Is the modified Harris hip score valid and responsive instrument for outcome assessment in the Indian population with pertrochanteric fractures?

Karthik Vishwanathan*, Keyur Akbari¹, Amit J. Patel

Department of Orthopaedics, Pramukhswami Medical College and Shri Krishna hospital, Gokal Nagar, Karamsad, 388325, India

ARTICLE INFO

Keywords:

Validity
Responsiveness
Psychometric properties
Ceiling floor effect
Harris hip score
Modified harris hip score
Pertrochanteric fracture
Intertrochanteric fracture
Internal fixation
Proximal femoral nail

ABSTRACT

Introduction: The original Harris hip score (HHS) does not contain ability to perform squatting and sitting cross legged as items in the questionnaire and hence a need was felt to modify the Harris hip score so that it could stay relevant in functional assessment of Indian patients in the rural setting. Validity, responsiveness and ceiling floor effect of the Harris hip score after internal fixation of pertrochanteric fracture has not been previously described. The objective of the study was to investigate construct validity, responsiveness and ceiling floor effects of the modified Harris hip score (mHHS).

Methods: For evaluation of construct validity two hypotheses were formulated: first, there would be no difference in mHHS in cohort of patients treated with short or long proximal femoral nail and second, patients younger than 65 years will have higher mHHS compared to patients older than 65 years postoperatively. Proportion of patients obtaining lowest score of 0 point (floor effect) and those obtaining highest score of 100 points (ceiling effect) was evaluated at one, three and six months postoperatively. It is recommended that the proportion of ceiling and floor effect should be lower than 15% in order to deduce satisfactory internal and content validity of an outcome instrument. Responsiveness was evaluated using distribution based methods (effect size and standardised response mean) and anchor based method (area under the curve using receiver operating curve). Ability to perform cross leg sitting and squatting at six months postoperatively were chosen as two different external anchors. Effect size and standardised response mean values higher than 0.80 and area under the curve value higher than 0.70 are indicators of adequate responsiveness of an outcome instrument.

Results: Eighty one consecutive patients with pertrochanteric hip fractures and treated with long and short proximal femoral nail were included in this prospective observational study. Six patients were lost to follow-up due to mortality and complete functional outcome data was available in 75 patients (92.6%). The mean age was 68 years (range: 38–89 years). The mHHS at one, three and six months postoperatively was 39.9 ± 9.5 , 61.6 ± 14.7 and 81.0 ± 15.9 respectively. The improvement in mHHS was significant at all time intervals. In accordance with the hypothesis, there was no significant difference in mHHS at one, three and six months postoperatively in patients treated with short or long proximal femoral nail. In accordance with the hypothesis, patients younger than 65 years had significantly better mHHS at one, three and six months postoperatively as compared to patients older than 65 years. There were no floor or ceiling effects at one, three and six months postoperatively. mHHS showed adequate internal responsiveness (Effect size = 4.34; standardised response mean = 4.26) and adequate external responsiveness (Area under curve = 0.77 and 0.89) using different external anchors.

Conclusion: The mHHS has adequate construct validity, internal validity and responsiveness to evaluate functional outcome of intramedullary nail fixation in pertrochanteric hip fractures in the Indian population.

1. Introduction

The Harris hip score (HHS) is a joint specific score that is completed

by both the clinician and the patient and consists of 10 items covering domains of pain, function, functional activities, deformity and hip range of motion. The HHS was initially described for assessment of

Abbreviations: AUC, Area under the curve; ES, Effect size; HHS, original Harris hip score; Long PFN, long proximal femoral nail; mHHS, modified Harris hip score; SRM, standardised response mean; Short PFN, short proximal femoral nail; ROC, receiver operating curve

* Corresponding author.

E-mail addresses: karthikvg@charutarhealth.org (K. Vishwanathan), keyur21088@yahoo.co.in (K. Akbari), dramit2000@hotmail.com (A.J. Patel).

¹ Present address: CHARUSAT Hospital, CHARUSAT Healthcare campus, Changa, Petlad, 388421, India.

