = "black", size = 1) + scale_fill_manual(values = c("blue",
  "orange")) + labs(x = "Group", y = "Magnitude (AUL)", title = "Magnitude (AUL) by group") +
  theme_bw()

print(mag_aul)
```



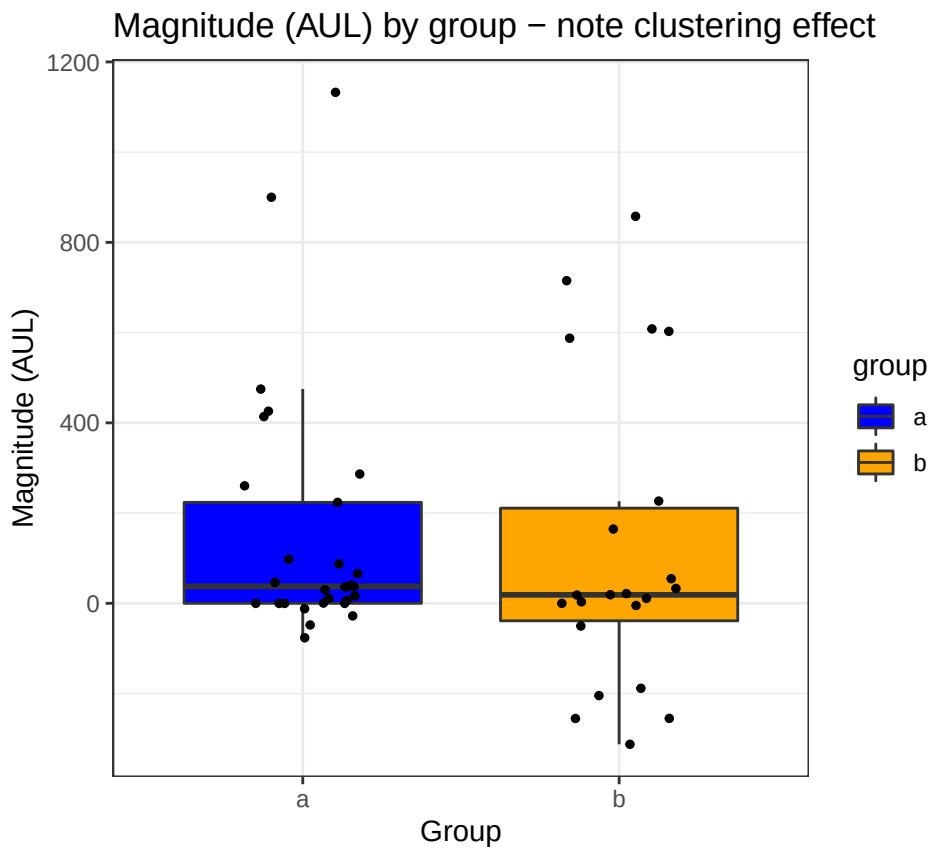

```
## prep data
mag_model_data <- main_outcome_data %>%
  filter(!is.na(aul_mag))

## enter data into model: main effects only, first
mag_model_main <- lm(aul_mag ~ distress_tot, data = mag_model_data)
summary(mag_model_main)
```

```
##
## Call:
## lm(formula = aul_mag ~ distress_tot, data = mag_model_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -454.76 -141.75  -66.13   29.85 1070.98
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    -58.28     124.74  -0.467   0.6424
## distress_tot   115.19      68.41   1.684   0.0986 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 299 on 49 degrees of freedom
## Multiple R-squared:  0.0547, Adjusted R-squared:  0.03541
## F-statistic: 2.836 on 1 and 49 DF,  p-value: 0.09856
```

```r
df1_mag <- ggpredict(mag_model_main, terms = "distress_tot")
# p-value > 0.05 but the CIs suggest there is a potential
# effect which would be interesting if genuine
ggplot(df1_mag, aes(x, predicted)) + geom_line() + labs(x = get_x_title(df1_mag),
  y = get_y_title(df1_mag), title = get_title(df1_mag)) + theme_bw() +
  coord_cartesian(ylim = c(0, 1200))
```

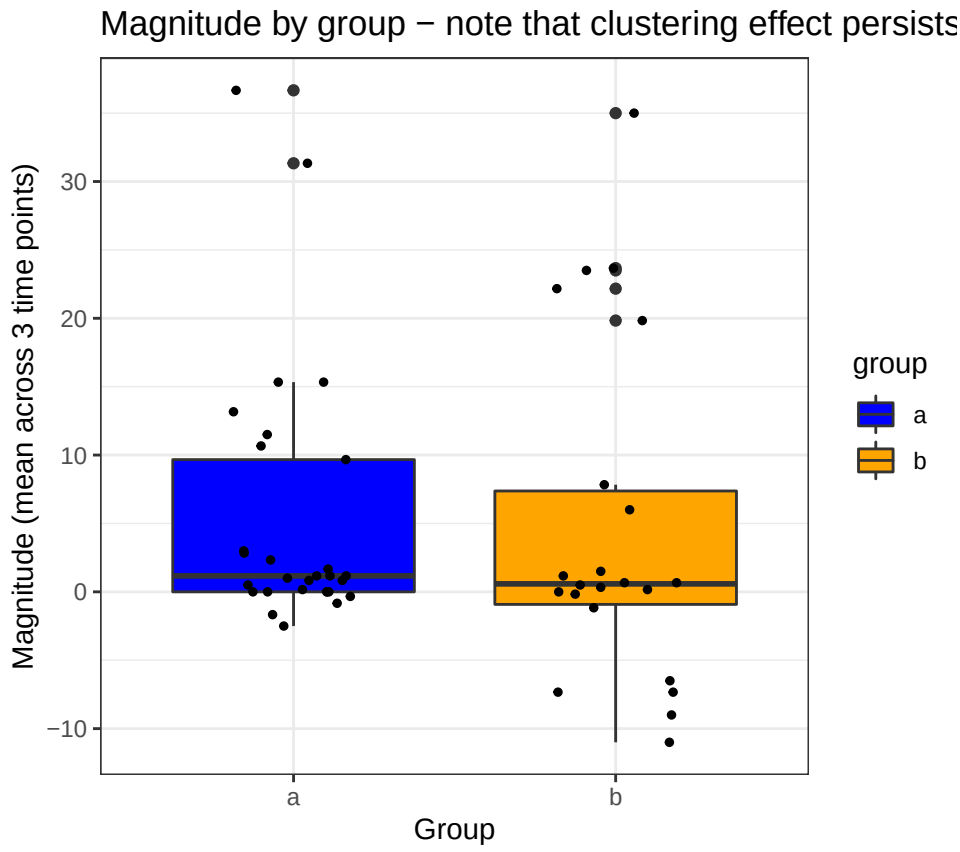



```r
# plot magnitude aul between groups
mag_aul2 <- ggplot(data = main_outcome_data) + aes(x = group,
  y = aul_mag, fill = group) + geom_boxplot(outlier.shape = NA) +
  geom_jitter(width = 0.2, shape = 21, color = "black", fill = "black",
    size = 1) + scale_fill_manual(values = c("blue", "orange")) +
  labs(x = "Group", y = "Magnitude (AUL)", title = "Magnitude (AUL) by group - note clustering effect") +
  theme_bw()

