ARE THERE RELIABLE AND VALID CLINICAL METHODS FOR MEASURING LEG LENGTH DISCREPANCY? Oosterveld F, Kamp N, de Jong R, ten Brinke A, van den Engh F; 1 Saxion University, Enschede, The Netherlands. 2 Hospital Medisch Spectrum Twente, Enschede, The Netherlands

PURPOSE: There is significant inconsistency regarding research into measurements of leg length discrepancy. RELEVANCE: The aim of this study was to investigate the intra observer reliability and validity of four different methods that are currently being used in clinical settings to determine leg length discrepancies. PARTICIPANTS: 25 Healthy male subjects, with a mean age of 49 years were enrolled in the study. Women were excluded from the study, because their reproductive organs cannot be protected adequately during the X-rays. In addition, also volunteers who had actual complaints of the musculoskeletal system and those who were unable to assume the desired position without complaints of pain were excluded. METHODS: The methods under study were 3 standing methods (the Calliet method, the Laser method and the Pelvic-leveling device method) and 1 supine method (the DeltaLeg method). The iliac crest symmetry was assessed and expressed on a nominal scale (left higher / even / right higher) and the difference, if present, was measured in cm (continuous scale). All measurements of the four different methods were performed twice by an experienced observer in the morning and twice by an inexperienced observer in the afternoon on the same subjects. Subjects were kept unaware of their own measurement results (blinded). To blind the observers each subject was assigned a different code number for each session and they were measured in a random order, without knowledge of the previous results. During the afternoon also an X-ray was made of each subject, showing the femur heads and the crista iliaca. This standing X-ray was used as a golden standard for validation. ANALYSIS: The data was analyzed using SPSS 12. To estimate intra observer agreement for nominal data Cohen’s Kappa was used and for continuous data Pearson’s correlation coefficient was calculated. Also, differences between the measurements were expressed in mean and standard deviation. To transform the group reliability results to the level of the individual patient from the SD’s the Smallest Detectable Difference (SDD) was calculated. RESULTS: Reliability for all the observers was relatively low. Cohen’s Kappa’s vary from 0.18 to 0.50 and SDD’s of 5.5 to 13.0 mm were found. The experienced observer had overall higher intra observer reliability than the inexperienced observer. Kappa’s (−0.05–0.17) and correlation coefficients (−0.12–0.39) showed very low validity of all methods as compared to X-ray. The reliability as well as validity increased to acceptable levels taking in account only the subjects in which a difference of ≥5 mm or ≥8 mm was found by the observers. CONCLUSIONS: This study showed that intra observer reliability and validity of the 4 tested measurement procedures in a healthy population is rather poor. The intra observer reliability and the validity of the methods improve with increasing leg length discrepancies. IMPLICATIONS: In general measured differences less than approximately 0.8 cm should not have any clinical or therapeutic consequences. KEYWORDS: measurements, validity, reliability. FUNDING ACKNOWLEDGEMENTS: None. CONTACT: f.g.j.oosterveld@saxion.nl

ETHICS COMMITTEE: Ethics committee of hospital Medisch Spectrum Enschede, The Netherlands