<https://doi.org/10.1016/j.jor.2017.12.001>

Received 5 September 2017; Accepted 30 December 2017

Available online 03 January 2018

0972-978X/ © 2018 Prof. PK Surendran Memorial Education Foundation. Published by Elsevier, a division of RELX India, Pvt. Ltd. All rights reserved.

functional outcome after mold arthroplasty for posttraumatic arthritis¹. The HHS has been used to evaluate functional outcome after pertrochanteric hip fracture^{2–4} and intracapsular fracture neck of femur⁵. Construct validity of an outcome instrument shows that all the domains of the instrument are indeed measuring what they are supposed to measure. Responsiveness of an outcome instrument is defined as the sensitivity of the instrument to detect clinically relevant change over a period of time⁶. Floor and ceiling effect pertains to the proportion of patients achieving minimum possible score and the maximum possible score respectively⁶. Validity and responsiveness are context specific terms and the evidence is applicable to a particular condition and specific intervention only⁷. The HHS is a valid and reliable outcome instrument to assess functional outcome after total hip replacement⁸. The cohort of patients undergoing surgical treatment for hip fractures is complex⁹ due to factors like cognitive impairment and co-morbidities and hence investigation is required to assess whether outcome instruments like HHS or mHHS which have been shown to be valid and responsive in cohort of patients with non-traumatic hip conditions continue to demonstrate validity and responsiveness in cohort of patients with proximal femur fracture. The responsiveness and floor ceiling effects of the Harris hip score has been described in intracapsular neck of femur fracture^{5,10}. However, though commonly used to assess functional outcome after intertrochanteric hip fractures, there is no report on validity, responsiveness and floor ceiling effect of either the Harris hip score or the modified Harris hip score after pertrochanteric fracture. Sitting in cross leg position and squatting on toes are essential activities for patients in the rural setting in the Indian sub-continent¹¹. The original Harris hip score does not contain ability to perform squatting and sitting cross legged as items in the questionnaire and hence a need was felt to modify the Harris hip score so that it could stay relevant in functional assessment of Indian patients in the rural setting.

The objective of the present study was to investigate construct validity, responsiveness and ceiling floor effects of the modified Harris hip score (mHHS) relevant to the Indian population with pertrochanteric fractures.

2. Methods

2.1. Ethics committee approval, inclusion and exclusion criteria

This prospective study was conducted in a university teaching hospital after obtaining approval from the Institutional research ethics committee. Patients with intertrochanteric fracture presenting to our hospital from 1st November 2014 to 1st May 2016 and treated with long or short proximal femoral nail were included in the study. Patients gave written informed consent to participate in the study.

2.2. Data collection

Data pertaining to the age, gender, type of implant (long or short proximal femoral nail), interval between injury and surgical intervention (in days), side affected, type of anaesthesia (general anaesthesia/spinal anaesthesia), co-morbidities, type of fracture (AO classification), pre-injury walking status, duration of operative procedure (in minutes) and amount of intra-operative blood loss (in millilitre) was collected and saved onto Microsoft Excel worksheet. Clinical and radiological follow-up was undertaken at one month, three months and six months after the surgical intervention.

2.3. Functional outcome assessment

At each follow-up at one, three and six months, functional outcome was assessed using the modified Harris hip score (Table 1). Two items pertaining to socks/shoes and sitting from the original Harris hip score were replaced with squatting and sitting cross legged respectively. This was done as patients in the Indian rural setting do not wear shoes and

Table 1
shows the domains and items of the modified Harris hip score.

Domains and items	Points
Pain	
None or ignores it	44
Slight, occasional, no compromise in activities	40
Mild pain, no effect on average activities, rarely moderate pain with unusual activity, may take diclofenac	30
Moderate pain, tolerable but makes concessions to pain, some limitation of ordinary activity and work: may require occasional pain medicine stronger than diclofenac	20
Marked pain, serious limitation of activities	10
Totally disabled, crippled, pain in bed, bedridden	0
Function: Gait.	
Limp	
None	11
Slight	8
Moderate	5
Severe or not able to walk	0
Support	
None	11
Cane for long walks	7
Cane most of the time	5
One crutch	3
Two canes	2
Two crutches or not able to walk	0
Distance walked	
Unlimited	11
Six blocks	8
Two or three blocks	5
Indoors only	2
Bed and chair	0
Functional activities	
Stairs	
Normally without using a rail	4
Normally using a rail	2
In any manner	1
Unable	0
Squatting	
With ease	4
With difficulty	2
Unable	0
Sitting cross legged	
With ease	5
With difficulty	3
Unable	0
Public transportation	
Able to use	1
Unable to use	0
Hip range of motion (Clinician assessed)	
Flexion (maximum = 140°)	
Abduction (maximum = 40°)	
Adduction (maximum = 40°)	
External rotation (maximum = 40°)	
Internal rotation (maximum = 40°)	
Range of motion scale (sum of the range of motion)	
211–300	5
161–210	4
101–160	3
61–100	2
31–60	1
0–30	0
Absence of deformity (Clinician assessed)	
• Less than 30° fixed flexion contracture – Yes/No	
• Less than 10° fixed abduction – Yes/No	
• Less than 10° fixed internal rotation in extension – Yes/No	
• Less than 3.2 cm limb length discrepancy – Yes/No	
If all 4 yes	4
If less than 4 yes	0

socks. The modified Harris hip score was scored from 0 (worst functional outcome and maximum pain) to 100 points (best functional outcome and least pain). Like the original Harris hip score, the interpretation of outcome using the modified Harris hip score was as

follows: < 70 (poor result), 70–79 (fair result), 80–89 (good result) and > 90 (excellent result)². Paired *t*-test was used to evaluate for statistically significant difference in the change of mHHS from one to three months postoperatively and from three to six months postoperatively. The level of significance was set at 5%.