print(mag_aul2)
```



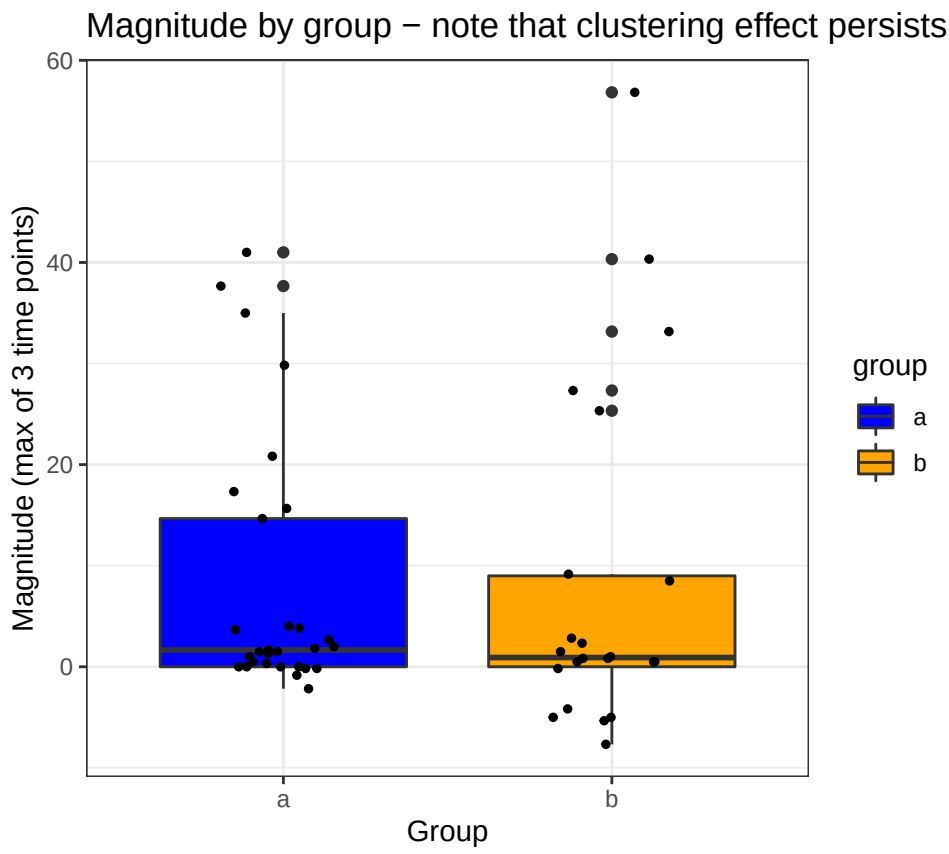

```
# try plotting mean per participant
mag_change_bl %>%
  group_by(pid) %>%
  summarise(mean_mag = mean(mag)) %>%
  left_join(main_outcome_data[, c(1:2)]) %>%
  filter(!is.na(group)) %>%
  ggplot(data = .) + aes(x = group, y = mean_mag, fill = group) +
  geom_boxplot() + geom_jitter(width = 0.2, height = 0, shape = 21,
  color = "black", fill = "black", size = 1) + scale_fill_manual(values = c("blue",
  "orange")) + labs(x = "Group", y = "Magnitude (mean across 3 time points)",
  title = "Magnitude by group - note that clustering effect persists") +
  theme_bw()
```

```
## Joining, by = "pid"
```




```
# try plotting max per participant
mag_change_bl %>%
  group_by(pid) %>%
  summarise(max_mag = max(mag)) %>%
  left_join(main_outcome_data[, c(1:2)]) %>%
  filter(!is.na(group)) %>%
  ggplot(data = .) + aes(x = group, y = max_mag, fill = group) +
  geom_boxplot() + geom_jitter(width = 0.2, height = 0, shape = 21,
  color = "black", fill = "black", size = 1) + scale_fill_manual(values = c("blue",
  "orange")) + labs(x = "Group", y = "Magnitude (max of 3 time points)",
  title = "Magnitude by group - note that clustering effect persists") +
  theme_bw()
```

```
## Joining, by = "pid"
```

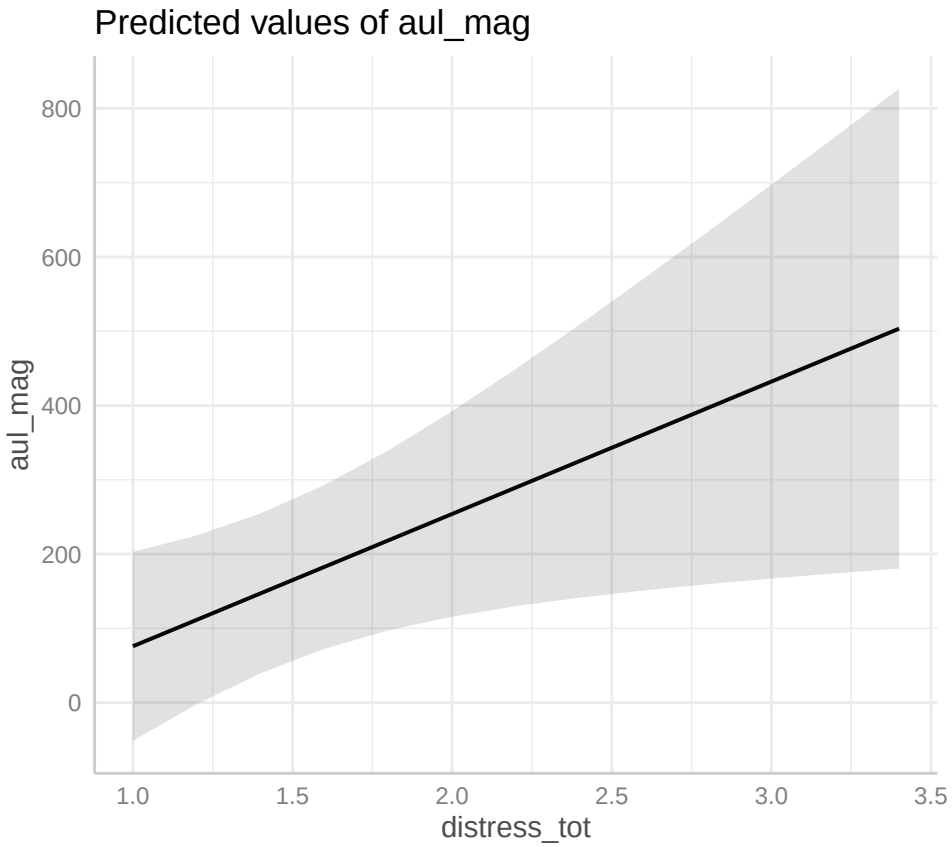



```
# Clustering effect seems truly representative of the data

## Test for group effect and examine marginal effects
mag_model_multiple <- lm(aul_mag ~ distress_tot + group, data = mag_model_data)
summary(mag_model_multiple)
```

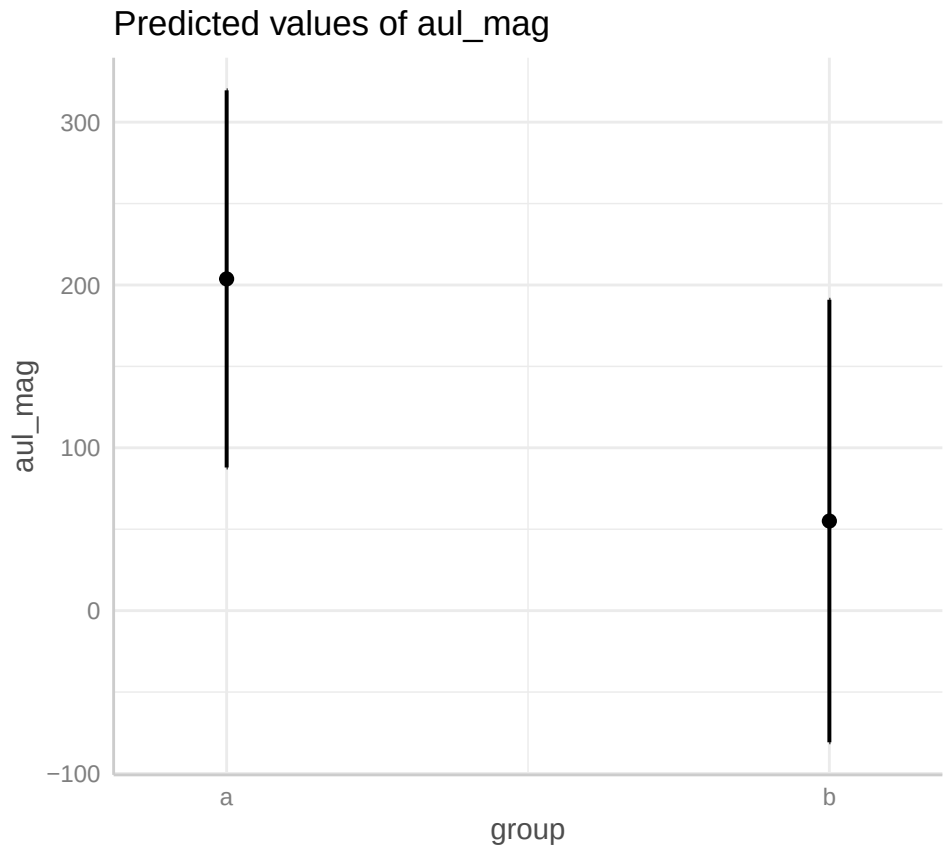
```
##
## Call:
## lm(formula = aul_mag ~ distress_tot + group, data = mag_model_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -408.69 -163.44  -73.88   67.36 1049.52
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -102.33     126.48  -0.809   0.4225
## distress_tot   178.18      79.24   2.249   0.0292 *
## groupb        -148.69      97.91  -1.519   0.1354
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 295.1 on 48 degrees of freedom
## Multiple R-squared:  0.09804,    Adjusted R-squared:  0.06046
## F-statistic: 2.609 on 2 and 48 DF,  p-value: 0.08404
```

