

## ORIGINAL ARTICLE

# Predicting Treatment Success after Scarf Osteotomy for Hallux Valgus using The American Orthopedic Foot and Ankle Society and Short Form Health Survey Scores

<sup>1</sup>Meng Zhu, <sup>2</sup>Jerry Y Chen, <sup>3</sup>Terry HL Teo, <sup>4</sup>Kevin OT Koo, <sup>5</sup>Inderjeet S Rikhraj

## ABSTRACT

**Aim:** Various hallux valgus corrective surgeries, including Scarf osteotomy, have demonstrated improvement in patients' fore-foot function and quality of life. However, no threshold values of these measures have been reported to define the success of surgery from the patients' perspective. This study aims to define treatment success threshold values of The American Orthopedic Foot and Ankle Society (AOFAS)-Health Management Information System (HMIS) and Short Form (36) Health Survey (SF-36) Scores for patients underwent Scarf osteotomy (SO) for hallux valgus.

**Materials and methods:** A retrospective analysis of patients who underwent scarf osteotomy between 2007 and 2013 was conducted. Patients were evaluated for AOFAS-HMIS, SF-36 score, satisfaction, and pain score at various time points. Treatment success was defined as a significant improvement in pain and satisfaction with surgery. Using receiver operating characteristic (ROC) analysis, threshold score for treatment success was defined as the cut-off value providing the largest sum of sensitivity and specificity.

**Results:** A total of 360 and 345 patients completed assessments at 6-month and 2-year after surgery, with a success rate of 70.0% and 77.4%, respectively. The AOFAS-HMIS and physical component score (PCS) improved significantly at both postoperative time points as compared to preoperatively. The ROC analyses revealed the excellent predictive value of AOFAS-HMIS for treatment success (AUC = 0.840 and 0.835).

**Conclusion:** The American Orthopedic Foot and Ankle Society-Health Management Information System of 84 and above at 6-month, or 89 and above at 2-year after surgery can be used to define treatment success of OS for hallux valgus. An increasingly higher threshold is required to define treatment success with time.

**Keywords:** Hallux valgus, Patient-reported outcome, Scarf osteotomy, The American Orthopedic Foot and Ankle Society-Health Management Information System, Threshold score, Treatment success.

**How to cite this article:** Zhu M, Chen JY, Teo THL, Koo KOT, Rikhraj IS. Predicting Treatment Success after Scarf Osteotomy for Hallux Valgus using The American Orthopedic Foot and

Ankle Society and Short Form Health Survey Scores. *J Foot Ankle Surg (Asia Pacific)* 2018;5(2):59-63.

**Source of support:** Nil

**Conflict of interest:** None

## INTRODUCTION

Hallux valgus is a common chronic forefoot condition that affects the function of foot and quality of life (QOL).<sup>1,2</sup> Studies have demonstrated how various types of surgeries, including Scarf osteotomy, can improve the function and QOL through a combination of patient-reported outcome measures (PROMs) and different clinical parameters.<sup>3-8</sup> While surgeons and healthcare professionals can monitor the changes in these scores for the progression of the condition and treatment response, there exists no threshold value to define treatment success in the patients' perspective.

The AOFAS score is the most frequently used measurement tool for assessing the function of foot and ankle in research,<sup>9,10</sup> and the AOFAS-HMIS has been shown to be reliable, valid, and responsive in clinical use.<sup>11-14</sup> The Short Form 36-Item Health Survey (SF-36, RAND Health, Santa Monica, California) is a widely-employed health-related QoL evaluation method.<sup>15</sup> Both AOFAS-HMIS and SF-36 score are the primary outcome measures used in the authors' institution, similar in many others, for observing treatment response. Therefore, the objective of this study is to define the threshold scores of AOFAS-HMIS and SF-36 for successful treatment of hallux valgus using scarf osteotomy.

## MATERIALS AND METHODS

This study was approved by the hospital's ethics committee and carried out in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki. Written informed consent was obtained from all the patients.

Patients who underwent scarf osteotomy for correcting hallux valgus deformity at the foot and ankle division of a tertiary hospital during 2007 and 2013 were included. Relief of pain and correction of deformity were the indications for surgery. The patients were followed-up for 2

<sup>1</sup>Resident, <sup>2,3</sup>Associate Consultant, <sup>4</sup>Consultant, <sup>5</sup>Senior Consultant

<sup>1-5</sup>Department of Orthopedic Surgery, Singapore General Hospital, Singapore

**Corresponding Author:** Meng Zhu, Resident, Department of Orthopedic Surgery, Singapore General Hospital, Singapore, e-mail: meng.zhu@mohh.com.sg

years. They were assessed by independent healthcare professionals preoperatively, at 6-month and 2-year-postoperatively, for forefoot function, QoL, and patient satisfaction. The AOFAS-HMIS was used to evaluate the forefoot function, which combines both PROMs and objective clinical evaluations. This scale assigned 40, 45, and 15 points to pain, function, and alignment respectively to give a maximally possible score of 100 points. The function component assessed includes daily and recreational activity limitation, footwear requirement, the presence of callus, as well as metatarsophalangeal and interphalangeal joint motion and stability.

