Reliability, validity and responsiveness of the Western Ontario McMaster Osteoarthritis Index (WOMAC) in the elderly population with a femoral neck fracture

Paul T.P.W. Burgers[1], Rudolf W. Poolman[2], Theodorus M.J. Van Bakel[1], Wim E. Tuinebreijer[1], Stephanie M. Zielinski[1], Mohit Bhandari[3], Peter Patka[4], Esther M.M. Van Lieshout[1] on behalf of the HEALTH and FAITH trial investigators*

[1] Trauma Research Unit Dept. of Surgery, Erasmus MC, University Medical Center Rotterdam, P.O. Box 2040, 3000 CA Rotterdam, the Netherlands

[2] Department of Orthopaedic Surgery, Joint Research, Onze Lieve Vrouwe Gasthuis, Amsterdam, the Netherlands

[3] Dept. of Clinical Epidemiology and Biostatistics, McMaster University, HSC 2C, 1200 Main Street West, Hamilton, ON, L8N 3Z5, Canada

[4] Dept. of Emergency Medicine, Erasmus MC, University Medical Center Rotterdam, P.O. Box 2010, 3000 CA Rotterdam, the Netherlands

Correspondence address:

E.M.M. van Lieshout, MSc PhD

Erasmus MC, University Medical Center Rotterdam

Trauma Research Unit, Department of Surgery

P.O. Box 2040, 3000 CA Rotterdam, The Netherlands

Phone: +31.10.7031050, Fax: +31.10.7032396

E-mail: e.vanlieshout@erasmusmc.nl

ABSTRACT

Background: Patient-reported outcome measures are gaining importance in clinical research. The WOMAC has been extensively evaluated in populations suffering from osteoarthritis, yet not in femoral neck fracture populations. This study aimed to determine the reliability, construct validity, and responsiveness of WOMAC, compared with SF-12 and EQ-5D, in elderly with a femoral neck fracture.

Methods: Reliability was tested by assessing Cronbach's alpha. Construct validity was determined by the Pearson correlation coefficient. Change scores were calculated from the 10 week and 12 months follow-up. Standardized response means (SRM; responsiveness) and floor and ceiling effects were determined. Analyses were performed for patients <80 versus ≥80 years of age.

Results: The WOMAC mean total score was 89 points prefracture in younger patients, increasing from 70 (10 weeks) to 81 (24 months). In the oldest old (*i.e.*, ≥ 80 years), these scores were 86, 75, and 78. The stiffness scores were 83, 67, and 76 in the younger group and 85, 80, and 80 in the oldest old. Pain scores were 92, 76, and 87 in the younger and 92, 84, and 93 in the older group. Function scores were 89, 68, and 79 in the younger and 84, 71, and 73 in the older group. Cronbach's alpha for pain, stiffness, function, and the total scale ranged 0.83-0.98 for the younger and 0.79-0.97 for the older group. Construct validity was good with 82% and 79% of predefined hypotheses confirmed in the younger and older group, respectively. Responsiveness was moderate. No floor effects were found. Moderate to large ceiling effects were found for the Pain and Stiffness scales at 10 weeks and 12 months in younger patients (18-36%) and in the oldest old (38-53%).

Conclusions: The WOMAC showed a good reliability, construct validity, and responsiveness in both age groups with a femoral neck fracture who were physically and mentally fit prefracture. The instrument is suitable for use in future clinical studies in these populations.

Level of Evidence: Diagnostic studies, level I

INTRODUCTION

Most patients with a femoral neck fracture are treated surgically[1, 2]. Objective determinants like mortality, complications, and revision surgery were traditionally used for evaluating quality of care. Nowadays, methods to assess the patient's perspective of treatment results are gaining importance[3]. These subjective patients' experiences are quantified using patient-reported outcome measures (PROMs). PROMs are available for assessing functional outcome and quality of life. For women at risk of hip fractures quality of life was profoundly threatened by falls and hip fractures[4], regardless of the type of fracture[5]. The problem in the hip fracture population is a complex assortment of issues ranging from baseline health and frailty, social isolation and support, mental status and joint function and pain. Different constructs can be assessed in the evaluation of hip fracture treatment. A frequently used disease-specific PROM is the WOMAC (Western Ontario and McMaster Osteoarthritis Index). To also cover the complexity of issues, the SF-12 (Short Form-12) and the EQ-5D (EuroQuality of Life, EuroQol-5D) are used for evaluating general health and health-related quality of life (HRQoL).

The WOMAC is a disease-specific 24-item questionnaire (scored on a 5-point Likert scale), measuring three domains: pain, stiffness, and function[10]. The WOMAC was designed for patients with osteoarthritis of the hip and knee[11]. Translation and cross-cultural validation was performed for different countries, including the Netherlands[12-15]. The WOMAC has frequently been used in orthopaedic patients, including hip fracture patients[16-18]. It has been extensively validated in patients who underwent knee and hip arthroplasty following osteoarthritis, but measurement properties in hip fracture patients is undetermined.

The aim of this study was to determine the reliability, validity, and responsiveness of the WOMAC, compared with SF-12 and EQ-5D in elderly who sustained a femoral neck fracture.

MATERIALS AND METHODS

Population

Between March 3, 2008 and February 14, 2011, 400 patients with a femoral neck fracture were enrolled in two multicenter trials; 150 were enrolled in the Dutch branch of the HEALTH trial, and 250 in the FAITH trial. Adult patients (aged ≥50 years) with a low energy femoral neck fracture without other major trauma, who were ambulant prefracture were considered eligible. Patients with a suspected pathological fracture, associated major injuries of the lower extremities, retained hardware or infection around the hip, bone metabolism disorder other than osteoporosis, cognitive impairment, dementia, or Parkinson's disease as well as patient who were not likely to be able to complete follow-up were excluded. Data on gender, age, American Society of Anesthesiologists (ASA) classification, walking status, and living situation were collected. The ethics committees approved this study and all patients signed informed consent.

Questionnaires

All 400 patients were asked to complete the three multidimensional questionnaires, the WOMAC, SF-12 and EQ-5D in clinic and at each visit to the outpatient department.

The SF-12 score ranges between 0 and 100 (best) and consists of a Physical Component Summary (PCS-12) and Mental Component Summary (MCS-12). The SF-12 represents a plausible alternative to the larger 36-item SF-36 for measuring health status, especially in large scale studies with a need to reduce questionnaire length[6]. The reliability and validity were determined in different populations[6], and countries, including the Netherlands[7]. The EQ-5D is a reliable and valid instrument for assessing HRQoL in elderly patients with a femoral neck fracture[8]. The domains mobility, self-care, daily activities, pain, and anxiety are tested on a three-point Likert scale, resulting in a utility score (EQ-US) that ranges from 0 to 1 (maximum)[9]. In addition, patients score their health status on a visual analog scale (EQ-VAS) that ranges from 0 (indicating worst possible health) to 100 (indicating best possible health)[9].

Data about the prefracture situation were gathered during visits within one week after surgery, other follow-up moments were at 10 weeks, six, 12 and 24 months after surgery. Questionnaires were also completed during a telephone interview at nine and 18 months. Patients did not receive any assistance in their completion of the questionnaires and did not consider the burden of completion too high. If necessary (*e.g.*, due to physical status early in recovery), interview-administered versions were used.