```
# Interesting that distress effect now statistically
# significant.  BUT model assumptions may still be
# violated, given clustering

df3_mag <- ggpredict(mag_model_multiple)
plot(df3_mag)
```

```
## $distress_tot
```



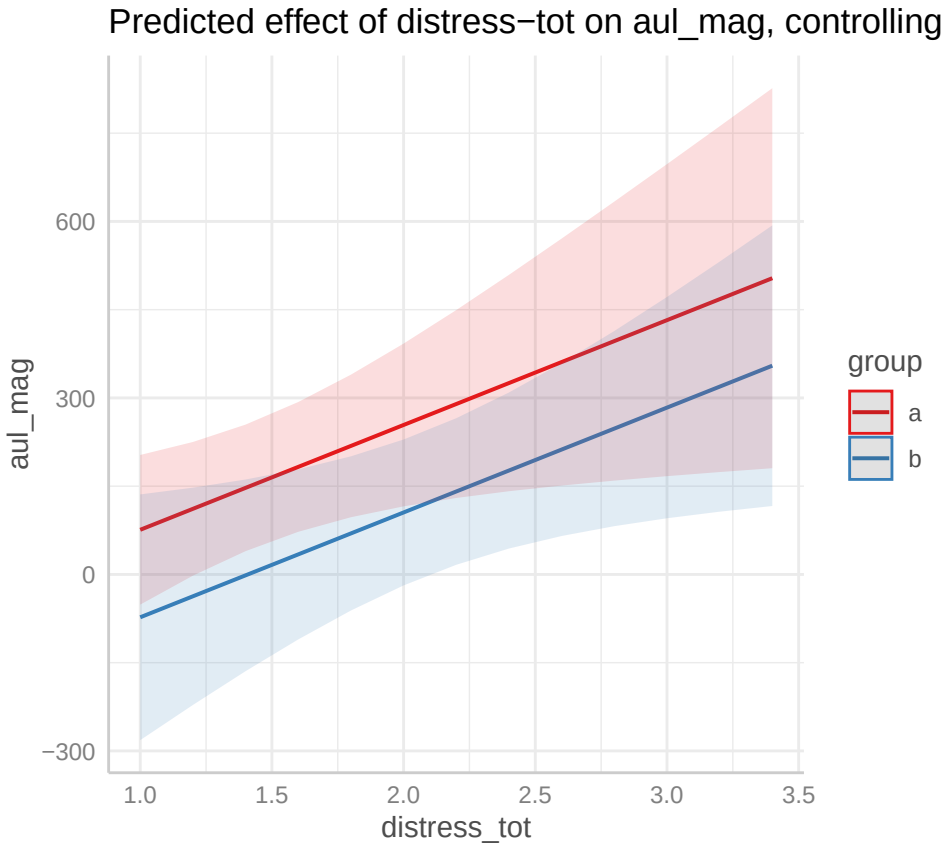

```
##
## $group
```




```
ggpredict(mag_model_multiple, terms = c("distress_tot", "group"))
```

```
## # Predicted values of aul_mag
##
## # group = a
##
## distress_tot | Predicted |           95% CI
## ------------------------------------------
##         1.00 |     75.85 | [-51.32, 203.03]
##         1.40 |    147.13 | [ 39.56, 254.69]
##         1.80 |    218.40 | [ 97.22, 339.58]
##         2.20 |    289.67 | [129.94, 449.41]
##         2.60 |    360.95 | [151.03, 570.86]
##         3.40 |    503.49 | [180.51, 826.47]
##
## # group = b
##
## distress_tot | Predicted |            95% CI
## -------------------------------------------
##         1.00 |    -72.84 | [-281.69, 136.01]
##         1.40 |     -1.57 | [-164.46, 161.33]
##         1.80 |     69.71 | [ -61.32, 200.74]
##         2.20 |    140.98 | [  16.40, 265.56]
##         2.60 |    212.25 | [  65.32, 359.18]
##         3.40 |    354.80 | [ 116.30, 593.29]
```

```
df2_mag <- ggpredict(mag_model_multiple, terms = c("distress_tot",
  "group"))
plot(df2_mag) + labs(title = "Predicted effect of distress-tot on aul_mag, controlling for group, in mul
```



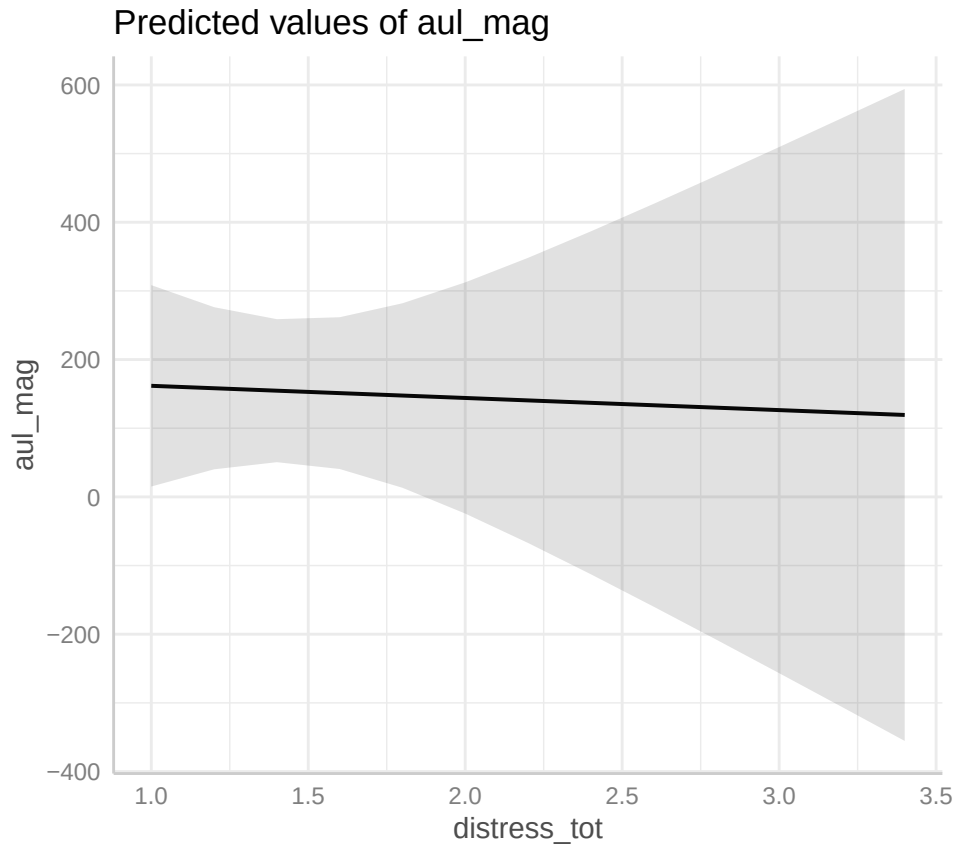

```
## Test for interaction effect and examine marginal effects
mag_model_interact <- lm(aul_mag ~ distress_tot * group, data = mag_model_data)
summary(mag_model_interact)
```

```
##
## Call:
## lm(formula = aul_mag ~ distress_tot * group, data = mag_model_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -423.58 -160.00  -95.71  118.43  971.44
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)            179.45     181.28   0.990   0.3273
## distress_tot           -17.68     120.52  -0.147   0.8840
## groupb                -706.72     281.57  -2.510   0.0156 *
## distress_tot:groupb    328.33     156.04   2.104   0.0407 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

