Original Experimental

Elisabeth Ørskov Rotevatn*, Mette Engan, Emilie Stensaker, Karl Ove Hufthammer, Lars Jørgen Rygh

# Measuring pain intensity through physical interaction in an experimental model of cold-induced pain: A method comparison study#

https://doi.org/10.1515/sjpain-2024-0011
received February 13, 2024; accepted April 16, 2024

**Abstract**

**Objectives** – Assessment of pain is challenging given its subjective nature. Standard pain assessment tools have limitations. We aimed to compare the verbal numeric rating scale (NRS) and Grasp, a novel handheld electronic device that reports pain by squeezing.

**Methods** – To compare Grasp and NRS, healthy adult volunteers were invited to undergo two subsequent standardised tests of cold-triggered pain using a cold pressor test (CPT) at a temperature of 3°C. Pain intensity was in a randomised manner reported by NRS (scale 0–10) or by squeezing Grasp (0–3 V) during the two CPTs. A third CPT was performed 1 to 14 days later where subjects reported pain by Grasp a second time in order to study the association of repeated Grasp measurements. Acceptable association was *a priori* considered as mean Kendall's $\tau$–$b$ coefficient ($\tau$–$b$) ≥ 0.7. The subjects reported their experience of using Grasp in a purpose-made questionnaire.

**Results** – In total, 102 subjects were included, and 96 subjects (56 females) completed all three tests. The association of pain intensity reported by Grasp and NRS was moderate with a mean $\tau$–$b$ of 0.53 (95% confidence interval [CI] 0.47–0.58). The association between the repeated Grasp measurements was weak with a mean $\tau$–$b$ of 0.43 (95% CI 0.37–0.48). Most subjects reported that Grasp was intuitive and easy to use.

**Conclusions** – Pain intensity reported by squeezing Grasp did not show acceptable association with pain intensity reported by NRS during CPTs. The association between pain intensity reported by Grasp during two CPTs on separate days was weak. Further improvements of the Grasp ball are needed before use in clinical settings.

**Keywords:** numeric rating scale, cold pressor test, patient-reported experience measures, Kendall's tau, randomisation

## 1 Introduction

Pain is defined as an "unpleasant sensory and emotional experience associated with, or resembling that associated with, actual or potential tissue damage" [1]. Hence, pain is always subjective and difficult to measure objectively precise. Still, as accurate evaluation of a patient's pain as possible is important for enabling optimal treatment, prevention of complications, and good long-term prognosis [2]. For patients with cognitive impairment, language issues, and for children, assessment of subjective symptoms can be particularly difficult [3–6].

A variety of tools assessing different aspects of pain have been developed and tested. The Numeric Rating Scale (NRS) and the visual analogue scale (VAS) are considered as gold standard methods for pain assessment in most situations [7]. Other tools are available for pain intensity assessment in children or cognitively impaired patients, e.g., the Face, Legs, Activity, Cry, Consolability Scale [8]. Since pain is a multifaceted and complex phenomenon, the available tools have different limitations, and no single tool or method is able to capture all dimensions of pain, including pain intensity [7]. Two major limitations in using

# Previous presentations: Some results from the study were previously presented as a poster at the EFIC 2023 congress in Budapest, Hungary, September 21, 2023.

*** Corresponding author: Elisabeth Ørskov Rotevatn,** Department of Anaesthesia and Intensive Care, Haukeland University Hospital, Postboks 1400, 5021, Bergen, Norway; Children and Youth Clinic, Haukeland University Hospital, Bergen, Norway; Department of Clinical Science, University of Bergen, Bergen, Norway,
e-mail: eoro@helse-bergen.no, tel: +4748341985
**Mette Engan:** Children and Youth Clinic, Haukeland University Hospital, Bergen, Norway
**Emilie Stensaker, Lars Jørgen Rygh:** Department of Anaesthesia and Intensive Care, Haukeland University Hospital, Postboks 1400, 5021, Bergen, Norway
**Karl Ove Hufthammer:** Centre for Clinical Research, Haukeland University Hospital, Bergen, Norway

standard tools are the need for advanced verbal communication skills and the ability to accurately recall previous pain [3]. Thus, development of additional and improved pain assessment tools is warranted.

A few previous studies have explored whether pain intensity can be reported through physical interaction. The concept builds on the idea of transferring the feeling of pain to another modality (handgrip strength) that can be obtained as a more objective proxy for pain. Hand dynamometers and a few other handheld devices with different technologies have shown some potential in smaller experimental studies [9–14]. These findings indicate a possibility of assessing pain intensity through handgrip strength with further development of these devices.

Recently, Grasp was developed. This is a novel device enabling patients to communicate pain nonverbally and whenever their symptoms occur. Grasp consists of a soft silicone ball with an inner pressure sensor that connects to software via Bluetooth (Figure 1). When squeezing the Grasp ball by hand, data on squeeze duration, force, and time point are stored in the software. Variation in squeeze force may represent the individual's variation in symptom intensity. The ability to express pain nonverbally and concurrently with symptom occurrence may prove useful. Hence, the Grasp system may be a valuable addition to existing pain assessment tools. The Grasp system's ability to assess pain intensity has not yet been investigated. The main objective of this study was to compare assessment of pain intensity by squeezing the Grasp ball to the gold standard method of the verbal NRS during a cold pressor test (CPT). The secondary objective was to assess the repeatability of pain intensity assessment using the Grasp ball during CPTs. Furthermore, we aimed to explore the study subjects' experience of using Grasp for pain assessment.