Data analysis

Data were analyzed with SPSS version 20.0 (IBM Corp. Released 2011. IBM SPSS Statistics for Windows, Armonk, NY: IBM Corp). Data are reported using the COSMIN checklist[19].

Outcome scores per follow up moment for the all instruments and subdomains were calculated. Unless mentioned otherwise, all further analyses were performed for two age groups (*i.e.*, <80 years and ≥ 80 years) and for 10 weeks and 12 months.

The reliability of the WOMAC instrument was tested by determining the Cronbach's alpha as measure of internal consistency. Cronbach's alpha was used as reliability parameter **as** there was no repetition measurement by different observers or at different time- points per follow-up. Qestionnaires aimed at measuring the same construct (WOMAC Pain, SF-12 Pain and EQ-Pain) were used as repetition measurement[20, 21]. A Cronbach's alpha within the range of 0.70 to 0.90 was considered acceptable.[20].

In the absence of a gold standard for hip function after a hip fracture construct validity was determined by calculating the Pearson correlation coefficient. Construct validity refers to the extent to which scores on a particular measure relate to other measures in a manner that is consistent with theoretically derived hypotheses concerning the concepts being measured[21]. Only questionnaires of which the WOMAC total score was available were included. Hypotheses for construct validity were defined before data analysis (**Appendix 2**). Correlation coefficients of 0.10-0.30 were considered weak, 0.3-0.6 moderate, and >0.6 strong[22].

Responsiveness is the ability of a questionnaire to detect clinically important changes over time [21]. Change scores of the instruments were calculated from the 10 week and 12 months follow up. To assess responsiveness, standardized response means (SRM) were calculated by dividing the mean change by the SD of this change. These effect estimates were interpreted according to Cohen: a SRM of 0.2–0.4 was considered a small effect, 0.5–0.7 moderate, and 0.8 large[23].

A floor or ceiling effect was considered present if more than 15% of the patients in a sample size of 50 patients achieved the lowest or highest possible score, respectively[21].

Missing data

As the raw data for individual items were analyzed, missing data were not imputed.

Source of Funding

Research grants were received from Stichting NutsOhra (The Netherlands) and The Netherlands Organization for Health Research and Development (The Netherlands). These organizations had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

RESULTS

Population

Of all 400 included patients 275 were aged <80 years and 125 were aged ≥ 80 years (Table 1); 250 patients (62.5%) were female and 386 (96.5%) lived independently. This was similar in both age groups. Prefracture, 325 patients (81.2%) walked without aid; 248 (90.2%) and 77 (61.6%) in the younger and older age groups, respectively. Of the total group, 289 patients (72.2%) had ASA class I/II; 223 (81.8%) in the younger age group versus 66 (52.8% in the oldest old. In the internal fixation group adverse events occurred in 41% of the patients, in the arthroplasty group this was the case in 49%, ranging from bladder infection to revision surgery.

Outcome scores

Patients aged <80 years reported a mean prefracture WOMAC Total score (Figure 1A and Table 2) of 89 points (SD 15), postfracture the score increased from 70 (SD 22) at 10 weeks to 81 (SD 20) after 24 months. In the oldest old, corresponding scores were 86 (SD 16), 75 (SD 21), and 78 (SD 18). Similar changes in scores were reported for WOMAC Stiffness, WOMAC Pain, and WOMAC Function (Figures 1B-D and Table 2).

SF-12 PCS scores decreased after fracture and increased again over time (Figure 2A and Table 2), yet SF-12 MCS remained more stable over time (Figure 2B and Table 2). Likewise, EQ-US scores decreased after fracture and increased again over time (Figure 2C and Table 2), whereas EQ-VAS remained more stable over time (Figure 2D and Table 2).

WOMAC reliability

Cronbach's alpha for the domains pain, stiffness, function and the total scale were between 0.83 and 0.98 for the younger age group and between 0.79 and 0.97 for the oldest old (**Table 3**).

Construct validity

All Pearson correlations were statistically significant at a $p<0.01$ (2-tailed; **Table 4**). In six main hypotheses, 28 components were predicted, of which 23 (82%) were correctly hypothesized a priori in the younger age group and 22 (79%) in the oldest old. Unconfirmed predictions were mainly underestimations (hypothesis 3). WOMAC Stiffness correlated moderately (*r*>0.35-0.55) with all other scores, while weak correlations (*r*= 0.1-0.3) were expected. WOMAC Pain correlated strongly with EQ-US (*r* = 0.74 in the younger age group and 0.72 in the oldest old), whereas moderate correlation (*r* 0.3-0.6) was expected.

Responsiveness

Descriptive statistics and responsiveness is presented in **Table 5**. The SRM was moderate for WOMAC Function (SRM 0.64) and the WOMAC Total score (SRM 0.66). WOMAC subscales Pain and Function showed small SRMs, ranging from 0.21 to 0.30.

Floor and ceiling effects

None of the WOMAC domains nor the Total score showed a floor effect. No ceiling effects were found for the Function domain or the Total score in either of the age groups (**Table 6**). Pain and Stiffness showed a ceiling effect at 10 weeks and 12 months postsurgery for the younger age group (18-36%) as well as for the oldest old (38-53%).

DISCUSSION

The current study was the first to determine reliability, validity, and responsiveness of the WOMAC, compared with the SF-12 and EQ-5D, in elderly patients who sustained a femoral neck fracture and who were physically and mentally fit prefracture. Results indicated that WOMAC is a reliable and valid PROM for these patients in patients aged <80 years and ≥ 80 years. Responsiveness was also sufficient, indicating the instrument can be used for measuring changes in scores over time.

(Sub)scores of all PROMs showed similar patterns over time. Scores reduced (most likely immediately) post surgery; from 10 weeks after surgery onwards a gradual increase over time was noted. After two years the scores approximated the prefracture scores, indicating a small residual decrease in mental and physical functioning. Due to differences in symptom evolution, the clinimetric results from osteoarthritis populations are not readily applicable to hip fracture populations. The minimal clinically important change (MIC) for hip fracture patients has not been reported, yet changes in WOMAC Pain and Function were larger than the MIC reported for hip replacement after osteoarthritis (10 points for pain and 8-9 points for function)[24, 25]. There may be two reasons why changes in WOMAC over time were smaller than observed after hip replacement for osteoarthritis (for pain: 53 points at baseline to 94 two years after surgery and from 50 to 92 for function)[26]. First, true baseline scores for hip fracture patients (*i.e.,* between fracture and surgery) would have been much lower. Although these scores are unknown (as asking patients to complete questionnaires before surgery is not feasible), we expect these scores to be even lower than preoperative scores in osteoarthritis patients[26]. Calculating changes in WOMAC scores relative to baseline would have resulted in much larger changes over time. Second, at the first follow-up moment (10 weeks) rehabilitation likely has progressed considerably already in the majority of patients, as participation required patients to be fit prefracture.

The construct validity of the WOMAC instrument was good, with 82% (**<80** years) and 79% (≥ 80 years) of the predefined hypotheses being true. Especially the strong correlations of WOMAC Function and Total score with SF-12 PCS, SF-12 Total score and EQ-US support the preferred use of WOMAC for assessing functional recovery in this population.