Patient QOL was assessed by SF-36, which consists of eight subscales: physical functioning, physical role, bodily pain, general health, vitality, social functioning, emotional role, and mental health. The individual scores were compounded into two higher-order summary scores, the PCS and mental component score (MCS), based on the most highly correlated subscales to simplify analyses without substantial loss of information.<sup>16</sup> These two summary scores have good validity in discriminating among different clinical groups, along with high test-retest reliability when used in a general population.<sup>17,18</sup>

During the follow-up assessment, the patients were assessed whether they are satisfied with the surgery. Besides, the pain was also rated on a 0 to 10 numerical scale, with 0 being no pain and 10 being the worst possible pain.<sup>19</sup> Treatment success was defined as satisfaction with surgery, as well as pain-free or significant improvement of pain at the follow-up. Significant improvement in pain was defined as a reduction of pain score of either no less than two points or at least 30% of baseline score, whichever is higher.<sup>20</sup>

**Statistical Analysis**

Statistical analyses were carried out using ROC analyses in Statistical Package for the Social Sciences (SPSS) 21.0 (IBM, Armonk, NY, USA). The outcome measures, namely AOFAS-HMIS, PCS, MCS, and their respective changes at the two postoperative time points, were tested as predictors, while the defined treatment success was set as the criterion. An area under curve (AUC) between 0.70 and 0.80 indicates acceptable discrimination and an AUC above 0.80 indicates excellent discrimination.<sup>21</sup> The

**Table 1:** Patient demographics and clinical outcome scores

	Pre-op (N = 364)	6-Month (N = 360)	2-Year (N = 345)
Gender, male/female	25/339	25/335	23/322
Side, left/right	191/173	190/170	182/163
AOFAS-HMIS	58 ± 17	81 ± 16	87 ± 13
PCS	46 ± 4	49 ± 10	50 ± 8
MCS	55 ± 10	55 ± 10	55 ± 11

threshold score of each outcome measure for predicting treatment success was defined as the cut-off value that provides the largest sum of sensitivity and specificity.<sup>22</sup> Analyses were performed separately for both 6-month and 2-year follow-up for comparison.

**RESULTS**

Out of 364 cases included, a total of 360 (99%) and 345 (95%) cases were included and evaluated at 6-month and 2-year follow-up, correspondingly. Table 1 shows the demographics, preoperative PROMs as well as PROMs at the two follow-up time points. Significant improvements were noted in the AOFAS-HMIS (81 vs. 58, p < 0.001; and 87 vs. 58, p < 0.001) and PCS (49 vs. 46, p < 0.001; and 50 vs. 46, p < 0.001) at 6-month and 2-year follow-up.

Overall, 78.1% of patients had successful hallux valgus surgery at 6-month and 79.1% of patients at 2-year. Details on the fulfillment of the treatment success criteria are presented in Table 2. The ROC analyses demonstrated excellent prediction for treatment success using AOFAS-HMIS at both follow-up time points (AUC = 0.840 and 0.835), as shown in Figure 1. Acceptable AUC values were also observed if changes in the scale were used as predictors (AUC = 0.716 and 0.708). However, PCS, MCS, or their respective changes were not sensitive enough to predict treatment success (ROC curves for using PCS and MCS as predictors were not shown). Table 2 also summarizes the AUC values with 95% confidence range for each outcome measure used as predictors.

An AOFAS-HMIS of 84 and above at 6-month predicts successful treatment of hallux valgus with a sensitivity of 77% and a specificity of 77%, and a score of 89 and above at 2-year has a sensitivity of 73% and a specificity of 82%.