2.4. Assessment of construct validity

The term ‘Construct validity’ refers to the extent to which scores of an outcome instrument are in concordance with hypothesis formulated regarding either the correlation of the outcome instrument with another previously validated outcome instrument or any difference in the values of the outcome instrument in different groups¹². A satisfactory or adequacy of construct validity shows that the outcome instrument is indeed measuring the construct that it is supposed to be evaluating.

The first hypothesis was that there would be no significant difference in the modified Harris hip score in patients that underwent fixation using short proximal femoral nail and long proximal femoral nail. This hypothesis was based on previous studies which showed no difference in the functional outcome after fixation using short and long proximal femoral nails^{13,14}. Independent sample *t*-test was used to evaluate statistical significance of difference of mean modified Harris hip score of long and short proximal femoral nail cohorts at one, three and six months postoperatively and the level of significance was set at 5%.

Outcome after hip fracture is better in patients younger than 65 years as compared to patients older than 65 years¹⁵. Studies on hip fractures have classified patients older than 65 years as a geriatric cohort¹⁶. Our second hypothesis was that the modified Harris hip score of patients in the age group of less than 65 years (non-geriatric cohort) would be significantly higher than the cohort of patients of age group more than or equal to 65 years (geriatric cohort). Independent sample *t*-test was used to evaluate statistical significance of difference of mean modified Harris hip score of non-geriatric and geriatric cohorts at one, three and six months postoperatively and the level of significance was set at 5%. Construct validity is considered to be adequate if the more than 75% of the pre-defined hypotheses are confirmed by the analysis⁶.

2.5. Assessment of internal validity

The number and proportion of patients achieving least possible score of 0 point (floor effect) and 100 points (ceiling effect) on the modified Harris hip score was evaluated at one, three and six months postoperatively. The absence of floor and ceiling effect are indicators of satisfactory internal validity of an outcome instrument¹⁷. The proportion of floor and ceiling effects must be lower than 15% for an outcome instrument to retain its validity⁶.

2.6. Assessment of responsiveness

Responsiveness was evaluated using distribution based methods (internal responsiveness) and anchor based methods (external responsiveness).

Internal responsiveness was estimated using the standardised response mean (SRM: mean of the change in modified Harris hip score divided by the standard deviation of the change in modified Harris hip score) and effect size (ES: mean of the change in modified Harris hip score divided by the standard deviation of the modified Harris hip score at initial assessment). ES and SRM are interpreted as per Cohen’s criteria wherein values of 0.2, 0.5 and 0.8 are considered as small, moderate and large effects respectively¹⁸. SRM values could be interpreted as values between 0.2 and 0.3 suggesting effects of placebo treatment, values between 0.3 and 0.5 suggesting effect of a moderately effective pharmaceutical agent and values more than 1.0 are usually seen after an effective surgical intervention¹⁸. ES and SRM value higher than 0.8 are indicators of satisfactory internal responsiveness¹⁹.

External responsiveness was determined by using two different external anchors (one anchor pertaining to ability to sit cross legged and the second anchor pertaining to the ability to squat). At each follow-up the patient response for ability to squat (1 – squat with ease/2 – squat with difficulty/3 – unable to squat) and ability to sit cross legged (1 – able to sit cross legged with ease/2 – able to sit cross legged with difficulty/3 – unable to sit cross legged) was recorded. Patients who were able to squat with ease and squat with difficulty were grouped as “Squatter cohort” and those unable to squat were termed as “Non-squatter cohort”. Similarly, those patients who were able to sit cross legged with ease and with difficulty were grouped as “Cross legged sitter cohort” and those unable to sit cross legged were grouped as “Non-cross leg sitter cohort”. Receiver operating curve (ROC) was plotted to check the ability of the modified Harris hip score to discriminate between patients who were able to squat and unable to squat and patients who were able to sit cross legged and unable to perform cross leg sitting. The change in mHHS from one month to six month postoperatively was used in the ROC. The area under the curve (AUC) was calculated from the ROC. The value of AUC ranges from 0 (no discriminatory ability) to 1 (perfect discriminatory ability). The value of AUC of 0.5 suggests that there is only 50% discriminatory ability of the outcome instrument thereby indicating that the discriminatory ability of the outcome instrument is no better than chance. AUC value more than 0.7 suggests satisfactory external responsiveness^{6,19}. The AUC value is considered significant if *p* value is less than 0.05 and if the lower limit of the 95% confidence interval of the AUC is higher than 0.5.