Responsiveness of the WOMAC instrument as a whole was moderate in patients **<80** years, as the SRM was 0.66. In the oldest old, the SRM was small (0.24). For hip fracture patients the SRM has not been shown before. However, for patients with hip osteoarthritis the SRM of the WOMAC exceeded 1.0[27, 28], this is mainly due to much larger changes in score between the two measurements. Larger changes for those patients are expected, as the WOMAC score is at the lowest prior to the hip replacement procedure. These true baseline scores cannot be determined for hip fracture patients. The Dutch version of the non-hip-specific Short Musculoskeletal Function Assessment questionnaire (SFMA) appeared to be moderately responsive (SRMs for subscales 0.17-0.47) for patients with a variety of musculoskeletal disorders[29]. In different Swedish hip fracture populations the SRMs were moderate to large (0.76-0.96) for the SMFA and small to large for the non-disease-specific instruments EQ-5D (SRM 0.01-1.14) and Nottingham Health Profile questionnaire (SRM 0.09-0.98)[30, 31].

There was no evidence for any floor effect of the WOMAC instrument or for a ceiling effect of WOMAC Function and Total score. In contrast, moderate to large ceiling effects (18-53%) were found for Pain and Stiffness at 10 weeks and 12 months in both age groups. This is similar as reported[32-34]. It may reflect the narrow discriminating capacity of these WOMAC domains in the studied population. Some selection bias might have played a role, as patients had to be physically and mentally fit to participate. Participants were ambulant and lived independently prefracture.

Strengths and limitations

WOMAC has been used for assessing functional outcomes and quality of life in hip fracture populations before[35, 36]. In the current study, it was validated in hip fracture populations for the first time. Rates of adverse events in these populations doesn't seem to differ from literature, they may have affect the scores, but not the validity of the questionnaire. We consider the novelty of validating WOMAC for hip fracture patients a strength of our study. Moreover, we consider the studied populations representative for a daily practice hip fracture population and believe the results apply to a generally fit hip fracture population. Whether or not WOMAC is also valid for use in frail elderly remains unknown, as patients that were non-ambulatory prefracture and patients with *e.g.*, Parkinson's disease, pathological fractures, and dementia were excluded. Another strength was the built-in very short period, only several days postfracture, for the self-reported, pre-injury disability evaluation to be completed, minimizing the risk of recall bias[37].

A limitation could be the use of arbitrary hypotheses for construct validation, although they were predefined in compliance with clinimetric evaluation guidelines[19]. Second, some selection bias might have led to overestimation of the outcomes, especially ceiling effects. This effect is also known from the Harris Hip Score which is frequently used in orthopedic research, but no gold standard exists for functional evaluation of hip fracture populations[38, 39].

Third, data completeness was not 100% at each time point. One can imagine that patients were not able to complete the forms especially if they were in bad condition. These 36% missing items causing a missing WOMAC Total score might have influenced the outcomes resulting in the current more favorable mean outcome scores and large ceiling effects.

CONCLUSION

The WOMAC, a widely used disease-specific questionnaire, shows adequate reliability and construct validity in patients aged 50 years or older with a femoral neck fracture who were physically and mentally fit prefracture. Responsiveness was better for younger patients than for patients ≥80 years. It is therefore a suitable instrument for use in future clinical studies in this population.

REFERENCES

1. Handoll HH, Parker MJ. Conservative versus operative treatment for hip fractures in adults. Cochrane Database Syst Rev. 2008 (3):CD000337.

2. Raaymakers EL. The non-operative treatment of impacted femoral neck fractures. Injury. 2002 Dec;33 Suppl 3:C8-14.

3. Rolfson O, Rothwell A, Sedrakyan A, Chenok KE, Bohm E, Bozic KJ, Garellick G. Use of patient-reported outcomes in the context of different levels of data. J Bone Joint Surg Am. 2011 Dec 21;93 Suppl 3:66-71. Epub 2012/01/25.

4. Salkeld G, Cameron ID, Cumming RG, Easter S, Seymour J, Kurrle SE, Quine S. Quality of life related to fear of falling and hip fracture in older women: a time trade off study. Bmj. 2000 Feb 5;320(7231):341-6.

5. Mendonca TM, Silva CH, Canto RS, Morales NM, Pinto RM, Morales RR. Evaluation of the health-related quality of life in elderly patients according to the type of hip fracture: femoral neck or trochanteric. Clinics (Sao Paulo). 2008 Oct;63(5):607-12.

6. Ware J, Jr., Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care. 1996 Mar;34(3):220-33.

7. Gandek B, Ware JE, Aaronson NK, Apolone G, Bjorner JB, Brazier JE, Bullinger M, Kaasa S, Leplege A, Prieto L, Sullivan M. Cross-validation of item selection and scoring for the SF-12 Health Survey in nine countries: results from the IQOLA Project. International Quality of Life Assessment. J Clin Epidemiol. 1998 Nov;51(11):1171-8.

8. Tidermark J, Zethraeus N, Svensson O, Tornkvist H, Ponzer S. Quality of life related to fracture displacement among elderly patients with femoral neck fractures treated with internal fixation. J Orthop Trauma. 2002 Jan;16(1):34-8.

9. Brooks R. EuroQol: the current state of play. Health Policy. 1996 Jul;37(1):53-72.
10. Bellamy N. WOMAC Osteoarthritis Index. 2012 [updated May 2012; cited 2013 21-08]; Available from: http://www.womac.org/womac/index.htm.
11. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: a health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. J Rheumatol. 1988 Dec;15(12):1833-40.
12. Basaran S, Guzel R, Seydaoglu G, Guler-Uysal F. Validity, reliability, and comparison of the WOMAC osteoarthritis index and Lequesne algofunctional index in Turkish patients with hip or knee osteoarthritis. Clin Rheumatol. 2010 Jul;29(7):749-56.
13. Soderman P, Malchau H. Validity and reliability of Swedish WOMAC osteoarthritis index: a self-administered disease-specific questionnaire (WOMAC) versus generic instruments (SF-36 and NHP). Acta Orthop Scand. 2000 Feb;71(1):39-46.
14. Roorda LD, Jones CA, Waltz M, Lankhorst GJ, Bouter LM, van der Eijken JW, Willems WJ, Heyligers IC, Voaklander DC, Kelly KD, Suarez-Almazor ME. Satisfactory cross cultural equivalence of the Dutch WOMAC in patients with hip osteoarthritis waiting for arthroplasty. Ann Rheum Dis. 2004 Jan;63(1):36-42.
15. Thumboo J, Chew LH, Soh CH. Validation of the Western Ontario and Mcmaster University osteoarthritis index in Asians with osteoarthritis in Singapore. Osteoarthritis Cartilage. 2001 Jul;9(5):440-6.
16. Jain R, Koo M, Kreder HJ, Schemitsch EH, Davey JR, Mahomed NN. Comparison of early and delayed fixation of subcapital hip fractures in patients sixty years of age or less. J Bone Joint Surg Am. 2002 Sep;84-A(9):1605-12. Epub 2002/09/05.

17. Edwards M, Baptiste S, Stratford PW, Law M. Recovery after hip fracture: what can we learn from the Canadian Occupational Performance Measure? Am J Occup Ther. 2007 May-Jun;61(3):335-44. Epub 2007/06/16.