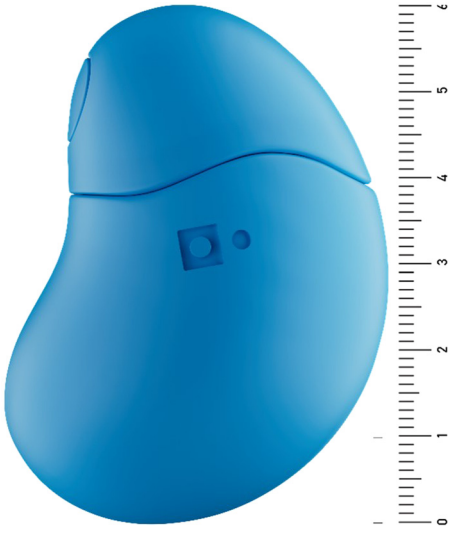


**Figure 1:** The Grasp device with a centimetre ruler.

## 1.1 Methods

To document the ability of Grasp to assess pain intensity via physical interaction, we designed an experimental method comparison study between Grasp and the verbal NRS (0–10).

## 1.2 Subjects

Included subjects were overtly healthy, between 18 and 60 years, and fluent in Norwegian language. Subjects were excluded if they had (a) any somatic condition except mild asthma, atopic eczema not affecting hands, or mild food or respiratory allergy; (b) any chronic or recurrent condition associated with pain; (c) previous hand surgery or other hand injuries; (d) any psychiatric disease requiring medical treatment ongoing or during the last 30 days; (e) history of syncopating if exposed to unpleasant stimuli; (f) pain medication consumption regularly or the last 48 h; or (g) alcohol consumption the last 24 h.

Based on sample size estimations (Section 1.6), we aimed for 100 volunteers to complete the study. We recruited subjects through advertisement at the University of Bergen, Haukeland University Hospital, and social media.

## 1.3 Grasp

Grasp is a soft bean-shaped ball containing a force-sensing resistor (FSR), a silicone inner plum and outer layer (Figure 1). Squeezing Grasp reduces FSR resistance, generating a voltage output (0–3 V) recorded every 0.2 s and reported to software via Bluetooth. We used six similar Grasp devices (version Hw1.2,60A,00-30, no-fluor). Grasp is manufactured by the Grasp AS company.

## 1.4 CPT

The CPT is a standardised experimental method for inducing pain reflecting key components of clinical pain [15,16]. Subjects participated in three CPTs submerging their non-

dominant hand in a 3°C water bath (FP40HE, Julabo Lobortechnik GmbH, Germany) for a maximum of 110 s. All participants sat upright in an armchair. The Grasp was held in the dominant hand with adduction in the shoulder and 90° forward in the elbow joint. Pain intensity was reported every ninth second, either verbally by NRS or physically by squeezing Grasp (increasing squeeze strength for increasing pain intensity). Pain intensity was expected to intensify during CPTs. If the pain became unbearable, subjects could withdraw their hand, recorded as pain tolerance time.

Before each CPT, subjects received standardised instructions on the test and the pain reporting method they were to use. For Grasp, each subject calibrated the ball by squeezing with maximum force, representing worst pain imaginable. For further education, they squeezed lightly for "mild pain," with medium strength for "moderate pain," and hard for "severe pain." The study nurse assessed whether the subjects managed to calibrate and squeeze appropriately during the training. During CPTs, subjects received no feedback on their reported pain levels. For NRS, ranging from 0 (no pain) to 10 (worst pain imaginable), subjects confirmed understanding by indicating appropriate numbers for mild, moderate, and severe pain.

On the first day of testing, each subject did two CPTs, one for each reporting method with a minimum 3-h gap. Subjects were randomised between NRS or Grasp starts. Within 14 days, a third CPT was performed, reporting pain intensity exclusively with Grasp (Figure S1, Supplementary Material).

Pain reports from NRS CPTs were recorded by a study nurse during tests, while for Grasp CPTs reports were saved in the software and extracted to data files after tests.

## 1.5 User experience

Participants were invited to complete a purpose-made questionnaire regarding utility and functionality of the pain assessment tools. The questionnaire, developed by literature review and inputs from user representatives and pain experts, featured statements rated on a five-point scale. Subjects also had the chance to provide open-ended comments. Graphs depicted pain reports for both tools during CPTs, and participants assessed which graph most accurately portrayed their pain.

## 1.6 Statistics

The results of questionnaires are reported as percentage of subjects choosing each response option.

To transform measurements done using Grasp into the NRS scale, we wanted an at least moderate-strength (possibly nonlinear) association between the two measurements. We decided to measure strength of association using Kendall's $\tau$–$b$ (tau-b), a scale-invariant measure capable of capturing also non-linear associations.

Here follows a slightly simplified explanation of how to calculate $\tau$–$b$ for Grasp vs NRS: We look at all *pairs* of measurements, i.e., measurements taken at two different time points for the same subject. For each pair, if the time point for the *highest* of the two Grasp values is also the time point for the *highest* of the two NRS values, we have a *concordant* pair. In the opposite situation, i.e., the time point for the *highest* of the two Grasp values is the time point for the *lowest* of the two NRS values, we have a *discordant* pair. We then calculate the proportion of concordant pairs minus the proportion of discordant pairs and get a value between −1 and 1. Higher values are better. Kendall's $\tau$–$b$ is calculated like this, but with a small correction in the case of ties. See Figure S2 (Supplementary Material) for visual explanation.