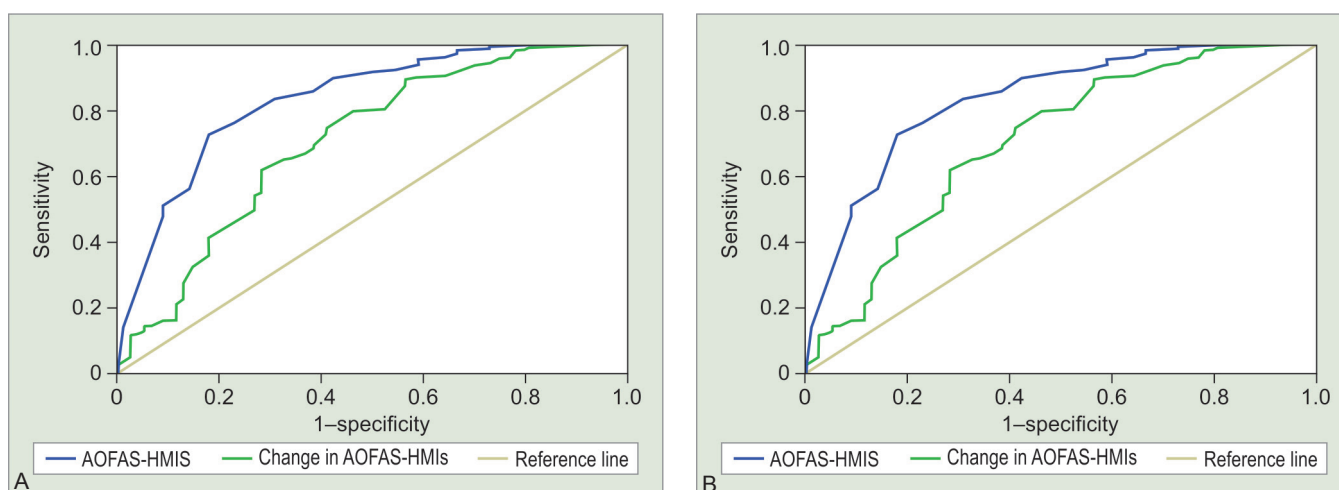
**DISCUSSION**

The most significant findings of the study are that AOFAS-HMIS is an excellent predictor of treatment

**Table 2:** Results for treatment success and area under curve (AUC) of Receiver Operating Characteristic (ROC) analysis.

	6-month (N = 360)	2-year (N = 345)
<i>Treatment success</i>		
Satisfaction criteria	281 (78.1%)	273 (79.1%)
Pain criteria	304 (84.4%)	318 (92.2%)
Treatment success	252 (70.0%)	267 (77.4%)
<i>AUC</i>		
AOFAS hallux MTP-IP scale	0.840 (0.795–0.885)	0.835 (0.782–0.888)
Change in AOFAS-HMIS	0.716 (0.654–0.738)	0.708 (0.638–0.779)
PCS	0.587 (0.518–0.657)	0.643 (0.573–0.714)
Change in PCS	0.605 (0.537–0.674)	0.574 (0.501–0.646)
MCS	0.679 (0.620–0.738)	0.617 (0.544–0.690)
Change in MCS	0.578 (0.512–0.643)	0.627 (0.553–0.702)





**Figs 1A and B:** Receiver operating characteristic (ROC) curve using AOFAS Hallux MTP-IP scale and its changes as predictors and treatment success as the criterion

success in hallux valgus correction using scarf osteotomy, and it can be used to interpret treatment outcomes both individually and in a cohort.

Hallux valgus is a common and complex deformity involving the first ray. Various clinical scores have been developed and validated to assess QOL and forefoot-specific function to guide management.<sup>12,13,23-26</sup> They give the convenience for surgeons to observe the score changes indetermining treatment response. However, treatment success cannot be accurately derived from the quantitative changes alone. Firstly, the change may be statistically meaningful, but clinically meaningless. Thus, minimal clinically important difference (MCID) has been proposed to be the smallest change in outcome scores which patients perceive as beneficial and meaningful.<sup>27,28</sup> While achieving the MCID is crucial for defining clinically meaningful improvement, it may not be considered as treatment success if the patient expectation is high. Secondly, the baseline score may affect the significance of the same magnitude of score change, with high baseline score requiring a possibly smaller magnitude of change to be construed as treatment success. Therefore, both absolute scores and their changes are considered as potential predictors in this study.

There is sparsity in the literature regarding the definition of treatment success in the context of hallux valgus corrective surgeries. Function scores and radiographic measurements are most commonly used in analyses and comparisons of surgical outcomes of hallux valgus correct surgeries. No definitive threshold values are provided in either absolute score or change in score to define treatment success. In addition, Schneider and Knahr found that using foot-specific clinical scores only partially addressed the important factors influencing the outcome of surgery, such as correction of footwear problems, alleviation of pain, and restoration of adequate walking.<sup>29</sup>

Criteria including pain and patient satisfaction are used in this study to define treatment success. This approach is more patient-centered and avoids physiological variables measured clinically or radiographically which may not be crucial in the patients' perspective. Only patients underwent scarf osteotomy procedures are included in the study to minimize heterogeneity.