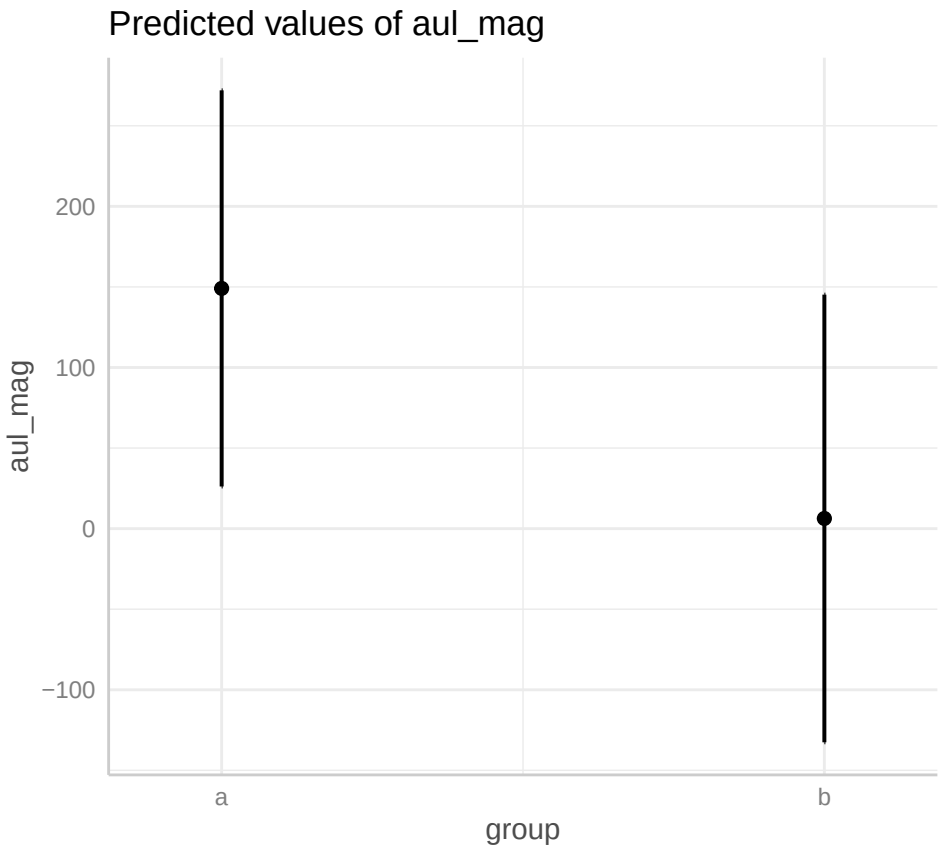
```
## Residual standard error: 285.1 on 47 degrees of freedom
## Multiple R-squared:  0.1757, Adjusted R-squared:  0.1231
## F-statistic: 3.339 on 3 and 47 DF,  p-value: 0.02705
```

```
# Interesting that distress effect now statistically
# significant.  BUT model assumptions may still be
# violated, given clustering

df4_mag <- ggpredict(mag_model_interact)
plot(df4_mag)
```

```
## $distress_tot
```




```
##
## $group
```




```
df4_mag_table <- ggpredict(mag_model_interact, terms = c("distress_tot",
  "group"))
print(df4_mag_table)
```

```
## # Predicted values of aul_mag
##
## # group = a
##
## distress_tot | Predicted |            95% CI
## ------------------------------------------
##         1.00 |    161.76 | [  15.14, 308.39]
##         1.40 |    154.69 | [  50.54, 258.85]
##         1.80 |    147.62 | [  13.26, 281.97]
##         2.20 |    140.55 | [ -67.08, 348.17]
##         2.60 |    133.47 | [-159.82, 426.77]
##         3.40 |    119.33 | [-355.45, 594.10]
##
## # group = b
##
## distress_tot | Predicted |            95% CI
## ------------------------------------------
##         1.00 |   -216.63 | [-458.81,  25.55]
##         1.40 |    -92.37 | [-271.03,  86.30]
##         1.80 |     31.89 | [ -99.50, 163.29]
##         2.20 |    156.15 | [  34.97, 277.33]
```

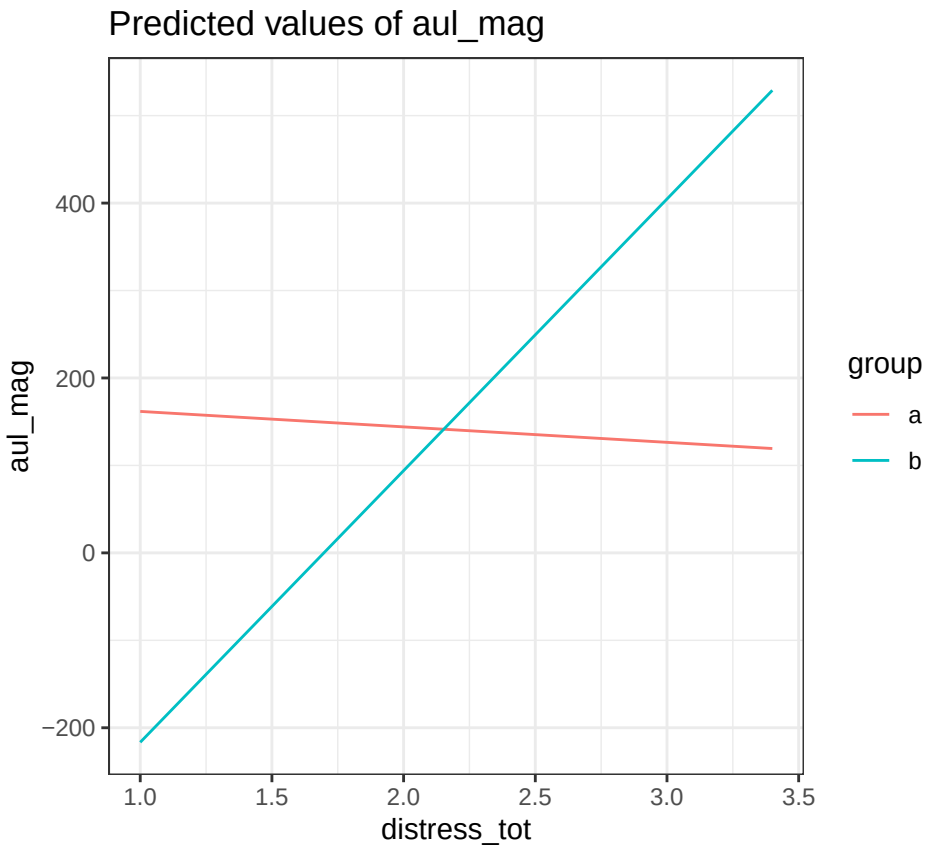
```
##         2.60 |    280.41 | [ 124.91, 435.91]
##         3.40 |    528.93 | [ 247.15, 810.71]
```

```
df4_mag_table %>%
  kbl(col.names = c("PSD", "Predicted magnitude", "standard error",
    "CI(lower bound)", "CI(upper bound)", names(df1)[-1:-5]),
    caption = "Estimated marginal means: predicted effect of PSD on the magnitude of SH in Model 3",
    booktabs = T) %>%
  kable_styling(latex_options = c("striped", "HOLD_position"))
```

Table 10: Estimated marginal means: predicted effect of PSD on the magnitude of SH in Model 3

PSD	Predicted magnitude	standard error	CI(lower bound)	CI(upper bound)	group
1.0	161.76481	74.81111	15.13773	308.39189	a
1.0	-216.62713	123.56334	-458.80683	25.55257	b
1.2	158.22818	60.24419	40.15175	276.30462	a
1.2	-154.49734	106.75081	-363.72509	54.73040	b
1.4	154.69156	53.14127	50.53658	258.84653	a
1.4	-92.36755	91.15610	-271.03023	86.29513	b
1.6	151.15493	56.39681	40.61922	261.69064	a
1.6	-30.23776	77.51773	-182.16972	121.69419	b
1.8	147.61830	68.55055	13.26169	281.97492	a
1.8	31.89203	67.04047	-99.50489	163.28894	b
2.0	144.08168	85.90535	-24.28972	312.45307	a
2.0	94.02182	61.36542	-26.25220	214.29584	b
2.2	140.54505	105.93514	-67.08400	348.17411	a
2.2	156.15161	61.82935	34.96830	277.33491	b
2.4	137.00843	127.38426	-112.66014	386.67699	a
2.4	218.28140	68.30729	84.40156	352.16123	b
2.6	133.47180	149.64364	-159.82434	426.76795	a
2.6	280.41119	79.33957	124.90849	435.91388	b
2.8	129.93518	172.39972	-207.96205	467.83241	a
2.8	342.54098	93.32476	159.62780	525.45415	b
3.0	126.39855	195.47910	-256.73343	509.53054	a
3.0	404.67077	109.13349	190.77305	618.56848	b
3.2	122.86193	218.77949	-305.93799	551.66185	a
3.2	466.80056	126.08168	219.68500	713.91611	b
3.4	119.32530	242.23713	-355.45074	594.10135	a
3.4	528.93035	143.76691	247.15237	810.70832	b