2.7. Statistical analysis

Quantitative data such as age, interval between injury and surgical intervention (in days), duration of operative procedure (in minutes) and amount of intra-operative blood loss (in millilitre), modified Harris hip score was presented as mean and standard deviation. Categorical data such as gender, type of implant (long or short proximal femoral nail), affected side, type of anaesthesia (general anaesthesia/spinal anaesthesia), co-morbidities, type of fracture (AO classification) and pre-injury walking status was presented as proportion and percentage. SPSS version 20 was used for statistical analysis.

3. Results

This prospective cohort study consisted of 81 consecutive patients who underwent proximal femoral nail fixation for pertrochanteric fractures of the hip from 1st November 2014 to 1st May 2016. Six patients were lost to follow up due to mortality and functional outcome data at one, three and six months after the surgical intervention was available for 75 patients.

3.1. Baseline demographic characteristics

The demographic features of the study population are given in Table 2.

3.2. Modified Harris hip score (mHHS)

The mean mHHS at one month, three months and six months postoperatively was 39.9 ± 9.5 , 61.6 ± 14.7 and 81 ± 15.9 respectively. Paired *t*-test showed that the improvement in mHHS from one to three months postoperatively was significant ($p < 0.0001$; 95% CI: 19.5 to 23.8) and also the improvement from three to six months postoperatively was significant ($p < 0.0001$; 95% CI: 17.2 to 21.6). At six months postoperatively, excellent, good, fair and poor outcome was seen in 23 patients (30.7%), 25 patients (33.3%), 19 patients (25.3%) and 8 patients (10.7%) respectively. Fig. 1 shows the trend in the improvement of mHHS postoperatively.

Table 2
shows demographic features of the cohort.

Demographic feature	*expressed as mean (standard deviation) {range}
N = 81	Rest expressed as count (percentage)
Age (years)*	68 (11.8) {38–89}
Gender	
Male	41 (50.6 %)
Female	40 (49.4 %)
Side affected	
Right hip	38 (46.9%)
Left hip	43 (53.1%)
AO classification	
A1.2	6 (7.4 %)
A1.3	3 (3.7 %)
A2.1	35 (43.2 %)
A2.2	18 (22.2 %)
A2.3	14 (17.3 %)
A3.1	4 (4.9 %)
A3.3	1 (1.2%)
Co-morbidities	
Hypertension	42 (51.9 %)
Diabetes Mellitus	12 (14.8%)
Ischemia heart disease	7 (8.6 %)
Chronic obstructive pulmonary disease	8 (9.9 %)
Previous cerebrovascular stroke (hemiparesis)	4 (4.9 %)
Pre-injury walking status	
Walking without aid	38 (46.9 %)
Walking with support	40 (49.4 %)
Bedridden	3 (3.7 %)
Type of anaesthesia	
Spinal anaesthesia	76 (93.8%)
General anaesthesia	5 (6.2%)
Type of implant	
Long proximal femoral nail	39 (48.1%)
Short proximal femoral nail	42 (51.9 %)
Injury surgery interval (days)*	3 (1.5) {1–10}
Duration of operative procedure (minutes)*	84 (23.3) {50–125}
Intra-operative blood loss (ml)*	262.4 (65.5) {150–500}

Fig. 1. shows statistically significant improvement in modified Harris hip score (mHHS) at all time intervals as there is no overlapping of their 95% confidence intervals. The interpretation of outcome using the modified Harris hip score was as follows: < 70 (poor result), 70–79 (fair result), 80–89 (good result) and > 90 (excellent result).

3.3. Construct validity

In the present study, 39 patients underwent fixation using the short proximal femoral nail (short PFN) and 36 patients underwent fixation

using the long proximal femoral nail (long PFN). In accordance with our first hypothesis, there was no significant difference in the mHHS obtained by patients fixed with either short or long proximal femoral nail. No significant difference in mHHS was observed at one (Short PFN: 38.3 ± 10.3 ; Long PFN: 41.7 ± 8.2 $p = 0.13$; 95% CI: -7.7 to 1.0), three (Short PFN: 61.1 ± 15.6 ; Long PFN: 62.1 ± 13.9 ; $p = 0.8$; 95%

Table 3
shows evaluation of internal responsiveness using distribution based methods.