Hunt and Hurwit reported that a total of 139 unique clinical outcome scale being used in the literature on foot and ankle topics during the period of 2002 to 2011, and the AOFAS scales were the most commonly used PROMs for evaluating foot and ankle conditions.<sup>9</sup> As AOFAS scales continue to gain popularity, the definitions of treatment success threshold scores for various foot and ankle procedures become more imminent. No similar definitions of treatment success threshold scores have been reported for PROMs in hallux valgus corrective surgeries. This study, the first of its kind, defines the threshold value of AOFAS-HMIS for treatment success of hallux valgus correction using scarf osteotomy.

Patient QOL is now recognized as one of the most important outcomes of surgery. Although there is a significant improvement in PCS postoperatively, the postoperative physical health status and improvement in physical health status were not meaningfully responsive for predicting treatment success. Pain relief is a significant contributor to the improvement of physical health QOL in hallux valgus,<sup>30</sup> correlating to treatment success. Other components in PCS, such as role physical and general health, may contribute insignificantly to treatment success, diluting the usefulness of the summary scores. The limitation of SF-36 in predicting postoperative improvement was also elucidated by others.<sup>31</sup>

There are a few limitations in this study. Firstly, there is a lack of radiographic assessment in determining treatment success. Studies have shown that the degree of



deformity and the amount of correction did not influence the outcome.<sup>1,2,32</sup> Furthermore, patients may have considered the cosmetic outcomes in the reported satisfaction score subjectively. Secondly, different study population characteristics (age, gender, and follow-up duration) and different surgical procedures may give different threshold scores. The relatively large sample size could partially reduce the bias; however, the reported values should be interpreted as a guide rather than definition.

In conclusion, AOFAS-HMIS has excellent predictive value in defining treatment success in hallux valgus correction using SO, facilitating interpretation of treatment outcomes. The threshold score for treatment success varies with time, requiring a higher score to constitute treatment success at the later time point.

