American Pain Society
RESEARCH
EDUCATION
TREATMENT
ADVOCACY
PUBLISHED BY ELSEVIER

The Journal of Pain, Vol 15, No 10 (October), 2014: pp 1008-1014
Available online at www.jpain.org and www.sciencedirect.com

Test-Retest Reliability of Pain-Related Brain Activity in Healthy Controls Undergoing Experimental Thermal Pain

Janelle E. Letzen,* Landrew S. Sevel,* Charles W. Gay,† Andrew M. O'Shea,* Jason G. Craggs,* Donald D. Price,‡ and Michael E. Robinson*

*Departments of *Clinical and Health Psychology, †Rehabilitation Sciences, and ‡Oral and Maxillofacial Surgery, University of Florida, Gainesville, Florida.*

Abstract: Although functional magnetic resonance imaging (fMRI) has been proposed as a method to elucidate pain-related biomarkers, little information exists related to psychometric properties of fMRI findings. This knowledge is essential for potential translation of this technology to clinical settings. The purpose of this study was to assess the test-retest reliability of pain-related brain activity and how it compares to the reliability of self-report. Twenty-two healthy controls (mean age = 22.6 years, standard deviation = 2.9) underwent 3 runs of an fMRI paradigm that used thermal stimuli to elicit experimental pain. Functional MRI summary statistics related to brain activity during thermal stimulation periods were extracted from bilateral anterior cingulate cortices and anterior insula. Intraclass correlations (ICCs) were conducted on these summary statistics and generally showed "good" test-retest reliability in all regions of interest (ICC range = .32–.88; mean = .71); however, these results did not surpass ICC values from pain ratings, which fell within the "excellent" range (ICC range = .93–.96; mean = .94). Findings suggest that fMRI is a valuable tool for measuring pain mechanisms but did not show an adequate level of test-retest reliability for fMRI to potentially act as a surrogate for individuals' self-report of pain.

***Perspective:** This study is one of the first reports to demonstrate the test-retest reliability of fMRI findings related to pain processing and provides a comparison to the reliability of subjective reports of pain. This information is essential for determining whether fMRI technology should be potentially translated for clinical use.*

***Key words:** Functional magnetic resonance imaging, test-retest reliability, anterior cingulate cortex, insula.*

Recent papers have proposed functional magnetic resonance imaging (fMRI) as a method to elucidate an objective biomarker for the diagnosis of pain syndromes.[1,5,6,32,33] One justification made for using such a biomarker in clinical practice is that the self-report of pain is unreliable, which can make diagnosis and treatment difficult. However, studies have found high test-retest reliability of subjective pain ratings for both acute[4] and chronic[17] pain, whereas this information is lacking for pain neuroimaging.[28] The reliability of fMRI findings in the study of pain is essential to determine before potential translation of this technology to clinical practice, as reliability establishes the upper bound for validity.

Test-retest reliability is a measure of the extent to which scores are consistent and free from error.[19] It is important to note that an individual's pain experience is not static over time, and pain intensity or unpleasantness can fluctuate.[28] However, the degree to which these pain ratings vary is predictable and therefore does not represent error.[30] Although the experience of pain can vacillate over time, subjective pain measures have been shown to reliably capture an individual's pain experience. Williamson and colleagues[34] conducted a meta-analysis of reliability studies on 3 commonly used rating scales: visual analog scales (VASs), numerical rating scales, and verbal rating scales. The authors concluded that all 3 instruments had strong reliability across studies

Received April 18, 2014; Revised June 19, 2014; Accepted June 24, 2014.
This research was funded by the National Center for Complementary and Alternative Medicine through grants to M.E.R. (R01AT001424-05A2) and J.E.L. (F31AT007898-01A1).
The authors have no conflicts of interest to report.
Address reprint requests to Michael E. Robinson, PhD, University of Florida, 101 South Newell Drive, Rm 3151, P.O. Box 100165, Gainesville, FL 32610-9165. E-mail: merobin@ufl.edu
1526-5900/$36.00

http://dx.doi.org/10.1016/j.jpain.2014.06.011

and were acceptable for clinical use. VASs showed the highest reliability across time points, with intraclass correlation (ICC) coefficients ranging from .97 to .99.[4,15] Jones and colleagues[18] found high repeatability of self-reported pressure pain thresholds across 4 consecutive days in pain-free women, with kappa coefficients ranging from .94 to .97. Commonly used questionnaires, such as the McGill Pain Questionnaire, have also been shown to have high reliability, with ICC coefficients ranging from .89 to .96 for total, sensory, affective, and average pain scores.[16] In general, higher reliability is found within shorter time spans because of the fluctuation of pain itself.[19]