*A priori*, we would prefer high values of $\tau$–$b$ for all subjects. However, considering that NRS and Grasp values were reported for different CPTs, we decided that a mean (i.e., average) $\tau$–$b$ value of 0.7 or higher would be acceptable in this initial study. To summarise the association between NRS and Grasp, we report the distribution of $\tau$–$b$ values, the mean value, and the corresponding 95% confidence interval (CI). The Grasp values used in calculations were maximal voltage output during a squeeze on Grasp. The CI was calculated based on the *t*-distribution.

We hypothesised a mean $\tau$–$b$ value of at least 0.7, and the sample size calculation was based on this reasoning: If most subjects had high $\tau$–$b$ values, a reasonable standard deviation would be 0.15. For 100 subjects, we could then report the mean $\tau$–$b$ with a precision of about 0.03 (i.e., the 95% CI would be: estimate ±0.03), which we deemed acceptable.

Since increase in reported pain intensity mostly occurred during the first half of CPTs, we did sub-analyses with calculation of mean $\tau$–$b$ values between NRS and Grasp and between repeated Grasp reports of this part alone.

Repeatability of pain assessments with Grasp was examined using Bland–Altman plots for each of the 12 measurement points. We also report the repeatability coefficient (i.e., 2.77 times the within-subject standard deviation) and the intraclass correlation coefficient (ICC). The ICC was calculated using a two-way mixed-effects model for absolute agreement and single measurements. ICC values <0.5 were considered poor, 0.5–0.75 moderate, 0.75–0.9 good, and > 0.90 excellent [17,18]. Statistical analyses were done in R version 4.2.3 [19].

# 2 Results

We screened 117 volunteers for inclusion, and 102 (59 females) of them were subsequently included. The mean age was 26 years (range 18–54 years). All subjects understood the instructions of the CPT and the NRS and managed to calibrate the Grasp unit according to protocol. Except for three, all subjects completed the first and second CPT. A third CTP was completed by 96 subjects (56 females) (Figure S3, Supplementary Material). One subject removed their hand from the water bath before the test was completed due to unbearable pain in all three tests (50, 40, and 65 s, respectively). However, the rest completed all CPTs with 110 s of pain stimulus.

The pattern of median reported pain intensity during the three CPTs was similar, with an increasing median and mean pain intensity the first 54–63 s (up to the time point 6–7) followed by a rather stable median and mean pain intensity during the rest of the test (Figure 2). The mean pain intensity pattern did not differ between randomisation groups. When every single measurement during the first report with Grasp was plotted against the corresponding reported NRS value at the same time point of the CPTs ($n$ = 1,179), there was a clear and steady increasing mean voltage with increasing NRS reports. Nevertheless, the range of different Grasp outputs within one single NRS value was wide (Figure 3). At most, 1% of the subjects reported the maximum NRS score (10) across the 12 time points, while up to 11% (test 1) and 15% (test 2) reported a maximum score (2.99 or 3.00 V) during CPTs with Grasp.

## 2.1 Association between Grasp and NRS

The mean Kendall's $\tau$–$b$ value for pain intensity reported by Grasp vs NRS for the 12 time points was $\tau$–$b$ 0.53 (95% CI 0.47–0.58, median 0.57). Twenty-eight percent of the subjects had an acceptable $\tau$–$b$ (≥0.70). The distribution of $\tau$–$b$ values is shown in Figure S4 (Supplementary Material). Since the increase in pain intensity mostly occurred during the first half of the CPT, individual $\tau$–$b$ was recalculated for this part of the test alone, showing a mean value of 0.60 (95% CI 0.53–0.67, median 0.69). About half of the subjects (46%) had an acceptable $\tau$–$b$ (≥0.70) during the first half of the CPT. For the complete CPTs, there were only minor