18. Macaulay W, Nellans KW, Iorio R, Garvin KL, Healy WL, Rosenwasser MP. Total hip arthroplasty is less painful at 12 months compared with hemiarthroplasty in treatment of displaced femoral neck fracture. HSS J. 2008 Feb;4(1):48-54. Epub 2008/08/30.

19. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, Bouter LM, de Vet HC. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res. 2010 May;19(4):539-49. Epub 2010/02/20.

20. Henrica C. W. de Vet CBT, Lidwine B. Mokkink, Dirk L. Knol. Measurement in Medicine A Practical Guide Cambridge University Press; 2011.

21. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, Bouter LM, de Vet HC. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007 Jan;60(1):34-42. Epub 2006/12/13.

22. Cohen J. Statistical power analysis for the behavioral sciences. New York: Academic Press; 1977. 474 p.

23. Cohen J. Statistical Power Analysis for the Behavioral Sciences (2nd Edition). second ed. Hillsdale (NJ): Lawrence Erlbaum Associates; 1988 July 1, 1988. 590 p.

24. Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, Bombardier C, Felson D, Hochberg M, van der Heijde D, Dougados M. Evaluation of clinically relevant changes in patient reported outcomes in knee and hip osteoarthritis: the minimal clinically important improvement. Ann Rheum Dis. 2005 Jan;64(1):29-33. Epub 2004/06/23.

25. Ehrich EW, Davies GM, Watson DJ, Bolognese JA, Seidenberg BC, Bellamy N. Minimal perceptible clinical improvement with the Western Ontario and McMaster Universities osteoarthritis index questionnaire and global assessments in patients with osteoarthritis. J Rheumatol. 2000 Nov;27(11):2635-41. Epub 2000/11/28.

26. Goodman SM, Ramsden-Stein DN, Huang WT, Zhu R, Figgie MP, Alexiades MM, Mandl LA. Patients with Rheumatoid Arthritis Are More Likely to Have Pain and Poor Function After Total Hip Replacements than Patients with Osteoarthritis. J Rheumatol. 2014 Seo;41(9):1774-80. Epub 2014/08/01.

27. Naal FD, Impellizzeri FM, von Eisenhart-Rothe R, Mannion AF, Leunig M. Reproducibility, validity, and responsiveness of the hip outcome score in patients with end-stage hip osteoarthritis. Arthritis Care Res (Hoboken). 2012 Nov;64(11):1770-5. Epub 2012/06/08.

28. Bilbao A, Quintana JM, Escobar A, Las Hayas C, Orive M. Validation of a proposed WOMAC short form for patients with hip osteoarthritis. Health Qual Life Outcomes. 2011;9:75. Epub 2011/09/23.

29. Reininga IH, el Moumni M, Bulstra SK, Olthof MG, Wendt KW, Stevens M. Cross-cultural adaptation of the Dutch Short Musculoskeletal Function Assessment questionnaire (SMFA-NL): internal consistency, validity, repeatability and responsiveness. Injury. 2012 Jun;43(6):726-33. Epub 2011/08/20.

30. Hedbeck CJ, Tidermark J, Ponzer S, Blomfeldt R, Bergstrom G. Responsiveness of the Short Musculoskeletal Function Assessment (SMFA) in patients with femoral neck fractures. Qual Life Res. 2011 May;20(4):513-21. Epub 2010/11/13.

31. Tidermark J, Bergstrom G. Responsiveness of the EuroQol (EQ-5D) and the Nottingham Health Profile (NHP) in elderly patients with femoral neck fractures. Qual Life Res. 2007 Mar;16(2):321-30. Epub 2006/10/13.

32. Escobar A, Quintana JM, Bilbao A, Arostegui I, Lafuente I, Vidaurreta I. Responsiveness and clinically important differences for the WOMAC and SF-36 after total knee replacement. Osteoarthritis Cartilage. 2007 Mar;15(3):273-80. Epub 2006/10/21.

33. Yang KG, Raijmakers NJ, Verbout AJ, Dhert WJ, Saris DB. Validation of the short-form WOMAC function scale for the evaluation of osteoarthritis of the knee. J Bone Joint Surg Br. 2007 Jan;89(1):50-6. Epub 2007/01/30.

34. Angst F, Aeschlimann A, Steiner W, Stucki G. Responsiveness of the WOMAC osteoarthritis index as compared with the SF-36 in patients with osteoarthritis of the legs undergoing a comprehensive rehabilitation intervention. Ann Rheum Dis. 2001 Sep;60(9):834-40. Epub 2001/08/15.

35. Sanz-Reig J, Lizaur-Utrilla A, Serna-Berna R. Outcomes in nonagenarians after hemiarthroplasty for femoral neck fracture. A prospective matched cohort study. Hip Int. 2012 Jan-Feb;22(1):113-8. Epub 2012/03/03.

36. Macaulay W, Nellans KW, Garvin KL, Iorio R, Healy WL, Rosenwasser MP. Prospective randomized clinical trial comparing hemiarthroplasty to total hip arthroplasty in the treatment of displaced femoral neck fractures: winner of the Dorr Award. J Arthroplasty. 2008 Sep;23(6 Suppl 1):2-8. Epub 2008/09/09.

37. Williamson OD, Gabbe BJ, Sutherland AM, Hart MJ. Does recall of preinjury disability change over time? Inj Prev. 2013 Aug;19(4):238-43. Epub 2012/12/01.

38. Wamper KE, Sierevelt IN, Poolman RW, Bhandari M, Haverkamp D. The Harris hip score: Do ceiling effects limit its usefulness in orthopedics? Acta Orthop. 2010 Dec;81(6):703-7. Epub 2010/11/30.

39. Nilsdotter A, Bremander A. Measures of hip function and symptoms: Harris Hip Score (HHS), Hip Disability and Osteoarthritis Outcome Score (HOOS), Oxford Hip Score (OHS), Lequesne Index of Severity for Osteoarthritis of the Hip (LISOH), and American Academy of Orthopedic Surgeons (AAOS) Hip and Knee Questionnaire. Arthritis Care Res (Hoboken). 2011 Nov;63 Suppl 11:S200-7. Epub 2012/05/25.

40. Cauley JA. Public Health Impact of Osteoporosis. J Gerontol A Biol Sci Med Sci. 2013 Jul 31.

41. Bhandari M, Sprague S, Schemitsch EH. Resolving controversies in hip fracture care: the need for large collaborative trials in hip fractures. J Orthop Trauma. 2009 Jul;23(6):479-84. Epub 2009/06/25.

42. de Groot IB, Reijman M, Terwee CB, Bierma-Zeinstra SM, Favejee M, Roos EM, Verhaar JA. Validation of the Dutch version of the Hip disability and Osteoarthritis Outcome Score. Osteoarthritis Cartilage. 2007 Jan;15(1):104-9.

43. Ostendorf M, van Stel HF, Buskens E, Schrijvers AJ, Marting LN, Verbout AJ, Dhert WJ. Patient-reported outcome in total hip replacement. A comparison of five instruments of health status. J Bone Joint Surg Br. 2004 Aug;86(6):801-8.

44. Kim SM, Moon YW, Lim SJ, Yoon BK, Min YK, Lee DY, Park YS. Prediction of survival, second fracture, and functional recovery following the first hip fracture surgery in elderly patients. Bone. 2012 Jun;50(6):1343-50.