	1 month postoperative	3 months postoperative	6 months postoperative
n	75	75	75
Mean modified Harris hip score (SD)	39.9 (9.46)	61.6 (14.70)	81.0 (15.87)
Mean of change in modified Harris hip score (SD of change score)	Not applicable	21.7 (9.37) ^a	19.4 (9.65) ^b
Effect size (ES)	Not applicable	2.29 ^c	1.32 ^d
Standardised response mean (SRM)	Not applicable	2.32 ^c	2.01 ^d

^a Change score calculated using 1 month postoperative value as baseline.

^b Change score calculated using 3 month postoperative value as baseline.

^c ES and SRM calculated using 1 month postoperative values as baseline.

^d ES and SRM calculated using 3 months postoperative values as baseline.

Fig. 2. shows the satisfactory ability of the modified Harris hip score (mHHS) to discriminate patients who were able to sit cross legged at six months postoperatively from patients who were unable to sit cross legged.

CI: -7.8 to 5.8) and six months (Short PFN: 79.8 ± 17.5 ; Long PFN: 82.3 ± 14.0 ; $p = 0.5$; 95% CI: -9.7 to 4.8) postoperatively suggesting that patient with both fixation devices achieve identical functional outcome.

There were 29 patients who were younger than 65 years and 46 patients who were aged more than or equal to 65 years. In accordance with our second hypothesis, patients younger than 65 years had significantly better mHHS at one (Age < 65 years: 44.1 ± 8.1 ; Age ≥ 65 years: 37.3 ± 9.4 ; $p = 0.002$; 95% CI: 2.6 to 11.0), three (Age < 65 years: 68.7 ± 10.9 ; Age ≥ 65 years: 57.1 ± 15.1 ; $p < 0.0001$; 95% CI: 5.5 to 17.6) and six months (Age < 65 years: 88.9 ± 14.6 ; Age ≥ 65 years: 76.0 ± 14.6 ; $p < 0.0001$; 95% CI: 6.1 to 19.9) postoperatively compared to patients older than 65 years.

3.4. Internal validity

The range of mHHS at one month, three and six months postoperatively was 7–53, 6–78 and 17–99 respectively. No floor or ceiling

Fig. 3. shows the satisfactory discriminatory ability of modified Harris hip score (mHHS) based on patient's ability to squat.

effect was observed at one, three and six months postoperatively. But six patients at six months follow-up secured 99 points (8%) suggesting closeness to ceiling effect but this was lower than the threshold value of 15%.

3.5. Responsiveness

The overall Effect Size (ES) and Standardised Response Mean (SRM) at the two extreme end points namely one month and six months postoperative was 4.34 and 4.26 respectively. The ES and SRM at adjacent postoperative follow-up duration from one month to three months and from three months to six months is depicted in Table 3. The ES and SRM at all time durations were higher than the recommended threshold value of 0.8.

At six months follow-up, nine patients were able to sit cross legged with ease (12%), 13 patients were able to perform cross leg sitting with difficulty (17.3%) and the remaining 53 patients were unable to perform cross leg sitting (70.7%). For calculation of the area under the curve, 22 patients (29.3%) were considered having ability to perform cross leg sitting and 53 patients (70.7%) were considered non-cross leg sitters. The area under the curve (AUC) using ability to perform cross legged sitting as external anchor was 0.89 (95% confidence interval: 0.80 to 0.99; $p < 0.0001$) and this was significant (Fig. 2).

At six months follow-up, 15 patients (20%) were able to squat easily; seven patients (9.3%) were able to squat with difficulty and 53 patients (70.7%) were unable to squat. For calculation of the area under the curve, 22 patients (29.3%) were labelled as having ability to squat and 53 patients (70.7%) were labelled as having inability to squat. The area under the curve (AUC) using ability to perform squatting as external anchor was 0.77 (95% confidence interval: 0.65 to 0.89; $p < 0.0001$) and the result was significant (Fig. 3).