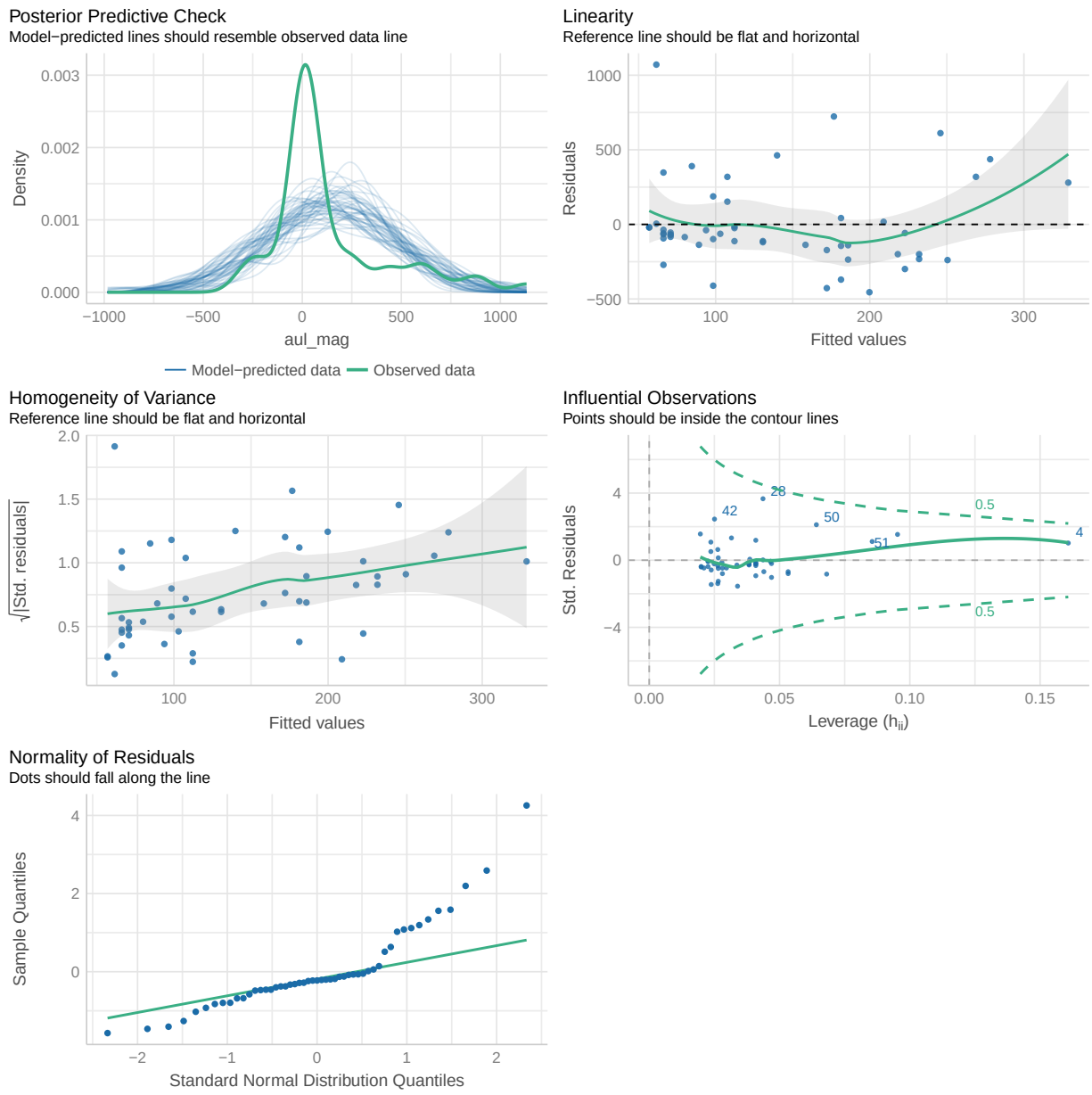
```
df5_mag <- ggpredict(mag_model_interact, terms = c("distress_tot",
  "group"))
ggplot(df5_mag, aes(x, predicted, colour = group)) + geom_line() +
  labs(x = get_x_title(df2_mag), y = get_y_title(df2_mag),
    title = get_title(df2_mag)) + theme_bw()
```