45. Calvert M, Blazeby J, Altman DG, Revicki DA, Moher D, Brundage MD. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. Jama. 2013 Feb 27;309(8):814-22.

46. Hoang-Kim A, Beaton D, Bhandari M, Kulkarni AV, Schemitsch E. The need to standardize functional outcome in randomized trials of hip fracture: a review using the ICF framework. J Orthop Trauma. 2013 Jan;27(1):e1-8.

Figure Legends

Figure 1. Changes in functional outcome scores over time

Mean and SD are shown for the WOMAC Total score (**A**), WOMAC Stiffness score (**B**), WOMAC Pain score (**C**), WOMAC Function score (**D**), and per follow up moment. Higher scores represent better functional outcomes.

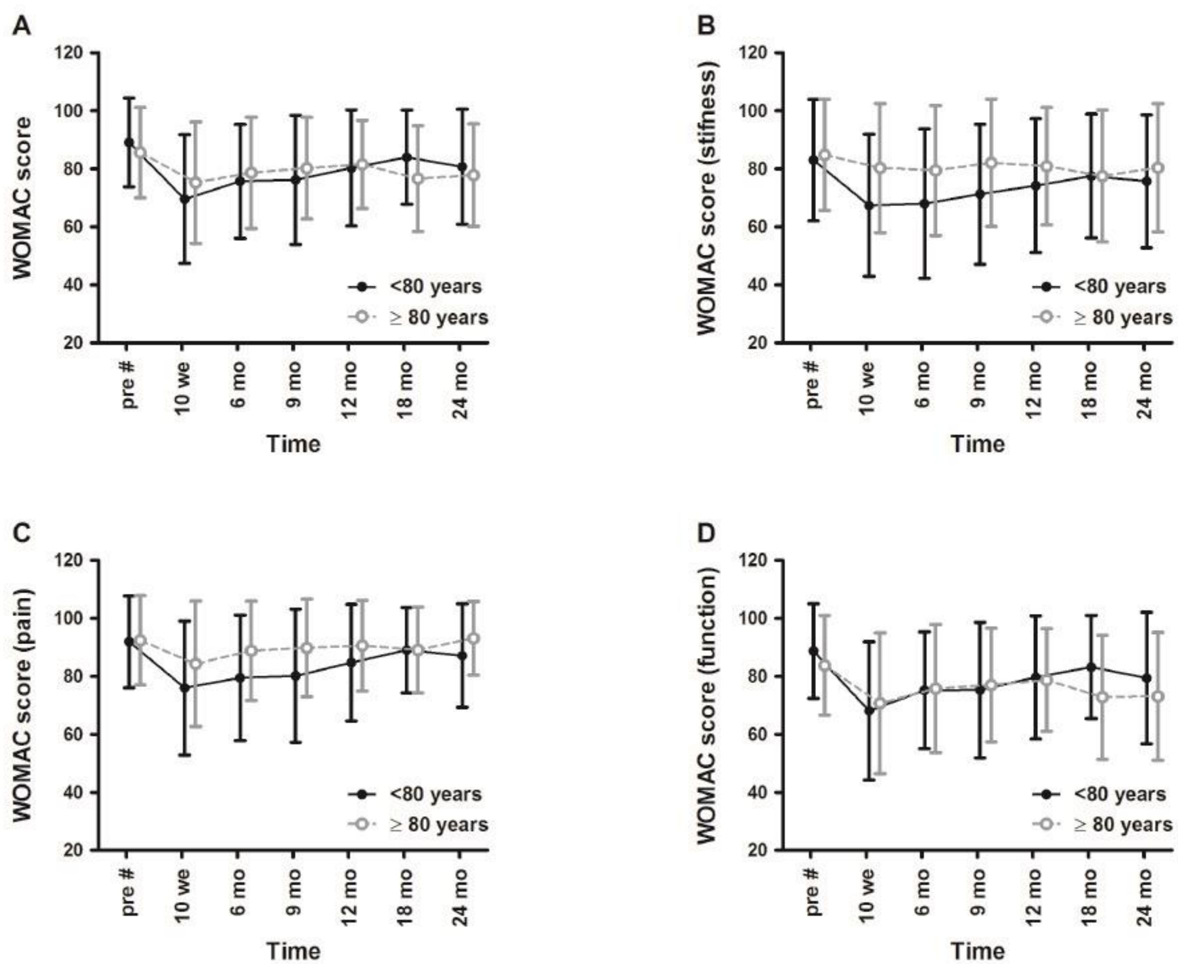

Figure 2. Changes in quality of life scores over time, separated by age group

(**A**) SF-12 Physical Component Summary (PCS) and (**B**) SF-12 Mental Component Summary (MCS) are from the Short Form-12 (SF-12) instrument; (**C**) EuroQol-5D Utility Score (EQUS) and (**D**) EuroQol-5D Visual Analog Scale (EQVAS) are from the EQ-5D score. Higher scores represent a better quality of life.

Mean and SD are shown per follow up moment.

For the SF-12, the dotted line shows the standardized mean (50 points) of the general population, the shaded area is the area within 1 SD deviation from this mean (10 points).