4. Discussion

Rai et al.,¹¹ modified the HHS to suit patients from the Indian sub-continent by including sitting in cross leg position and squatting on toes in their questionnaire. However; in their questionnaire there were significant modifications from the HHS like exclusion of function sub-domains (distance walked and limp), functional activities sub-domains (stair climbing and use of public transportation), deformity assessment, inclusion of new domains (patient satisfaction, trendelenberg sign assessment). In the original HHS and in our modification of HHS there was no negative marking but in the HHS modified by Rai et al. there was negative marking of 15 points if trendelenberg sign was positive. In the questionnaire by Rai et al., functional activities like squatting and ability to sit in cross leg position were evaluated as binary response (can do/unable to do) whereas in our questionnaire these activities were evaluated using three responses (with ease/with difficulty and unable to do). Having multiple options helps to capture true change in functional outcome more appropriately. The HHS modified by Rai et al. had 25 points for pain domain and only four responses as compared to 44 points for six responses in the original HHS and in our study. Rai et al. quantified pain as occasional, moderate and severe and there was no clear description of how these could be assessed objectively. Rai et al. evaluated hip range of motion using adjectival criteria (normal, moderate reduction and severe reduction) and in the absence of clear cut objective criteria of what would constitute moderate reduction and severe reduction could lead to lower interobserver and intraobserver agreement. The HHS modified by Rai et al. had significant changes from the original HHS not only in terms of domains but also in terms of points allotted to various domains. Our modification of HHS was pertaining to only two items wherein items pertaining to ability to wear shoes/socks and sitting were replaced by ability to squat on toes and sitting in cross leg position. We did not make any alteration to the marking scheme of individual sub-domains and the overall domains. The HHS modified by Rai et al. was evaluated in a cohort of patients

with intracapsular fracture of neck of femur treated with bipolar hemiarthroplasty and there was no report of psychometric properties like validity, reliability and responsiveness of their modified HHS. The proportion of patients achieving excellent, good, fair and poor results as per our mHHS at six months postoperatively was comparable to the results observed in a previous study² that used HHS.

4.1. Construct validity

Known groups method, convergent and divergent validity evaluation are methods to assess construct validity of an outcome instrument. We used the known groups method and all our results were in accordance with the two clinical hypothesis that were prior formulated.

Studies have shown that HHS²⁰ and mHHS²¹ tend to decrease with age especially in the geriatric population that sustained intertrochanteric hip fracture and treated with proximal femoral nail^{20,21}. Our results are in accordance with results from previously published studies^{13,14} that used the original HHS for assessment of functional outcome after pertrochanteric fractures after long and short proximal femoral nail fixation found no difference in outcome at one year postoperative follow-up.

4.2. Floor and ceiling effects (Internal validity)

Floor and ceiling effects are indicators of content validity of an outcome instrument²². Frihagen et al.,⁵ used the original HHS to assess functional outcome after neck of femur fracture at four and 12 months postoperatively and reported absence of any floor or ceiling effect. A systematic review on studies reporting the HHS highlighted that the pooled ceiling effect could be as high as 20%²². However, the mean duration of follow-up in the systematic review was extremely long [5.9 years] (range: 0.25 years to 15.2 years postoperatively) and the primary intervention procedure was primary total hip arthroplasty for arthritis and fracture. It is expected that if normally functioning individuals gain maximal score of an outcome instrument (ceiling effect) and in long term follow-up studies it is reasonable to observe that after a successful intervention most participants of the study should achieve the maximal score. Just like the original HHS, our mHHS did not demonstrate ceiling or floor effect at one, three and six months follow-up in hip fracture.

4.3. Responsiveness

Though it remains debatable as to what constitutes the most ideal method to evaluate responsiveness of an outcome instrument, effect size (ES) and standardised response mean (SRM) are the most commonly reported methods in literature¹⁰. Area under the curve assessment (AUC) using receiver operating curve is the most common anchor based method described to evaluate external responsiveness of an outcome instrument. Honkavaara et al.,¹⁰ reported responsiveness of the original HHS in cohort of young patients (mean age – 57 years) with intracapsular neck of femur fracture and treated with closed reduction and internal fixation. ES and SRM were not reported for HHS and the AUC for HHS was 0.53 thereby, suggesting poor responsiveness of the HHS. Possible explanation for this difference could be the use of different external anchors and the difference in duration of assessment. Hookavaara et al. used the current health status as external anchor whereas we used ability to squat and sit in cross leg position as external anchors. Also, Hookavaara et al. assessed AUC using change in HHS from four to 24 months to determine responsiveness whereas we used change in mHHS from one to six months to determine external responsiveness. HHS tends to improve in the first six months, tends to plateau between six to nine months and then starts deteriorating gradually even below the values achieved at one month postoperatively. Long term functional outcome studies on HHS in pertrochanteric fractures have shown that HHS value at 36 months postoperatively was