## REFERENCES

- Saro C, Jensen I, Lindgren U, Fellander-Tsai L. Quality-of-life outcome after hallux valgus surgery. *Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation*. 2007 Jun;16(5):731-738.
- Thordarson D, Ebramzadeh E, Moorthy M, Lee J, Rudicel S. Correlation of hallux valgus surgical outcome with AOFAS forefoot score and radiological parameters. *Foot & ankle international*. 2005 Feb;26(2):122-127.
- Johnson KA, Cofield RH, Morrey BF. Chevron osteotomy for hallux valgus. *Clinical orthopaedics and related research*. 1979 Jul-Aug(142):44-47.
- Mitchell CL, Fleming JL, Allen R, Glenney C, Sanford GA. Osteotomy-bunionectomy for hallux valgus. *The Journal of bone and joint surgery American volume*. 1958 Jan;40-a(1):41-58.
- McBride ED. The McBride bunion hallux valgus operation. *The Journal of bone and joint surgery American volume*. 1967 Dec;49(8):1675-83.
- Lapidus PW. The author's bunion operation from 1931 to 1959. *Clinical orthopaedics*. 1960;16:119-135.
- Mann RA, Rudicel S, Graves SC. Repair of hallux valgus with a distal soft-tissue procedure and proximal metatarsal osteotomy. A long-term follow-up. *The Journal of bone and joint surgery American volume*. 1992 Jan;74(1):124-129.
- Weil LS. OSfor correction of hallux valgus. Historical perspective, surgical technique, and results. *Foot and ankle clinics*. 2000 Sep;5(3):559-580.
- Hunt KJ, Hurwit D. Use of patient-reported outcome measures in foot and ankle research. *The Journal of bone and joint surgery American volume*. 2013 Aug 21;95(16):e118(1-9).
- Kitaoka HB, Alexander IJ, Adelaar RS, Nunley JA, Myerson MS, Sanders M. Clinical rating systems for the ankle-hind-foot, midfoot, hallux, and lesser toes. *Foot & ankle international*. 1994 Jul;15(7):349-353.
- Ibrahim T, Beiri A, Azzabi M, Best AJ, Taylor GJ, Menon DK. Reliability and validity of the subjective component of the American Orthopaedic Foot and Ankle Society clinical rating scales. *The Journal of foot and ankle surgery*. 2007 Mar-Apr;46(2):65-74.
- Baumhauer JF, Nawoczenski DA, DiGiovanni BF, Wilding GE. Reliability and validity of the American Orthopaedic Foot and Ankle Society Clinical Rating Scale: a pilot study for the hallux and lesser toes. *Foot & ankle international*. 2006 Dec;27(12):1014-1019.
- Dawson J, Doll H, Coffey J, Jenkinson C. Responsiveness and minimally important change for the Manchester-Oxford foot questionnaire (MOXFQ) compared with AOFAS and SF-36 assessments following surgery for hallux valgus. *Osteoarthritis and cartilage / OARS, Osteoarthritis Research Society*. 2007 Aug;15(8):918-931.
- SooHoo NF, Vyas R, Samimi D. Responsiveness of the foot function index, AOFAS clinical rating systems, and SF-36 after foot and ankle surgery. *Foot & ankle international*. 2006 Nov;27(11):930-934.
- Hays RD, Sherbourne CD, Mazel RM. The RAND 36-Item Health Survey 1.0. *Health economics*. 1993 Oct;2(3):217-227.
- Ware J, Kosinski M, Keller S. SF-36 Physical and Mental Health Summary Scales: A User's Manual. The Health Institute, Boston, MA, 1994.
- VanderZee KI, Sanderman R, Heyink J. A comparison of two multidimensional measures of health status: the Nottingham Health Profile and the RAND 36-Item Health Survey 1.0. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation*. 1996 Feb;5(1):165-174.
- VanderZee KI, Sanderman R, Heyink JW, de Haes H. Psychometric qualities of the RAND 36-Item Health Survey 1.0: a multidimensional measure of general health status. *International journal of behavioral medicine*. 1996;3(2):104-122.
- Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). *Arthritis care & research*. 2011 Nov;63 Suppl 11:S240-52.
- Farrar JT, Young JP, Jr, LaMoreaux L, Werth JL, Poole RM. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain*. 2001 Nov;94(2):149-158.
- David W. Hosmer Jr. SL, Rodney X. Sturdivant. *Applied Logistic Regression*. 3rd ed: Wiley:New York; Apr 2013. 528 p.
- Youden WJ. Index for rating diagnostic tests. *Cancer*. 1950 Jan;3(1):32-35.
- Chen L, Lyman S, Do H, Karlsson J, Adam SP, Young E, et al. Validation of foot and ankle outcome score for hallux valgus. *Foot & ankle international*. 2012 Dec;33(12):1145-1155.
- Cook JJ, Cook EA, Rosenblum BI, Landsman AS, Roukis TS. Validation of the American College of Foot and Ankle Surgeons Scoring Scales. *The Journal of foot and ankle surgery : official publication of the American College of Foot and Ankle Surgeons*. 2011 Jul-Aug;50(4):420-429.
- Schrier JC, Palmieri LN, Verheyen CC, Jansen J, Koeter S. Patient-reported outcome measures in hallux valgus surgery. A review of literature. *Foot and ankle surgery : official journal of the European Society of Foot and Ankle Surgeons*. 2015 Mar;21(1):11-15.
- Spruce MC, Bowling FL, Metcalfe SA. A longitudinal study of hallux valgus surgical outcomes using a validated patient centred outcome measure. *Foot (Edinburgh, Scotland)*. 2011 Sep;21(3):133-137.
- King MT. A point of minimal important difference (MID): a critique of terminology and methods. *Expert review of*

- pharmacoeconomics & outcomes research. 2011 Apr;11(2):171-184.
28. Chan HY, Chen JY, Zainul-Abidin S, Ying H, Koo K, Rikhranj IS. Minimal Clinically Important Differences for American Orthopaedic Foot & Ankle Society Score in Hallux Valgus Surgery. *Foot & ankle international*. 2017 May;38(5):551-517.
29. Schneider W, Knahr K. Surgery for hallux valgus. The expectations of patients and surgeons. *International orthopaedics*. 2001;25(6):382-385.
30. Mickle KJ, Munro BJ, Lord SR, Menz HB, Steele JR. Cross-sectional analysis of foot function, functional ability, and health-related quality of life in older people with disabling foot pain. *Arthritis care & research*. 2011 Nov;63(11):1592-1598.
31. McGuigan FX, Hozack WJ, Moriarty L, Eng K, Rothman RH. Predicting quality-of-life outcomes following total joint arthroplasty. Limitations of the SF-36 Health Status Questionnaire. *The Journal of arthroplasty*. 1995 Dec;10(6):742-747.
32. Chong A, Nazarian N, Chandrananth J, Tacey M, Shepherd D, Tran P. Surgery for the correction of hallux valgus: minimum five-year results with a validated patient-reported outcome tool and regression analysis. *The bone & joint journal*. 2015 Feb;97-b(2):208-214.