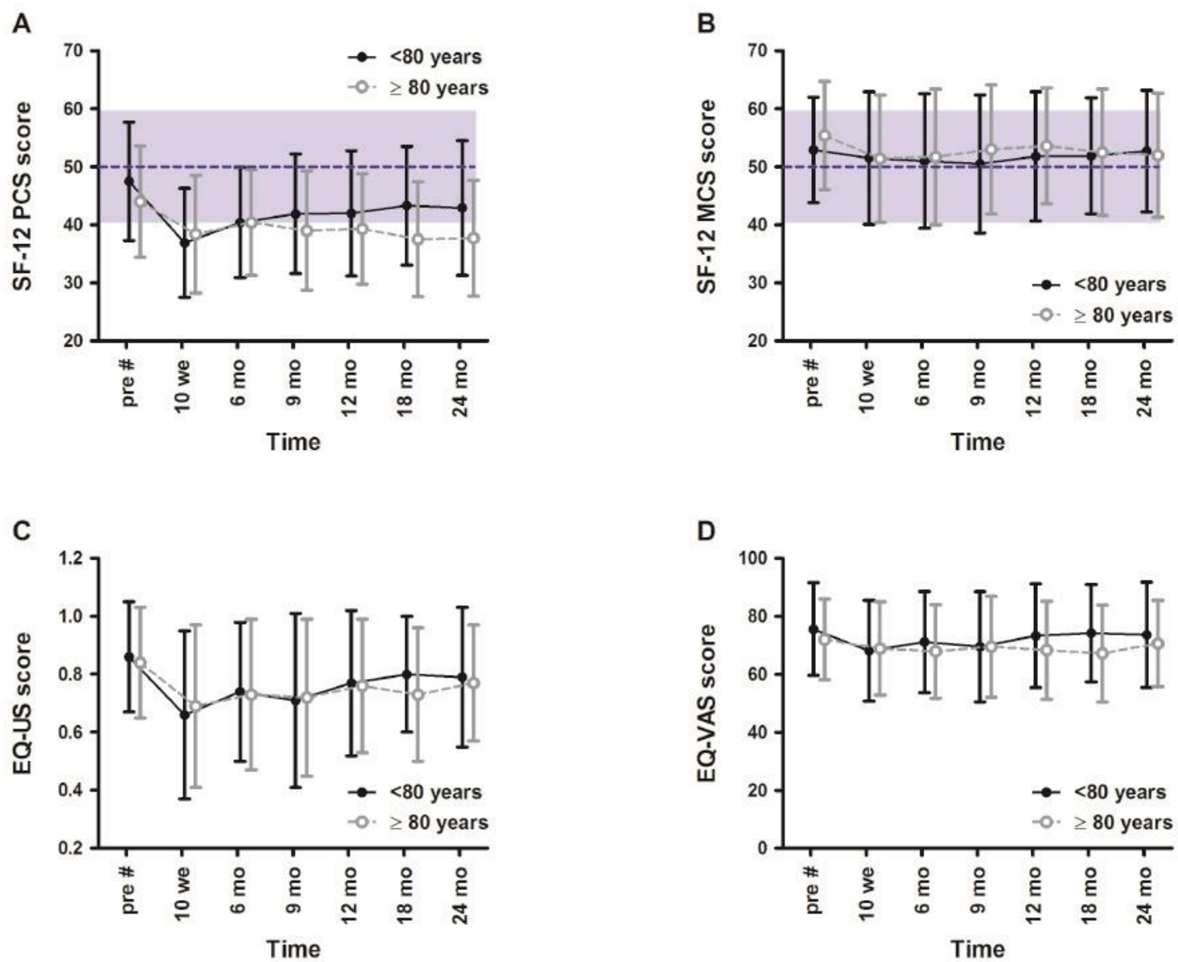

Table 1. Demographic description of the study population, separated by age group

	Total (N=400)	Age <80 years (N=275)	Age ≥ 80 years (N=125)
Female	251 (62.8%)	173 (62.9%)	78 (62.4%)
Mean age (years)	74 (10)	69 (8)	85 (4)
ASA I/II	289 (72.3%)	223 (81.1%)	66 (52.8)
Walking without aids prefracture	325 (81.3%)	248 (90.2%)	77 (61.6%)
Living independently prefracture	386 (96.5%)	268 (97.5%)	118 (94.4%)

SD, standard deviation; ASA, American Society of Anesthesiologists.

Data are shown as number (%) or as mean (SD).

Table 2. Outcomes scores of the study population, separated by age group

Instrument	Age <80 years (N=275)			Age ≥ 80 years (N=125)		
	Prefracture	10 weeks	2 years	Prefracture	10 weeks	2 years
WOMAC Stiffness	83 (21)	67 (25)	76 (23)	85 (19)	80 (23)	80 (22)
WOMAC Pain	92 (16)	76 (23)	87 (18)	83 (15)	84 (22)	93 (13)
WOMAC Function	89 (16)	79 (23)	69 (24)	84 (17)	71 (24)	73 (22)
WOMAC Total score	89 (15)	70 (22)	81 (20)	86 (16)	75 (21)	78 (18)
SF-12 PCS	47 (10)	37 (9)	43 (12)	44 (10)	38 (10)	38 (10)
SF-12 MCS	53 (9)	51 (11)	53 (11)	55 (9)	51 (11)	52 (11)
EQ US	0.86 (0.19)	0.66 (0.29)	0.79 (0.24)	0.84 (0.19)	0.69 (0.29)	0.77 (0.20)
EQ VAS	76 (16)	68 (17)	74 (18)	72 (14)	69 (16)	71 (15)

Data are shown as mean (SD).

Table 3. Reliability of the WOMAC instrument at 10 weeks and 12 months in patients who sustained a femoral neck fracture, separated by age group

WOMAC domain	Number of items	Cronbach's alpha (N)			
		Age <80 years (N=275)		Age ≥ 80 years (N=125)	
		10 weeks	12 months	10 weeks	12 months
Stiffness	2	0.83 (229)	0.83 (204)	0.79 (113)	0.79 (95)
Pain	5	0.92 (191)	0.92 (186)	0.85 (67)	0.83 (70)
Function	17	0.97 (107)	0.97 (130)	0.97 (22)	0.93 (36)
Total scale	24	0.98 (105)	0.98 (128)	0.97 (21)	0.94 (35)

N; number of available questionnaires with all items completed per domain and with all items completed.

The Cronbach's alpha is given with the number of patients included in the analysis between brackets.

Table 4. Construct validity of the WOMAC domains and WOMAC Total score and the domains of the SF-12 and EQ-5D instruments at 10 weeks , separated by age group

		WOMAC							
		Age <80 years (N=275)				Age ≥ 80 years (N=125)			
		Stiffness	**Pain**	**Function**	**Total score**	**Stiffness**	**Pain**	**Function**	**Total score**
SF-12 PCS	***r***	**0.35**	**0.56**	**0.71**	**0.70**	**0.41**	**0.52**	**0.74**	**0.710**
	95% CI	**(0.23-0.46)**	**(0.46-0.65)**	**(0.62-0.77)**	**(0.61-0.77)**	**(0.24-0.55)**	**(0.37-0.64)**	**(0.61-0.83)**	**(0.58-0.81)**
	N	222	218	177	177	110	110	72	72
SF-12 MCS	***r***	**0.41**	**0.47**	**0.51**	**0.52**	**0.45**	**0.59**	**0.57**	**0.64**
	95% CI	**(0.29-0.51)**	**(0.36-0.57)**	**(0.40-0.61)**	**(0.40-0.62)**	**(0.29-0.59)**	**(0.46-0.70)**	**(0.39-0.71)**	**(0.47-0.76)**
	N	222	218	177	177	110	110	72	72
SF-12 Pain	***r***	0.55	**0.74**	**0.77**	**0.79**	0.51	**0.78**	**0.59**	**0.66**
	95% CI	(0.45-0.63)	**(0.68 -0.80)**	**(0.70-0.82)**	**(0.72-0.84)**	(0.36-0.63)	**(0.70-0.84)**	**(0.42-0.72)**	**(0.50-0.77)**
	N	228	224	179	179	113	112	73	73
SF-12 Total	***r***	0.50	**0.67**	**0.78**	**0.78**	0.55	0.71	**0.81**	**0.84**
	95% CI	(0.39-0.59)	**(0.59-0.74)**	**(0.71-0.83)**	**(0.71-0.83)**	(0.40-0.67)	(0.60-0.79)	**(0.71-0.88)**	**(0.75-0.90)**
	N	222	218	177	177	110	110	72	72

EQ-US	***r***	0.42	0.74	**0.76**	**0.78**	0.52	0.72	**0.77**	**0.80**
	95% CI	(0.30-0.52)	(0.67-0.79)	**(0.69-0.82)**	**(0.71-0.83)**	(0.37-0.64)	(0.61-0.80)	**(0.66-0.85)**	**(0.70-0.87)**
	N	228	223	179	179	113	112	73	73
EQ-Pain	***r***	0.44	**0.80**	**0.72**	**0.75**	0.49	**0.75**	**0.53**	**0.60**
	95% CI	(0.33-0.54)	**(0.74-0.84)**	**(0.64-0.78)**	**(0.68 -0.81)**	(0.34-0.62)	**(0.65-0.82)**	**(0.34-0.68)**	**(0.43-0.73)**
	N	229	224	179	179	113	112	73	73
EQ-VAS	***r***	**0.39**	**0.51**	**0.66**	**0.65**	**0.41**	**0.55**	**0.63**	**0.66**
	95% CI	**(0.27-0.49)**	**(0.41-0.60)**	**(0.57-0.73)**	**(0.56-0.73)**	**(0.25-0.56)**	**(0.40-0.67**)	**(0.47-0.75)**	**(0.50-0.77)**
	N	229	224	179	179	112	111	72	72

N; number of available questionnaires with all items completed per domain and with all items completed.