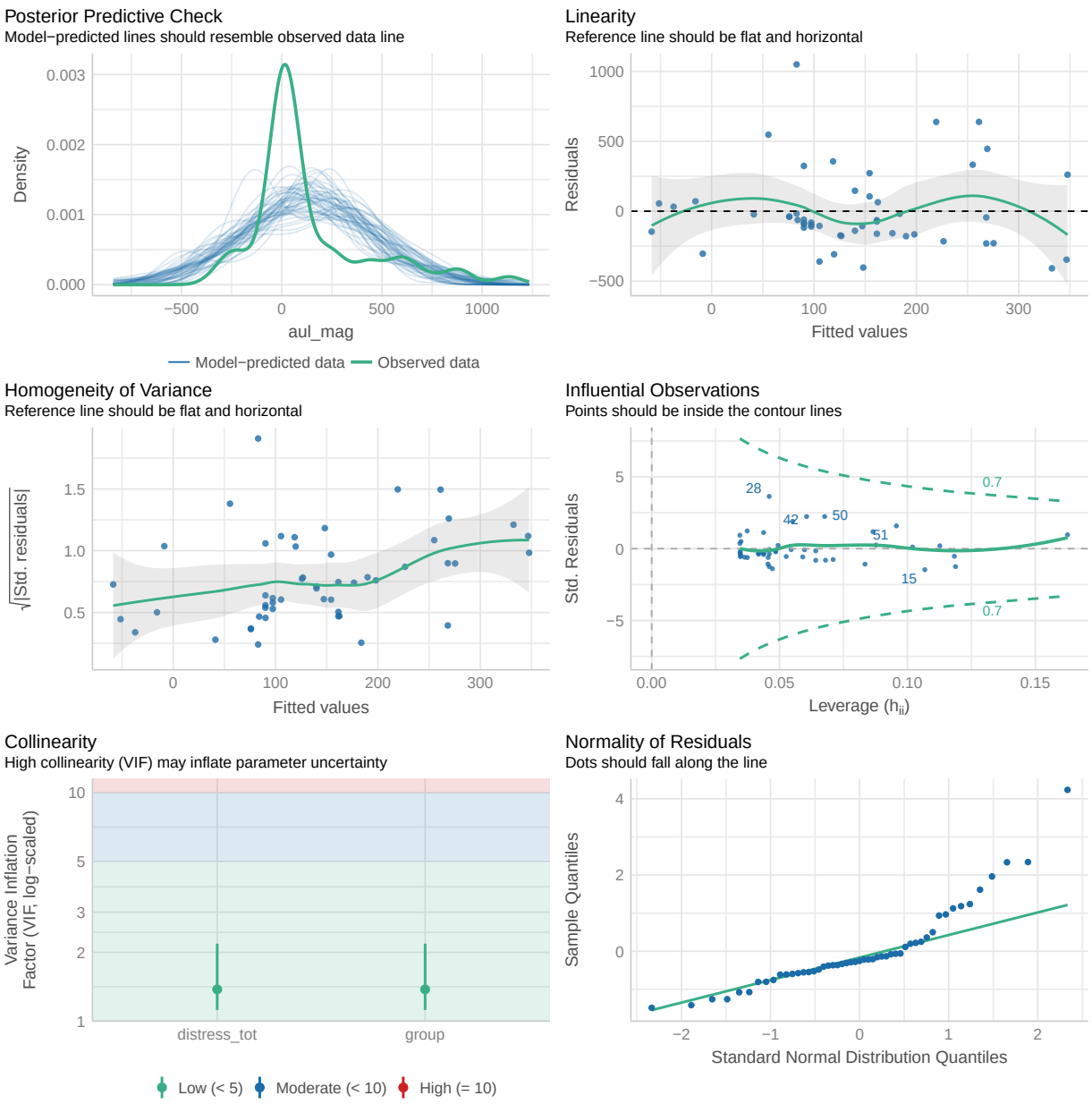

```
anova(mag_model_main, mag_model_multiple, mag_model_interact)
```

```
## Analysis of Variance Table
##
## Model 1: aul_mag ~ distress_tot
## Model 2: aul_mag ~ distress_tot + group
## Model 3: aul_mag ~ distress_tot * group
##   Res.Df     RSS Df Sum of Sq      F  Pr(>F)
## 1     49 4380198
## 2     48 4179393  1    200805 2.4709 0.12268
## 3     47 3819576  1    359817 4.4276 0.04074 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

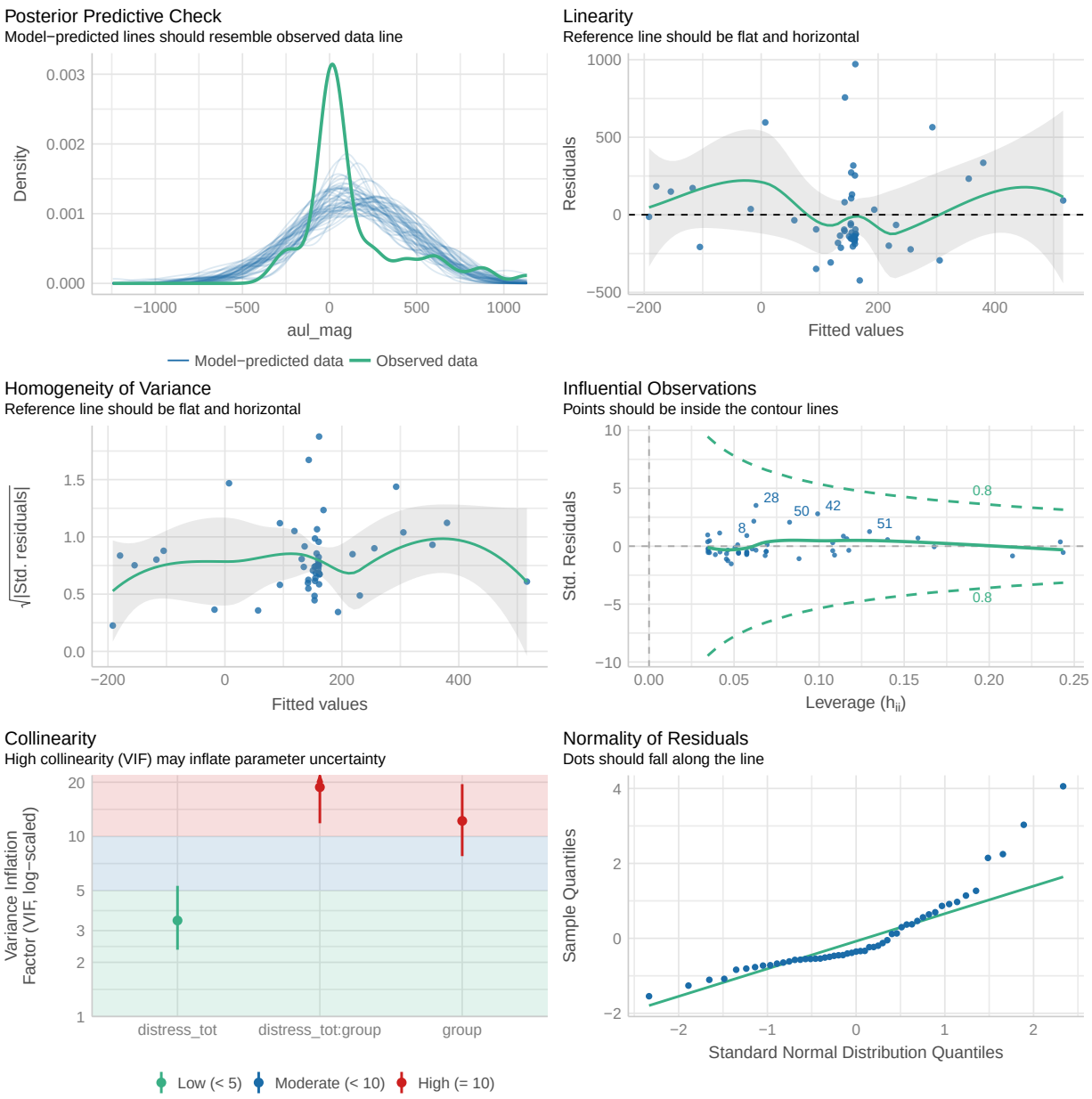
```
# assess model assumptions
check_model(mag_model_main)  # looks poor. see scatterplot - clustering?
```




```
check_model(mag_model_multiple)  # concerns with violation of assumptions
```




```
check_model(mag_model_interact)  # concerns with violation of assumptions
```



Magnitude model summary:

Formal modelling of the relationship between magnitude and PSD (mag_model_main) showed that the magnitude was not statistically predicted by PSD ($F(1,49) = 2.84$, $p=0.09$). However, important model assumptions were violated, rendering the model uninformative, possibly due to the clustering of magnitude values in group b. In the magnitude AUL, in group 'b', we noticed clustering. We investigated the possibility that the clustering of magnitude data was an artefact of our calculation process and found that the clustering was present in the raw data (see plots above). No current modelling technique can adequately and appropriately account for the clustering effect observed in group b.

Initially, we planned to test the relationship between the magnitude of SH and (i) PSD and group (mag_model_multiple) and (ii) the interaction effect between PSD and group (mag_model_interact) if mag_model_main confirmed that PSD predicted the magnitude of SH. However, we opted to test mag_model_multiple and mag_model_interact anyway because our plots suggested a possible difference in the relationships between magnitude and PSD between groups: group b showed a stronger positive

relationship between magnitude and PSD; however, this relationship was not observed in group a. Given the learning priority of the current research project, we opted to proceed with testing mag_model_multiple and mag_model_interact to investigate the pattern seen in our plots, although we realised the model assumptions would likely be problematic. mag_model_multiple confirms that the magnitude was predicted by PSD (p = 0.03) but not group (p = 0.14), ($F(2,48) = 2.61$, p=0.08). The inclusion of 'group' as an independent variable in mag_model_multiple did not improve the model fit (ANOVA comparing mag_model_main and mag_model_multiple $Pr(>F) = 0.12$).

mag_model_interact reflected what was seen in our plots: the magnitude was predicted by group (0.02) and the interaction between PSD and group (0.04) but not predicted by PSD (0.88), ($F(3,47) = 3.34$, p=0.03). Including the interaction effect between PSD and group showed a significant difference in fit (ANOVA comparing mag_model_multiple and mag_model_interact p=0.04). The estimated marginal means for mag_model_interact support noteworthy differences in the relationship between the magnitude of SH and PSD between the groups. However, this is likely strongly influenced by the clustering of the four values. Overall, it was impossible to generate an acceptable statistical model of the relationships between magnitude and (i) PSD and (ii) group for these data.

magnitude model assumptions:

- mag_model_main violated four (posterior predictive check, linearity, homogeneity of variance, and normality of residuals) of five assumptions, and the assumption of influential observations was deemed to have been met.
- mag_model_multiple violated three (posterior predictive check, homogeneity of variance, and normality of residuals) of six assumptions. The assumptions of linearity, collinearity and influential observations were deemed to have been met.
- mag_model_interact violated five (posterior predictive check, homogeneity of variance, linearity, collinearity and normality of residuals) of six assumptions, and the assumption of influential observations was deemed to have been met.

Unplanned analysis 1: Does the peak of the surface area correspond temporally with the peak of the magnitude of SH?

To answer this question, we plotted the surface area and magnitude of SH for each time point after the induction by each group.