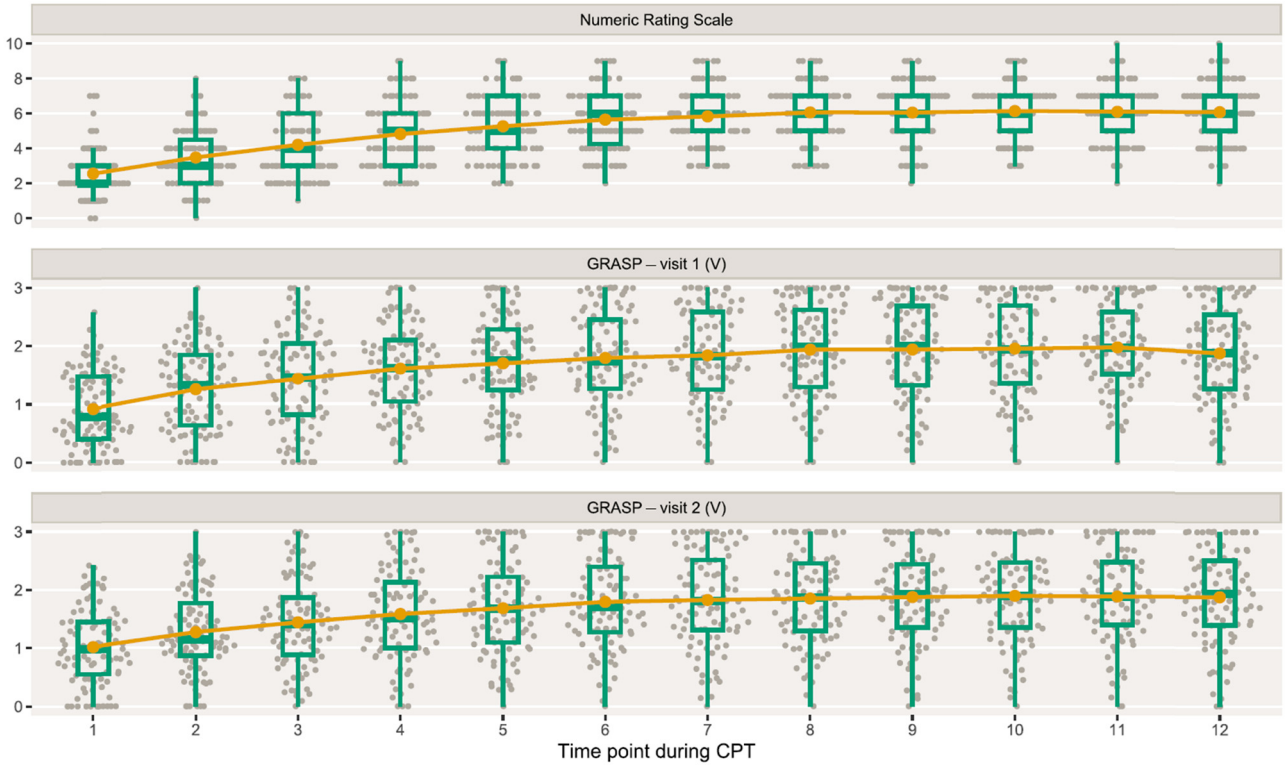


**Figure 2:** Reported pain intensity at the 12 time points during the three CPTs. The points have been jittered horizontally to avoid overplotting. The orange dots represent the mean values. Time points: 1 = 9 s, 2 = 18 s, 3 = 27 s, 4 = 36 s, 5 = 45 s, 6 = 54 s, 7 = 63 s, 8 = 72 s, 9 = 81 s, 10 = 90 s, 11 = 99 s, 12 = 108 s. CPT: Cold pressor test. NRS: Numeric Rating Scale.

differences in mean $\tau$–$b$ for different subgroups: stratification by sex, randomisation group, different Grasp devices, or preference of reporting method (results not shown).

Repeated Grasp measurements ($n$ = 96) showed a mean $\tau$–$b$ of 0.43 (95% CI 0.37–0.48, median 0.43). When considering only the first half of the CPT, the mean $\tau$ was 0.46 (95% CI 0.39–0.54, median 0.56). A proportion of 28% of the subjects had a $\tau$–$b \geq 0.70$ during the first half of the CPT.

The repeatability of pain intensity reports with Grasp was similar across all 12 time points. The Bland–Altman plots had an overall mean difference of approximately 0.0 V (range −0.1 V to + 0.1 V), showing no systematic difference between the first and the second Grasp report. The limits of agreement ranged between ±1.1 and ±1.4 V (Figure 4, Table S1, Supplementary Material). The repeatability coefficient estimates varied between 1.13 and 1.40, and the ICC estimates between 0.59 and 0.72 (Table S1, Supplementary Material).

When calibrating to maximal hand grip strength, 54% reached the maximum of the scale of 2.99 or 3 V.

## 2.2 User experience

One hundred subjects completed the questionnaire addressing user experience and preferences for pain assessment tools. Based on Table 1 and free-text comments, we found that the subjects overall found the Grasp ball intuitive and easy to use. Furthermore, 27% of the subjects reported that they would prefer Grasp for pain reporting before NRS. In contrast, 55% preferred the NRS to Grasp, and 18%