The Pearson correlation coefficient (*r*) is given with its 95% confidence interval in brackets.

Correlation is significant at the 0.01 level (2-tailed) for all comparisons.

Correlations that were predicted correctly are given in boldface.

Table 5. Responsiveness of the WOMAC domains and WOMAC Total score, separated by age group

	Age <80 years (N=275)				Age ≥ 80 years (N=125)			
	10 weeks Mean (SD)	12 months Mean (SD)	Change Mean (SD)	SRM (N)	10 weeks Mean (SD)	12 months Mean (SD)	Change Mean (SD)	SRM (N)
WOMAC Stiffness	67 (24)	74 (23)	5.5 (25.5)	0.21 (185)	80 (22)	81 (20)	-0.3 (25.2)	-0.01 (93)
WOMAC Pain	76 (23)	85 (20)	6.9 (22.8)	0.30 (184)	84 (22)	91 (16)	5.1 (19.4)	0.26 (92)
WOMAC Function	68 (24)	80 (21)	11.7 (18.2)	0.64 (145)	71 (24)	79 (18)	5.0 (19.1)	0.26 (54)
WOMAC Total score	70 (22)	80 (20)	10.8 (16.4)	0.66 (144)	75 (21)	82 (15)	4.1 (16.6)	0.24 (54)

Scores at 10 weeks and at 12 months as well as the difference between these scores (change), are shown as median with SD.

The Standardized Response Mean (SRM) is given with the number of patients used in the analysis between brackets.

Table 6. Ceiling effect of the WOMAC domains and WOMAC Total score, separated by age group

	Age <80 years (N=275)			Age ≥ 80 years (N=125)		
	N	10 weeks N (%)	12 months N (%)	N	10 weeks N (%)	12 months N (%)
WOMAC Stiffness	224	40 (17.9%)	72 (35.6%)	112	43 (38.4%)	50 (52.6%)
WOMAC Pain	229	46 (20.1%)	53 (26.1%)	113	48 (42.5%)	38 (40.0%)
WOMAC Function	179	7 (3.9%)	26 (13.7%)	73	3 (4.1%)	8 (10.3%)
WOMAC Total score	179	2 (0.1%)	13 (6.9%)	73	2 (2.7%)	7 (9.0%)

The number of patients reporting the maximum score of 100 points are given with the percentage given in brackets. N represents the total number of questionnaires used for the analysis. None of the instruments demonstrated a floor effect (*i.e.*, 0 points).

Appendix 1. List of international collaborators

HEALTH and FAITH trial investigators*

HEALTH trial:

Research grants were received from the following: Canadian Institutes of Health Research (CIHR) (PI: M Bhandari, Co-PI: GH Guyatt), National Institutes of Health (NIH) (PI: TA Einhorn), The Netherlands Organisation for Health Research and Development (ZonMw) (PI: EMM van Lieshout), Sophies Minde Foundation for Orthopaedic Research (PI: L Nordsletten and F Frihagen), and McMaster Surgical Associates (PI: M Bhandari). Dr. Bhandari was also funded, in part, by a Canada Research Chair in Musculoskeletal Trauma which is unrelated to the present study (McMaster University, Hamilton, ON, Canada). The funding sources had no role in design or conduct of the study; the collection, management, analysis, or interpretation of the data; or the preparation, review, or approval of the manuscript. Dr. Mohit Bhandari had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Author Contributions:

Steering Committee: Mohit Bhandari (Chair), PJ Devereaux, Gordon H Guyatt, Thomas A. Einhorn, Lehana Thabane, Emil H. Schemitsch, Ken Korval, Frede Frihagen, Rudolph W. Poolman, Kevin Tetsworth.

Global Methods Centre: Mohit Bhandari (Principal Investigator); Sheila Sprague (Research Program Manager)**;** Marilyn Swinton, Taryn Scott, Paula McKay, Kim Madden (Research Coordination); Diane Heels-Ansdell, (Statistical Analysis); Lisa Buckingham, Aravin Duraikannan (Data Management) **(McMaster University)**

US Methods Centre: Thomas A. Einhorn (Principal Investigator); Heather Silva (Research Coordination) **(Boston University Medical Center)**.

Netherlands Methods Centre: Martin J. Heetveld (Principal Investigator); Rudolf W. Poolman (Co-Principal Investigator), Esther M.M. Van Lieshout (Research Coordination), Paul T.P.W. Burgers (Trial Coordination) **(Erasmus Medical Centre, Rotterdam)**

Central Adjudication Committee: Mohit Bhandari (Chair), Gregory J. Della Rocca, Susan Liew, Thomas A. Einhorn, Rudolph W. Poolman, Robert Haverlag, Martin Heetveld.

Data Safety Monitoring Board (CIHR): John Antoniou (Chair), Tim Ramsay, Earl R. Bogoch, Andrew Trenholm.

Data Safety Monitoring Board (NIH): Stephen Lyman (Chair), Madhu Mazumdar, Kevin J. Bozic, Mark Luborsky, Stuart Goodman, Susan Muray.

HEALTH Investigators:

The following persons participated in the HEALTH trial.

Canada

Foothills Medical Centre - Rob Korley, Richard Buckley, Paul Duffy, Shannon Puloski, Kimberly Carcary, Melissa Lorenzo. *St. Michael's Hospital* - Emil H. Schemitsch, Michael D. McKee, Jeremy A. Hall, Aaron Nauth, Daniel Whelan, Timothy R. Daniels, Earl R. Bogoch, James P Waddell, Henry Ahn, Milena R. Vicente, Jennifer T. Hidy, Melanie T. MacNevin. *Sunnybrook Health Sciences Centre* - Hans Kreder, Terry Axelrod, Richard Jenkinson, Markku Nousiainen, David Stephen, Veronica Wadey, Monica Kunz, Katrine Milner, Ria Cagaanan, Melanie MacNevin. *Vancouver General Hospital* **-** Peter J. O'Brien, Piotr A. Blachut, Henry M. Broekhuyse, Pierre Guy, Kelly A. Lefaivre, Gerard P. Slobogean, Raman Johal, Irene Leung. *Queen Elizabeth II Health Sciences Centre* - Chad Coles, Ross Leighton, C. Glen Richardson, Michael Biddulph, Michael Gross, Michael Dunbar, J. David Amirault, David Alexander, Catherine Coady, Mark Glazebrook, David Johnston, William Oxner, Gerald Reardon, Ivan Wong, Kelly Trask, Shelley MacDonald. *Memorial University of Newfoundland* - Andrew Furey, Craig Stone, Minnie Parsons. *University of British Columbia/Fraser Health Authority -* Trevor Stone, Mauri Zomar, Robert McCormack, Kelly Apostle, Dory Boyer, Farhad Moola, Bertrand Perey, Darius Viskontas, Karyn Moon, Raely Moon. *Hôpital du Sacré-Coeur de Montréal* – Yves Laflamme, Benoit Benoit, Pierre Ranger, Michel Malo, Julio Fernandes, Karine Tardif, Julie Fournier. *Hôpital Maisonneuve-Rosemont* – Pascal André Vendittoli, Vincent Massé, Alain G. Roy, Martin Lavigne, Daniel Lusignan.