```
sa <- raw_sa %>%
  rename(time = time_sa)
mag <- mag_change_bl %>%
  rename(time = time_mag)
both <- full_join(sa, mag) %>%
  pivot_longer(cols = c("area", "mag"), names_to = "outcome",
    values_to = "value")
```

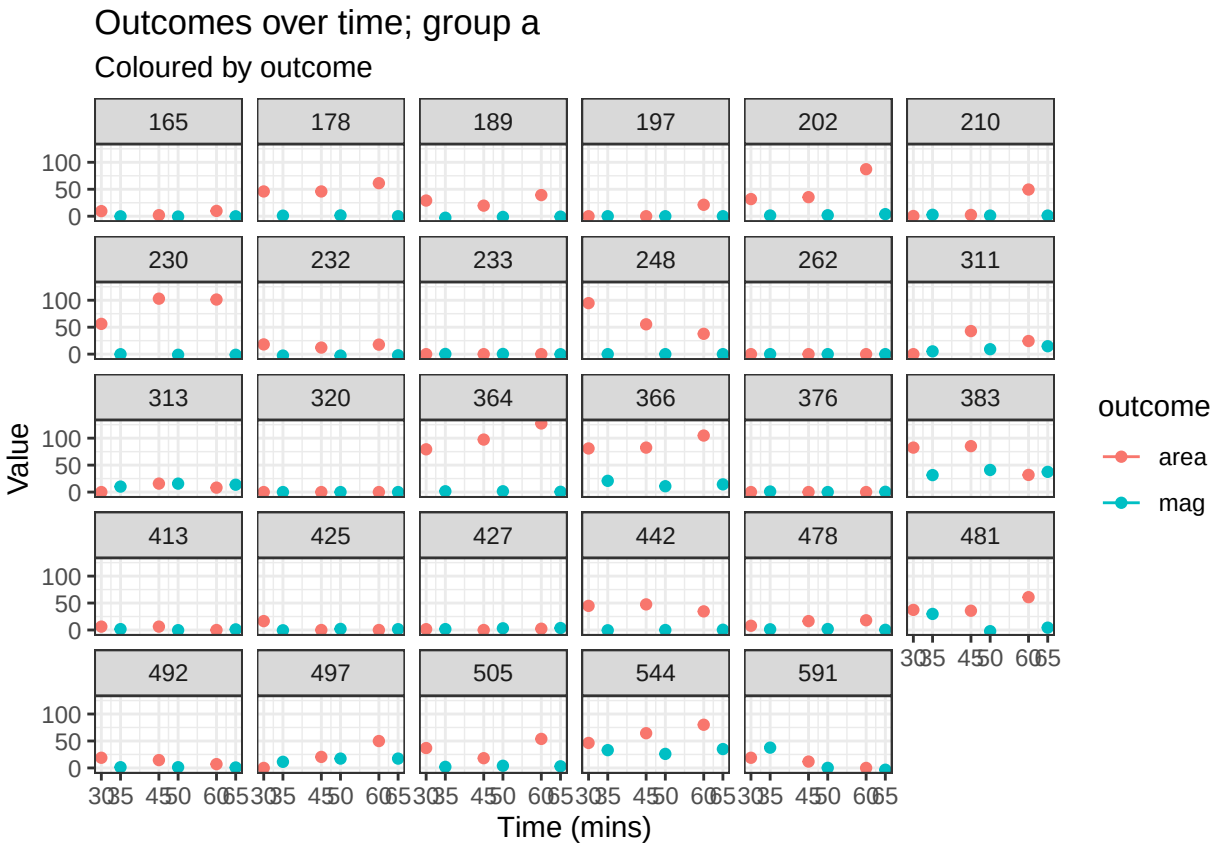
```
## Joining, by = c("pid", "group", "time", "participation_status")
```

```
## plot surface area by individual over time plot group a
both %>%
  filter(group == "a") %>%
  ggplot(.) + aes(x = time, y = value, group = outcome, colour = outcome,
  fill = outcome) + geom_point() + geom_line(aes(group = outcome)) +
  facet_wrap(~pid, nrow = 5) + labs(title = "Outcomes over time; group a",
  subtitle = "Coloured by outcome", x = "Time (mins)", y = "Value") +
```

```
  scale_x_continuous(breaks = c(30, 35, 45, 50, 60, 65)) +
  theme_bw()
```

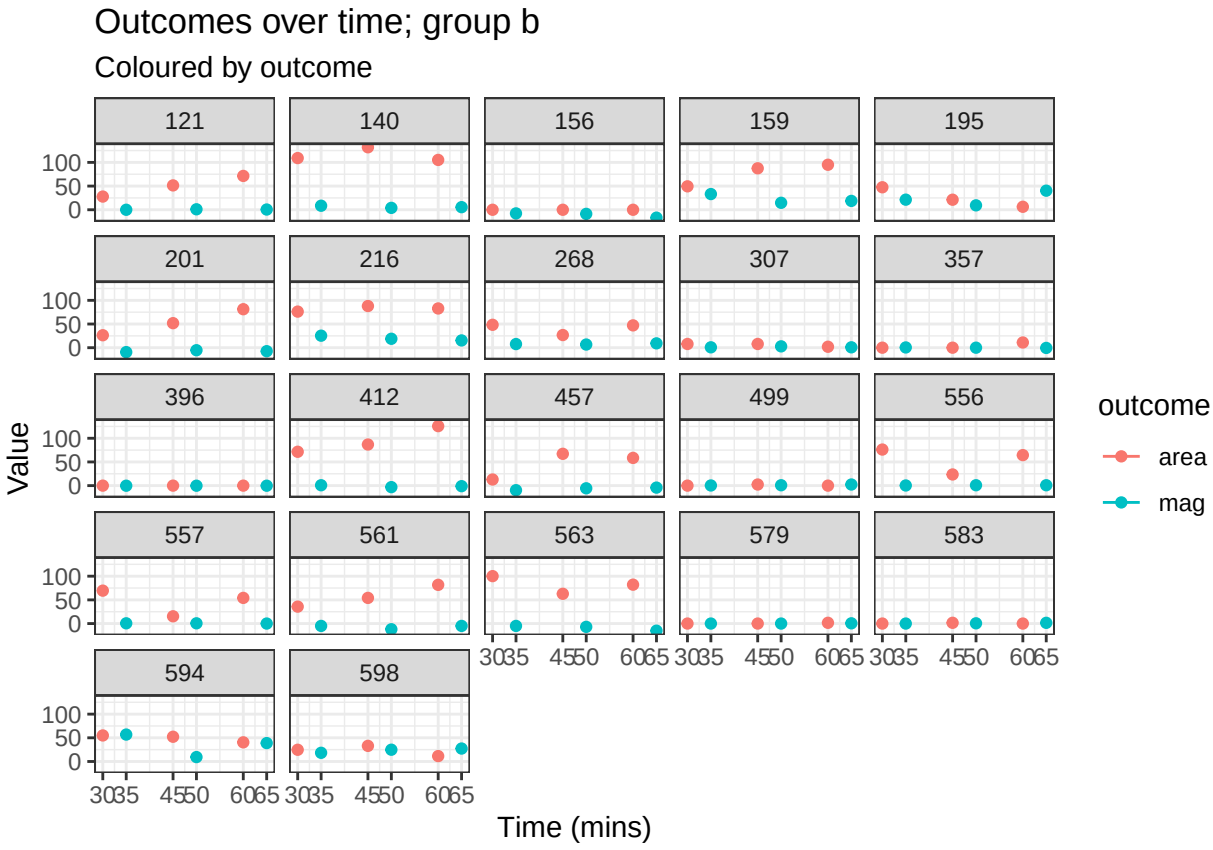
```
## Warning: Removed 174 rows containing missing values (geom_point).
```

```
## Warning: Removed 2 row(s) containing missing values (geom_path).
```




```
both %>%
  filter(group == "b") %>%
  ggplot(.) + aes(x = time, y = value, group = outcome, fill = outcome,
  colour = outcome) + geom_point() + geom_line(aes(group = outcome)) +
  facet_wrap(~pid, nrow = 5) + labs(title = "Outcomes over time; group b",
  subtitle = "Coloured by outcome", x = "Time (mins)", y = "Value") +
  scale_x_continuous(breaks = c(30, 35, 45, 50, 60, 65)) +
  theme_bw()
```