lower than the HHS at one month postoperatively⁴. Frihagen et al.,⁵ described the responsiveness of HHS in elderly patients (mean age – 83 years) with intracapsular neck of femur fracture and treated with either internal fixation or hemiarthroplasty. The study reported moderate SRM (0.75) and AUC of 0.76 which was comparable to our result. ES was not evaluated in that study. One possible explanation for SRM being low than that observed in our study could be that Frihagen et al. utilized the change in HHS from four to 12 months to determine SRM whereas we utilized the change in mHHS from one to six months to evaluate SRM. Honkavaara et al.,¹⁰ showed that the value of SRM for EQ5D and SF-36 tends to decrease at long term follow-up at 12 months and 24 months. Hence, in short term follow-up though score obtained tends to improve, in the long term it shows a trend to plateau and then deteriorate. We chose ability to sit cross legged and ability to squat as external anchors as they are extremely relevant to patients in the Indian subcontinent. There is no “gold standard” method described to choose an external anchor. It would be interesting to see whether the external responsiveness of mHHS would change with use of other external anchors like change in the clinical condition of the patient or patient satisfaction after surgical intervention.

A recent systematic review on quality of life after hip fractures concluded that the health related quality of life improved in the first six months after the hip fracture²³. Studies have described the assessment of functional outcome at six months after fixation of pertrochanteric hip fractures^{2,3}. Most studies on evaluation of psychometric properties of outcome instruments after hip fracture have assessed outcome at one month, three months, six months and twelve months after surgical intervention⁹. There is no specific recommendation on what would constitute ideal time for assessment of psychometric properties.

Strengths of our study include use of both distribution based and anchor based methods to determine internal and external responsiveness. The internal validity of our study was enhanced by lower chance of assessment bias as the evaluation was done by an orthopaedic resident who was not the operating surgeon and lower chance of loss to follow-up bias because the loss to follow-up due to mortality was 7.4%. All patients who were alive participated in the study at one, three and six months. Long-term follow-up in patients with proximal femoral fracture is challenging due to loss to follow-up bias due to high mortality²⁴. If the incidence of loss to follow-up is less than 5% then there is low risk of loss to follow-up bias however; loss of more than 20% patients, leads to higher incidence of loss to follow-up bias and it endangers the internal validity of the study²⁵. Ours is the first study to describe the validity and responsiveness of the HHS modified for Indian population.

The external validity of the study is limited only to the cohort of patients with pertrochanteric hip fracture treated with proximal femoral nail and generalisable only to patients from Indian sub-continent and in countries where squatting and sitting cross legged is part of daily activities.

5. Conclusion

The present study has established that the modified Harris hip score that includes items pertaining to squatting and sitting cross legged has demonstrated satisfactory construct validity, internal validity and responsiveness in cohort of patients with pertrochanteric fractures and treated with proximal femoral nail.

Author's contribution

1. Karthik Vishwanathan – study design, study set up, data collection, data compiling, statistical analysis, manuscript writing and manuscript revision
2. Keyur Akbari – study design, study set up, data collection, manuscript writing and revision
3. Amit Patel – study design, study set up, data review, data analysis, manuscript writing and revision

Conflict of interest

None.