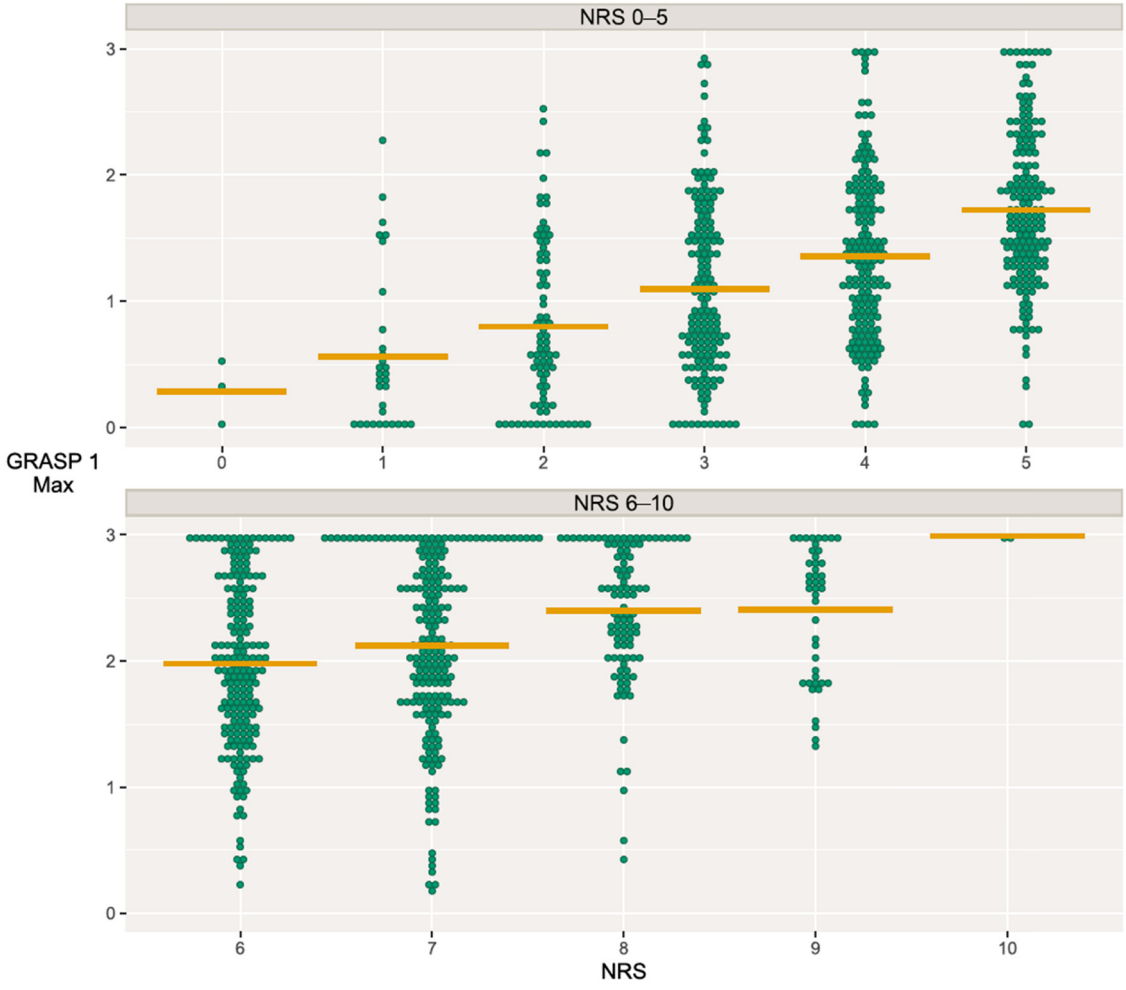


**Figure 3:** Individually reported pain intensity with the Grasp ball (0–3 volts) during a CPT and reported pain intensity with the verbal NRS (0–10) at the corresponding time point for a different CPT. Number of observations: 1,179. Number of subjects: 99. The orange lines represent the mean voltages. GRASP 1 Max: Maximum registered voltage during one squeeze on the Grasp ball (0–3 volts), NRS: Numeric Rating Scale (0–10).

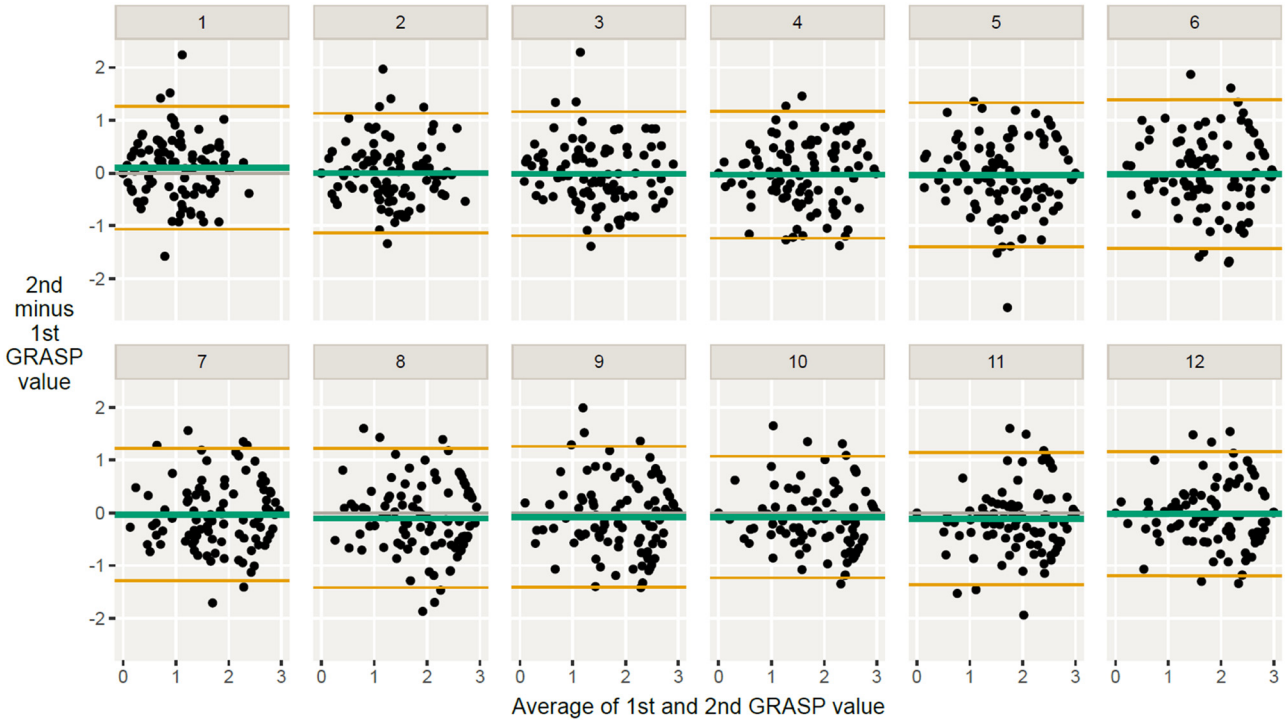


**Figure 4:** Bland–Altman plot showing agreement between the first and second report with Grasp for the 12 time points. The horizontal green lines show the mean difference between the first and second Grasp value. The orange lines show the limits of agreement. Time points: 1 = 9 s, 2 = 18 s, 3 = 27 s, 4 = 36 s, 5 = 45 s, 6 = 54 s, 7 = 63 s, 8 = 72 s, 9 = 81 s, 10 = 90 s, 11 = 99 s, 12 = 108 s.