```
## Warning: Removed 132 rows containing missing values (geom_point).
## Removed 2 row(s) containing missing values (geom_path).
```



Unplanned analysis 1 summary:

- Our plot shows that in both groups the surface area and magnitude peak did not consistently correspond temporally. In group 'a', the surface area peak was greater than the magnitude peak in 22 of 29 participants. In group 'b', the surface area peak was greater than the magnitude peak in 13 of 22 participants.

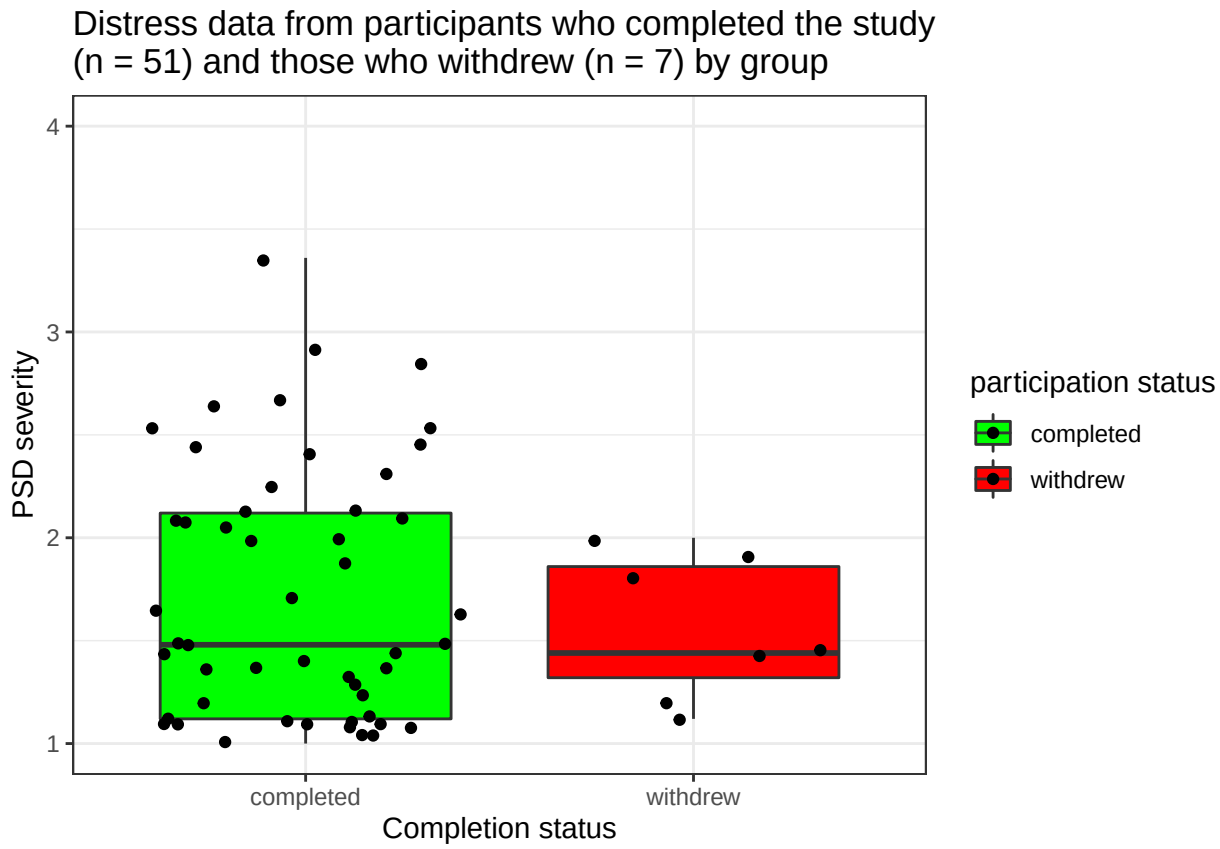
Unplanned exploratory analysis 2: Did the withdrawal of participants who cited the induction as painful bias our results?

After seven participants withdrew from the study citing the painfulness of the induction procedure, we wished to ascertain whether this had biased our results. Therefore, we will plot PSD by completion status to shed light on the possibility that we had lost results from a particularly distressed subgroup.

```
began_procedure %>%
  mutate(`participation status` = case_when(str_detect(participation_status,
    "completed") ~ "completed", str_detect(participation_status,
    "withdrew") ~ "withdrew", str_detect(participation_status,
    "missing") ~ "completed", str_detect(participation_status,
    "excluded") | str_detect(participation_status, "ineligible") ~
    "excluded")) %>%
  filter(`participation status` != "excluded") %>%
  ggplot(.) + aes(x = `participation status`, y = distress_tot,
  fill = `participation status`) + scale_fill_manual(values = c("green",
```

```
  "red")) + geom_boxplot(outlier.shape = NA) + geom_jitter() +
  labs(title = "Distress data from participants who completed the study\n(n = 51) and those who withdrew
    x = "Completion status", y = "PSD severity") + scale_y_continuous(limits = c(1,
  4)) + theme_bw()  # PSD range for those who withdrew: 1.12 to 2
```

```
## Warning: Removed 1 rows containing missing values (geom_point).
```




```
# How many participants scored more than 2 for PSD in our
# sample?
psd_more_than2 <- began_procedure %>%
  filter(distress_tot > "2") %>%
  nrow()
glue("{psd_more_than2} of 51 participants scored more than 2 for PSD severity, which is greater than the
```

19 of 51 participants scored more than 2 for PSD severity, which is greater than the PSD severity reported by the seven participants who withdrew (PSD severity ranged between 1.12 to 2).

Unplanned exploratory analysis 2 summary We plotted PSD severity plotted by completion status. Of the seven participants who withdrew PSD severity ranged between 1.12 to 2. Therefore, the withdrawal of these participants did not bias our results.

Conclusions

We aimed to investigate the relationships between self-reported PSD and SH in people with HIV who reported persistent pain or had no pain. We hypothesised that we would observe a positive relationship between PSD

and SH outcomes (surface area and magnitude) in both groups and that the persistent pain group would have a greater surface area and magnitude of experimentally induced SH than the no pain group.

- The induction check confirmed that our induction was successful because our sample showed evidence of the surface area (in 39 of 51) and magnitude (21 of 51) of experimentally induced SH; however, the SH was short-lived.
- Blinding of participants to study aim and Researcher 2 to group membership was maintained. Consequently, we can explore the relationships between PSD, SH and persistent pain.
- There was a positive relationship between the surface area and PSD ($F(1,49) = 6.49$, $p=0.01$).
- Our plots suggest that group b showed a stronger positive relationship between surface area and PSD than group a (see Primary analysis above).
- The surface area of SH was not predicted by group membership ($p = 0.93$).
- The surface area of SH was observed more frequently in group 'a' (24 of 29 participants) than in group 'b' (15 of 22) (see SH data preparation: plots on group surface area overtime faceted by participant).
- For the surface area of SH the 'sa_model_main' was found to be a suitable fit for our data, although the normality of residuals assumption may still need to be met (see Primary analysis model assumptions (sa_model_main)).
- There was no relationship between PSD and the magnitude of SH ($F(1,49) = 2.84$, $p=0.09$).
- Plots illustrating the relationship between magnitude and PSD severity, across the sample suggested a possible difference between magnitude and PSD. When we plotted this relationship by group, we observed that group b showed a stronger positive relationship between magnitude and PSD than group a (see Secondary analysis above). However, we couldn't confirm this because we couldn't find a suitable statistical model.
- The magnitude of SH was observed more frequently in group 'b' (13 of 22 participants) than in group 'a' (8 of 29).
- For magnitude of SH, eight participants displayed hypoalgesia: three in group 'a' (232, 481, and 591) and five in group 'b' (156, 201, 457, 561 and 563). Notably, of the eight participants who developed hypoalgesia, seven developed an apparent area of SH, which was relatively small in participants of this subset who were in group 'a' but large in those in group 'b'.
- None of the models for magnitude of SH were suitable for our data.
- Plots on PSD severity and participation status show that PSD severity was not different between those who withdrew or completed the study.
- Our plot shows that in both groups the surface area and magnitude peak did not consistently correspond temporally: the surface area peak was greater than the magnitude peak in 35 of 51 participants.
- The withdrawal of participants who cited the induction as painful did not bias our results.