References

- Harris WH. Traumatic arthritis of the hip after dislocation and acetabular fractures: treatment by mold arthroplasty. *J Bone Joint Surg Am.* 1969;51:737–755.
- Babhulkar S. Unstable trochanteric fractures: issues and avoiding pitfalls. *Injury.* 2017;48(4):803–818. <http://dx.doi.org/10.1016/j.injury.2017.02.022>.
- Hopp S, Wirbel R, Ojodu I, Pizanis A, Pohlemann T, Fleischer J. Does the implant make the difference? – Prospective comparison of two different proximal femur nails. *Acta Orthop Belg.* 2016;82(2):319–331.
- Zhang H, Zeng X, Zhang N, Zeng D, Xu P, Zhang L, et al. INTERTAN nail versus proximal femoral nail antirotation-Asia for intertrochanteric femur fractures in elderly patients with primary osteoporosis. *J Int Med Res.* 2017;45(4):1297–1309. <http://dx.doi.org/10.1177/0300060517710584>.
- Frihagen F, Grotle M, Madsen JE, Wyller TB, Mowinckel P, Nordsletten L. Outcome after femoral neck fractures: a comparison of Harris Hip Score, Eq-5d and Barthel Index. *Injury.* 2008;39(10):1147–1156. <http://dx.doi.org/10.1016/j.injury.2008.03.027>.
- Terwee CB, Bot SDM, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60:34–42.
- Irrgang JJ, Lubowitz JH. Measuring arthroscopic outcome. *Arthroscopy.* 2008;24:718–722.
- Söderman P, Malchau H. Is the Harris hip score system useful to study the outcome of total hip replacement? *Clin Orthop Relat Res.* 2001;384:189–197.
- Hutchings L, Fox R, Chesser T. Proximal femoral fractures in the elderly: how are we measuring outcome? *Injury.* 2011;42(11):1205–1213. <http://dx.doi.org/10.1016/j.injury.2010.12.016>.
- Honkavaara N, Al-Ani AN, Campenfeldt P, Ekström W, Hedström M. Good responsiveness with EuroQol 5-dimension questionnaire and short form (36) health survey in 20–69 years old patients with a femoral neck fracture: a 2-year prospective follow-up study in 182 patients. *Injury.* 2016;47(8):1692–1697. <http://dx.doi.org/10.1016/j.injury.2016.05.021>.
- Rai AK, Agarwal R, Singh S, Ratan R. The BHU bicentric bipolar prosthesis in fracture neck femur in active elderly. *J Trauma Manag Outcomes.* 2008;2(1):7. <http://dx.doi.org/10.1186/1752-2897-2-7>.
- Scholtes VA, Terwee CB, Poolman RW. What makes a measurement instrument valid and reliable? *Injury.* 2011;42(3):236–240. <http://dx.doi.org/10.1016/j.injury.2010.11.042>.
- Li Z, Liu Y, Liang Y, Zhao C, Zhang Y. Short versus long intramedullary nails for the treatment of intertrochanteric hip fractures in patients older than 65 years. *Int J Clin Exp Med.* 2015;8:6299–6302.
- Okcu G, Ozkayin N, Okta C, Topcu I, Aktuglu K. Which implant is better for treating reverse obliquity fractures of the proximal femur: a standard or long nail? *Clin Orthop Relat Res.* 2013;471(9):2768–2775. <http://dx.doi.org/10.1007/s11999-013-2948-0>.
- Holt G, Smith R, Duncan K, Hutchison JD, Gregori A. Epidemiology and outcome after hip fracture in the under 65s-evidence from the Scottish Hip Fracture Audit. *Injury.* 2008;39(10):1175–1181. <http://dx.doi.org/10.1016/j.injury.2008.04.015>.
- Suhm N, Kaelin R, Studer P, et al. Orthogeriatric care pathway: a prospective survey of impact on length of stay, mortality and institutionalisation. *Arch Orthop Trauma Surg.* 2014;134(9):1261–1269. <http://dx.doi.org/10.1007/s00402-014-2057-x>.
- Sciascia AD, Morris BJ, Jacobs CA, Edwards TB. Responsiveness and internal validity of common patient-reported outcome measures following total shoulder arthroplasty. *Orthopedics.* 2017;40(3):e513–e519. <http://dx.doi.org/10.3928/01477447-20170327-02>.
- Schouffoer AA, van der Giesen FJ, Beaat-van de Voorde LJ, Wolterbeek R, Huizinga TW, Vliet Vlieland TP. Validity and responsiveness of the Michigan Hand Questionnaire in patients with systemic sclerosis. *Rheumatology (Oxford).* 2016;55(8):1386–1393. <http://dx.doi.org/10.1093/rheumatology/kew016>.
- Naal FD, Impellizzeri FM, Rippstein PF. Which are the most frequently used outcome instruments in studies on total ankle arthroplasty. *Clin Orthop Relat Res.* 2010;468:815–826.
- Tang P, Hu F, Shen J, Zhang L, Zhang L. Proximal femoral nail antirotation versus hemiarthroplasty: a study for the treatment of intertrochanteric fractures. *Injury.* 2012;43(6):876–881. <http://dx.doi.org/10.1016/j.injury.2011.11.008>.
- Ruecker AH, Rupprecht M, Gruber M, et al. The treatment of intertrochanteric fractures: results using an intramedullary nail with integrated cephalocervical screws and linear compression. *J Orthop Trauma.* 2009;23(1):22–30. <http://dx.doi.org/10.1097/BOT.0b013e31819211b2>.
- Wamper KE, Siersevelt IN, Poolman RW, Bhandari M, Haverkamp D. The Harris hip score: do ceiling effects limit its usefulness in orthopedics? *Acta Orthop.* 2010;81(6):703–707. <http://dx.doi.org/10.3109/17453674.2010.537808>.
- Peeters CM, Visser E, Van de Ree CL, Gossens T, Den Ouden BL, De Vries J. Quality of life after hip fracture in the elderly: a systematic literature review. *Injury.* 2016;47(7):1369–1382. <http://dx.doi.org/10.1016/j.injury.2016.04.018>.
- Nazrun AS, Tzar MN, Mokhtar SA, Mohamed IN. A systematic review of the outcomes of osteoporotic fracture patients after hospital discharge: morbidity, subsequent fractures, and mortality. *Ther Clin Risk Manag.* 2014;10:937–948. <http://dx.doi.org/10.2147/TCRM.S72456>.
- Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. *Lancet.* 2002;359(9308):781–785.