reported no preference between the two methods (Table 1). When both pain reports were illustrated as graphs, 34% reported that the pain report given with Grasp gave a more correct picture than the NRS report, 51% reported that the pain report given with NRS gave a more correct picture, and 15% found no difference between the two reports (Table 1).

## 3 Discussion

In this experimental comparison study, we found that the association between pain intensity reported on an NRS and through physical interaction via squeezes on the Grasp ball showed a mean $\tau$–$b$ of 0.53, which was lower than our pre-specified cut-off of acceptable value (0.70). There was a high degree of variation in $\tau$–$b$ values. Half of the subjects reported pain with acceptable $\tau$–$b$ during the first half of the CTP where mean pain intensity increased. The $\tau$–$b$ between repeated tests with the Grasp ball was low (mean 0.43), and the ICC estimates were moderate. About a quarter of subjects preferred reporting pain via physical interaction, while about half of the subjects preferred reporting pain using NRS.

We have found no previous studies that tested comparable devices through cold-induced pain. Other studies have used heat pain, electrical stimuli, clinical cancer pain, or imaginary pain [9–14]. Previous smaller studies have tested comparable devices for pain reporting in different ways. One study compared a device called “Pain mouse” to VAS (0–100) ($n$ = 13) and found it able to distinguish four pain levels (no, low, medium, and high pain), but when repeating the test after 1 week, pain reports on all four pain levels were significantly lower [14]. Another study compared a device called “the eEgg” to a hand dynamometer through heat stimuli ($n$ = 50) and found a moderate to strong correlation between applied force to the two devices ($r$ = 0.353–0.614) [10]. A third small study ($n$ = 15) included patients with cancer pain in reporting pain intensity on a finger dynamometer (FD), VAS, and NRS at four time points and found a moderate to strong association between FD and NRS (Kendall’s correlation coefficient = 0.47–0.68) and weak to moderate association between FD and VAS (0.38–0.46) [13].

One important difference between reporting pain with verbalisation of numbers compared to squeezing is that subjects may more easily recall the previous numbers they verbally reported than how hard they squeezed. Changing the number one reported 9 s ago is therefore

**Table 1:** Results from questionnaire

| | Very good (%) | Good (%) | Acceptable (%) | Poor (%) | Very poor (%) | Total (%) |
|---|---|---|---|---|---|---|
| To what extent did you find squeezes on Grasp an easy way to grade pain level? | 6 | 38 | 28 | 7 | 1 | 100 |
| To what extent did you find the numeric scale an easy way to grade pain level? | 25 | 51 | 22 | 1 | 1 | 100 |
| To which degree did squeezes on Grasp express your experience of pain level correctly? | 8 | 42 | 44 | 5 | 1 | 100 |
| To which degree did the numeric scale (NRS) express your experience of pain level correctly? | 24 | 57 | 18 | 0 | 1 | 100 |
| To which degree did you find squeezes on Grasp a good way to express pain level? | 16 | 43 | 37 | 3 | 1 | 100 |
| | **Numbers (NRS) (%)** | | **Squeezes on Grasp (%)** | | **No difference (%)** | **Total (%)** |
| Which method would you prefer for grading and reporting pain level? | 55 | | 27 | | 18 | 100 |
| Which illustration (graph) of your pain intensity reports do you think gives a better picture? | 51 | | 34 | | 15 | 100 |

Number of completed questionnaires: 100. NRS: Numeric Rating Scale.

less likely than changing squeezes if the pain intensity feels equal. One-third of subjects reported that the pain report using Grasp gave a more correct picture than the NRS report when both were illustrated by graphs (Table 1). Some subjects explained this with Grasp being able to catch more details in the pain experience than NRS, which only has 11 levels.

To be able to fully understand a patient's pain, comprehensive examination and knowledge about the individual patient over a longer time span is often needed. In the setting of today's busy healthcare system, efficient and accurate tools for pain measurement will be useful. Both pain reported on a 0–10 number scale, and squeezing this new device, Grasp, represents a considerable reduction of the complex phenomenon of pain. Previous experience of using NRS for pain assessment has raised this particular concern [20,21]. The present study adds to the knowledge about nonverbal reporting of pain through interaction with handheld devices. Even though squeezing is a relatively simple way of communicating, this physical manner of pain reporting may be a valuable supplement and possibly even a preferred tool for some patients, e.g., for children and for patients with limited verbal skills. The questionnaires in this study revealed that the Grasp method was easily adapted, and despite more subjects preferring the NRS, a considerable number of subjects preferred the Grasp. Conveying pain through hand grip may be a richer and more intuitive method than reducing pain experience to a verbally expressed number to some people. In addition to what we explored in the current study, the Grasp system has an inbuilt possibility of letting patients report pain at home or wherever they are, concurrently with the pain experience over time period. With the storage of reports and visualisation of pain trajectory in an informative digital interface, the patient's pain experience may be understood better and more in-depth by both clinicians and patients themselves [22]. Another advantage to the Grasp method over a mobile application with NRS could be that it in some situations (e.g., situations with intense pain or during exercise sessions) seems more convenient than the procedure of logging into the phone, opening the application, and writing pain scores.