Although the reliability of subjective pain report has been examined among different measures, there is a paucity of reliability studies in pain neuroimaging. In general, few studies have examined the test-retest reliability of fMRI data, which could be lower than commonly expected in the field.[3] Brain activity within the default mode network demonstrated good reproducibility over 3 separate time points.[22] However, Brandt and colleagues[7] examined the test-retest reliability of a novelty encoding paradigm and concluded that results were difficult to interpret at the single-subject level because of poor reliability. In a meta-analysis of fMRI reliability studies, Bennett and Miller[3] found an average ICC coefficient of .5 across several cognitive tasks in healthy controls. The authors also concluded that test-retest reliability was typically poorer among studies of clinical populations.

Very few studies have examined the test-retest reliability of brain activity associated with pain processing, which is imperative to determine before this technology can be used clinically, as suggested by others.[5,6,20,33] The purpose of the present study was to measure the test-retest reliability of 2 brain regions associated with pain processing in healthy controls and to compare it with the reliability of self-report. A meta-analysis examining fMRI studies of experimental pain showed that bilateral anterior insula (aINS) and right anterior cingulate cortex (ACC), among 2 other regions, had the highest likelihood of being activated by noxious stimuli.[13] We limited our a priori regions of interest (ROIs) to the ACC and aINS because they were suggested to best reflect pain perception[1] and sensitivity to changes in self-report,[8] respectively. Although the ACC is involved in the affective component of pain processing, the aINS is involved in both affective and cognitive-evaluative components of pain.[13]

Methods

This study is a secondary data analysis from a larger, NIH-funded fMRI project in process examining mechanisms and temporal properties of placebo analgesia. For the parent study, individualized "pain" and "placebo" temperature thresholds were established during a screening visit using VAS responses to thermal quantitative sensory testing outside of the MRI scanner. Individuals who qualified to participate in the study then completed 1 baseline fMRI visit wherein only thermal "pain" temperatures were applied, with no placebo conditioning or other manipulation. This baseline visit was followed by 2 additional scanning sessions, each separated by 1 week, and participants underwent placebo conditioning prior to either the second or third scanning session. Before each scanning session, participants completed the State-Trait Anxiety Inventory and the Pennebaker Inventory of Limbic Languidness. Data included in the present analyses were from the parent study's baseline visit and represent only brain activity and self-report associated with thermal, experimental pain. Methods described below represent procedures used for the baseline visit.

Participants

Data from 22 healthy, pain-free individuals were analyzed in this study (mean age = 22.6 years, standard deviation = 2.9; 13 females). Nine participants identified as white, 4 as Asian, 5 as Hispanic, and 4 as African American. Participants were excluded if they met the following criteria: 1) current enrollment in another research study that could influence participation in the present study, 2) use of pain-related medications that could not be stopped 7 days prior to testing (eg, nonsteroidal anti-inflammatory drugs, antihistamines, antidepressants, anticonvulsants, migraine medications, and cough suppressants), 3) history of psychiatric, psychological, or neurologic disorder, as well as medical conditions associated with chronic pain, 4) current medical condition that could affect study participation, 5) positive pregnancy test result in females, 6) presence of ferromagnetic metal within the body, and 7) inability to provide informed consent. The parent study was approved by the University of Florida Institutional Review Board. All participants provided written informed consent.

Experimental Materials

Thermal stimuli during fMRI scanning periods were delivered using an MR-compatible, Peltier-element-based stimulator (Medoc Thermal Sensory Analyzer, TSA-2001; Medoc, Ramat Yishai, Israel). Temperatures produced by this device range from 33°C to 51°C. Participants reported subjective pain ratings to these stimuli using a computerized VAS, anchored by "no pain" and "the most imaginable pain."

Experimental Procedures

The present study used a within-subjects design to assess the test-retest reliability of pain-related brain activity across 3 fMRI runs. Because of individual differences in pain perception, each participant completed quantitative sensory testing during a screening visit prior to baseline fMRI scanning. Thermal pulses were delivered on the dorsal aspect of each foot, beginning at 43°C and increasing by 1°C until tolerance or 51°C was reached. Participants rated pain intensity on a VAS after each pulse. Temperatures for "pain" stimuli used during the baseline fMRI visit were determined for each individual based on the lowest temperature rated between 40°C and 60°C.