We wanted to adjust squeeze outputs to individually calibrated maximal force, allowing recalculation into percentages that easily translate into NRS, but more subjects than expected reached the highest possible output voltage when asked to squeeze to indicate the worst imaginable pain. This shows that a considerable number of subjects had maximal handgrip strength beyond the scale of the Grasp ball, introducing a ceiling effect and making calibration adjustment difficult. Stronger resistance in the ball,

resulting in fewer subjects reaching the ceiling, would, to a better degree, mimic NRS and allow calibration.

For about half of the subjects, there was a strong association between methods when pain intensity increased. But still, finding only a moderate overall association between methods and low to moderate repeatability of Grasp, there is a clear need for improvement of Grasp before usage in clinical settings. Improvements to both physical design, the properties of silicone, and the sensor must be considered, along with education and user instructions. Validation against a hand dynamometer would also be desirable. In the present experiment, subjects were only briefly introduced to the method and tried Grasp for only a few minutes before starting the experiment. It is likely that better instructions and more time to get familiar with the tool could improve results.

Several subjects expressed a wish for a feedback mechanism in the Grasp system that could assure them that the pain intensity they wished to express was correctly obtained. One can think of both feedback mechanisms within the ball itself (e.g., lights in different colours or vibrations) and in an accompanying smartphone application. Such a feedback mechanism could lower demands for improvement of the technical precision of the ball and the method and most likely improve the reliability of the method.

### 3.1 Strengths and limitations

A major strength of this study was the highly controlled pain stimulus of the CPT that has been extensively used in previous studies [15,16]. Even though pain is affected by several uncontrollable factors like genotype, sex, and age, the pain exposure in this study was as similar as possible. Another strength of the study was that Grasp was compared to NRS, which has been validated thoroughly in several settings [7,23,24]. In the Tromsø study, more than 10,000 individuals performed CPTs reporting experienced pain using NRS in the same way as the present study [16,25]. The mean pain intensity reported using the NRS in the Tromsø study was in line with the pain reports using the NRS in this study. But in contrast to the Tromsø study, where about 30% of the subjects withdrew their hand from the water before the test was completed, only one out of 99 subjects did so in our study. This may be explained by the subjects of this study being healthier, younger, and primarily university students and, therefore, more prone to complete tests than subjects selected from the general population.

There are several limitations to this study. First, we only measured subjects' experience of pain intensity reported on the NRS once. Exploring the repeatability of reports with NRS in this particular setting for comparison to the Grasp method could have been informative. Previous studies have indicated that the NRS method is reliable in this setting [16,23–25], but calculating Kendall's $\tau$–$b$ and ICC for NRS, in addition, would have given more assurance to the evaluation of the Grasp method. Second, some subjects indicated different subjective pain experiences in the three tests in their questionnaire. This individual difference in experience may have influenced the repeatability of Grasp reports and the association between the measurements of the two methods. However, there was no *mean* difference in Grasp reports between the two tests. If pain was reported using both tools during the same CPT, variation of experienced pain intensity would have been eliminated. However, concurrent reports using both NRS and Grasp could impact the way subjects squeezed the Grasp ball guided by the NRS and in this way bias the properties of Grasp.

The phenomenon of pain conditioning may have influenced the results by either facilitating or inhibiting pain stimulus in a different way in the first and the second CPT; however, we found no *mean* difference in NRS reports between the two randomisation groups, indicating that pain conditioning played a minor role in this setting. The same applies to the phenomenon of pain modulation; there was a similar pattern of pain intensity development during CPTs measured with Grasp and NRS, suggesting that pain modulation of squeezing a ball was minor in this setting.

Another limitation was that the subjects were somewhat uniform regarding age and socio-economic status and not representative of the general population. Furthermore, the pain stimulus in this study was experimentally applied. Use in patients with a variety of illnesses and disabilities is unexplored.

## 4 Conclusions

Pain intensity reported by squeezing Grasp did not show an acceptable association with pain intensity reported by NRS during CPTs. The association between pain intensity reported by Grasp during two CPTs on separate days was weak. The findings of this study suggest that further improvements of the Grasp ball are needed before use in clinical settings. We suggest adding a feedback mechanism in the ball that will ensure patients correctly reported experienced pain intensity.

**Acknowledgements:** We thank all participants for participating in the pain tests. We also thank the Grasp AS team for software and hardware development and support. Professor Audun Stubhaug at the University of Oslo kindly lend us the ice water bath. Professor Frode Guribye at the University of Bergen contributed to planning the study and discussing implications of the results. Louise Sandal Løkeland, M.Phil. in information science and researcher at the University of Bergen, contributed to planning the study and provided technical support.

**Research ethics:** The Norwegian Medicines Agency and the regional ethics committee approved the study (REK KULMU B number 460216), and it was conducted in concordance with the Helsinki Declaration.

**Informed consent**: All subjects received oral and written information about the study and about the option to withdraw consent at any time. Health risks of the CPT were evaluated before the study, and the CPT was assessed to be a safe and low-risk procedure, with syncopating being the most frequent adverse event.

**Author contributions:** The authors have accepted responsibility for the entire content of this manuscript and approved its submission. LJR, ME, and KOH designed the study. ES and EØR included subjects and conducted the pain tests. KOH and EØR performed the statistical analyses. All authors contributed to the interpretation of the results. EØR wrote the manuscript with valuable inputs from ES, KOH, ME, and LJR. All authors reviewed and approved the final manuscript.

**Competing interests:** The authors state no conflict of interest.

**Research funding:** The study was financed through a grant from the Norwegian Research Council.

**Data availability:** The raw anonymised data can be obtained on request from the corresponding author.

**Artificial intelligence/Machine learning tools:** Not applicable.

**Supplementary Material:** This article contains supplementary material (followed by the link to the article online).

# References

[1] Raja SN, Carr DB, Cohen M, Finnerup NB, Flor H, Gibson S, et al. The revised international association for the study of pain definition of pain: concepts, challenges, and compromises. Pain. 2020;161(9):1976–82.

[2] Sinatra R. Causes and consequences of inadequate management of acute pain. Pain Med. 2010;11(12):1859–71.

[3] Cook AK, Niven CA, Downs MG. Assessing the pain of people with cognitive impairment. Int J Geriatr Psychiatry. 1999;14(6):421–5.

[4] Hauer J, Houtrow AJ. Pain assessment and treatment in children with significant impairment of the central nervous system. Pediatrics. 2017;139(6):85–112.

[5] Ford B, Snow AL, Herr K, Tripp-Reimer T. Ethnic differences in nonverbal pain behaviors observed in older adults with dementia. Pain Manag Nurs. 2015;16(5):692–700.

[6] Ali PA, Johnson S. Speaking my patient's language: bilingual nurses' perspective about provision of language concordant care to patients with limited English proficiency. J Adv Nurs. 2017;73(2):421–32.

[7] Breivik H, Borchgrevink PC, Allen SM, Rosseland LA, Romundstad L, Breivik Hals EK, et al. Assessment of pain. Br J Anaesth. 2008;101(1):17–24.

[8] Merkel SI, Voepel-Lewis T, Shayevitz JR, Malviya S. The FLACC: a behavioral scale for scoring postoperative pain in young children. Pediatr Nurs. 1997;23(3):293–7.

[9] Adams AT, Murnane EL, Adams P, Elfenbein M, Chang PF, Sannon S, et al. Keppi: A tangible user interface for self-reporting pain. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems; 2018.

[10] Böing-Meßing DM, Tomschi F, Cegla T, Hilberg T. The eEgg: evaluation of a new device to measure pain. Front Physiol. 2022;13:832172.

[11] Gracely RH, Dubner R. Reliability and validity of verbal descriptor scales of painfulness. Pain. 1987;29(2):175–85.

[12] Gracely RH, McGrath P, Dubner R. Validity and sensitivity of ratio scales of sensory and affective verbal pain descriptors: manipulation of affect by diazepam. Pain. 1978;5(1):19–29.

[13] Wilkie D, Lovejoy N, Dodd M, Tesler M. Cancer pain intensity measurement: concurrent validity of three tools—finger dynamometer, pain intensity number scale, visual analogue scale. Hosp J. 1990;6(1):1–13.

[14] Schaffner N, Folkers G, Käppeli S, Musholt M, Hofbauer GF, Candia V. A new tool for real-time pain assessment in experimental and clinical environments. PLoS One. 2012;7(11):e51014.

[15] Treister R, Nielsen CS, Stubhaug A, Farrar JT, Pud D, Sawilowsky S, Oaklander AL. Experimental comparison of parametric versus nonparametric analyses of data from the cold pressor test. J Pain. 2015;16(6):537–48.

[16] Samuelsen P-J, Nielsen CS, Wilsgaard T, Stubhaug A, Svendsen K, Eggen AE. Pain sensitivity and analgesic use among 10,486 adults: the Tromsø study. BMC Pharmacol Toxicol. 2017;18(1):1–8.

[17] Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med. 2016;15(2):155–63.

[18] Vaz S, Falkmer T, Passmore AE, Parsons R, Andreou P. The case for using the repeatability coefficient when calculating test–retest reliability. PloS One. 2013;8(9):e73990.

[19] R-Core-Team R. A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2023. https://wwwR-projectorg/.

[20] Baamer RM, Iqbal A, Lobo DN, Knaggs RD, Levy NA, Toh LS. Utility of unidimensional and functional pain assessment tools in adult postoperative patients: a systematic review. Br J Anaesth. 2022;128(5):874–88.

[21] Wood BM, Nicholas MK, Blyth F, Asghari A, Gibson S. Assessing pain in older people with persistent pain: the NRS is valid but only provides part of the picture. J Pain. 2010;11(12):1259–66.
[22] Chapman CR, Donaldson GW, Davis JJ, Bradshaw DH. Improving individual measurement of postoperative pain: the pain trajectory. J Pain. 2011;12(2):257–62.
[23] Ferreira-Valente MA, Pais-Ribeiro JL, Jensen MP. Validity of four pain intensity rating scales. Pain. 2011;152(10):2399–404.
[24] Mitchell LA, MacDonald RA, Brodie EE. Temperature and the cold pressor test. J Pain. 2004;5(4):233–7.
[25] Kambur O, Kaunisto MA, Winsvold BS, Wilsgaard T, Stubhaug A, Zwart JA, et al. Genetic variation in P2RX7 and pain tolerance. Pain. 2018;159(6):1064.