United States

Colorado Orthopedic Consultants – Craig Davis, Philip Stull, Stewart Weinerman, Peter Weingarten, Steven Lindenbaum, Michael Hewitt, Rebecca Danielwicz, Janell Baker, *Rubin Institute for Advanced Orthopaedics* – Michael Mont, Donald E. Delanois, Bhaveen Kapadia, Kimona Issa, and Marylou Mullen, *Mayo Clinic* – Andrew Sems, Barb Foreman, *Rothman Institute* – Javad Parvizi, Tiffany Morrison, *Orthopaedic Associates of Hartford* – Courtland Lewis, Stephanie Caminiti, *Boston University Medical Center* – Thomas A Einhorn, Paul Tornetta III, William R Creevy, Heather Silva, Michelle J. Lespasio, Hope Carlisle, *Lahey Clinic* - Andrew Marcantonio, Michael Kain, Lawrence Specht, and John Tilzey, John Garfi, *University of Pennsylvania* – Samir Mehta, John L. Esterhai Jr, Jaimo Ahn, Derek Donegan, Annamarie Horan, Kelly McGinnis, *Emory University School of Medicine* – James Roberson, Thomas Bradbury, Greg Erens, Kyle Webb, *Indiana University* – Brian Mullis, Karl Shively, Andrew Parr, Janos Ertl, Ripley Worman, Mark Webster, Judd Cummings, Valda Frizzell, Molly Moore, *Orthopaedic Associates of Michigan* – Clifford B. Jones, James R. Ringler, Debra L. Sietsema, Jane E. Walker, *Texas Tech University* – Enes Kanlic, Amr Abdelgawad, Juan Shunia, *Mission Hospital Research Institute* – Charles DePaolo, Susan Sutherland, Rachel Alosky, *Duke University Medical Center* – Robert Zura, Maria Manson, *Park Nicollet Institute* – Gregg Strathy, Kathleen Peter, Paul Johnson, and Meaghan Morton, *St. Elizabeth Health Center* - James Shaer, Tyson Schrickel, Barbara Hileman, Marina Hanes, Elisha Chance. *Texas Institute for Hip and Knee Surgery* - E. Matthew Heinrich, David Dodgin, Michele LaBadie, *University of California Irvine* - David Zamorano, Martin Tynan, Ran Schwarzkopf, John A Scolaro, Ranjan Gupta, Samuel Bederman, Nitin Bhatia, Bang Hoang, Douglas Kiester, Neil Jones, Gregory Rafijah, Damon Alavekios, Jason Lee, Akshay Mehta, Steven Schroder, Tom Chao, Vincent

Colin, Phuc (Phil) Dang, Stephen Keun Heng, Gregory Lopez, Samuel Galle, Sohrab Pahlavan, Duy L Phan, Minal Tapadia, Christopher Bui, Nickul Jain, Tyler Moore, Nathan Moroski, Deeba Pourmand. *University of Utah* -Erik N. Kubiak, Jeremy Gililland, David Rothberg, Christopher Peters, Christopher Pelt, Ami R. Stuart, Kirby Corbey. *Marshall University* - Franklin D. Shuler, James Day, Tigran Garabekyan, Felix Cheung, Ali Oliashirazi, Jonathon Salava, Linda Morgan, Timothy Wilson-Byrne, and Mary Beth Cordle.

Netherlands

Amphia Ziekenhuis - Leon H.G.J. Elmans, Joost A.A.M. van den Hout, Adrianus J.P. Joosten, Ad F.A. van Beurden, Stefan B.T. Bolder, Denise Eygendaal, Adrianus F.C.M. Moonen, Rutger C.I. van Geenen, Eric A. Hoebink, Robert Wagenmakers, Wouter van Helden; *Deventer Ziekenhuis* - Hans-Peter W. van Jonbergen, Herbert Roerdink, Joost M. Reuver, Alexander F.W. Barnaart, Elvira R. Flikweert; *Diaconessenhuis Leiden* - Rover Krips, J. Bernard Mullers, Hans Schüller; *Flevoziekenhuis* - Mark L.M. Falke, Frans J. Kurek, Adrianus C.H. Slingerland; *Gelderse Vallei* - Jan P. van Dijk, Wouter H. van Helden; *Gelre Ziekenhuizen* - Hugo W. Bolhuis, Pieter H.J. Bullens, Mike Hogervorst, Karin E . de Kroon, Rob H . Jansen, Ferry Steenstra, Eric E.J. Raven; *IJsselland Ziekenhuis* - W. Peter J. Fontijne, Saskia C. Wiersma, Bastiaan Boetes, Edgar J.T. ten Holder; *Leids Universitair Medisch Centrum* - Huub J.L. van der Heide, Jochem Nagels, Enrike H.M.J. van der Linden-van der Zwaag; *Medisch Centrum Haaglanden* - Stefan B. Keizer, Jan-Willem A. Swen, Peter H.C. den Hollander, Bregje J.W. Thomassen; *Onze Lieve Vrouwe Gasthuis* - Rudolf W. Poolman, Willem Jan Kleyn Molekamp, Frank R.A.J. de Meulemeester, Arthur E.B. Kleipool, Robert Haverlag, Maarten P. Simons, Eduard L.A.R. Mutsaerts; *Ruwaard van Putten Ziekenhuis* - Rob Kooijman, Roelf R. Postema, René J.T.M. Bleker, Harald I.H. Lampe; *Slotervaartziekenhuis* - Lein Schuman, John Cheung, Frank van Bommel, W. Paul C.A. Winia, Daniel Haverkamp, Harm van der Vis; *Spaarne Ziekenhuis* - Peter A. Nolte, Michel P.J. van den Bekerom, Tjitte de Jong, Arthur van Noort, Diederik A. Vergroesen, Bernard G. Schutte *Tergooiziekenhuizen* - Harm M . van der Vis, Lijkele Beimers, Jasper de Vries, Arthur W. Zurcher, G.H. Rob Albers, Maarten Rademakers, Stefan Breugem, Ibo van der Haven, Pieter Jan Damen, Gythe H. Bulstra, Martin M. Campo, Mathijs P. Somford, Daniël Haverkamp.

International

The Alfred - Susan Liew, Harvinder Bedi, Ashley Carr, Andrew Chia, Steve Csongvay, Craig Donohue, Stephen Doig, Elton Edwards, Max Esser, Richard Freeman, Andrew Gong, Doug Li, Russell Miller, Lu Ton, Otis Wang, Ian Young, Adam Dowrick, Zoe Murdoch, Claire Sage. *Oslo University Hospital* – Frede Frihagen, Lars Nordsletten, John Clarke-Jenssen, Geir Hjorthaug, Anne Christine Brekke, Elise Berg Vesterhus, *Ringerike Sykehus Hospital* – Ingunn Skaugrud. *Hospital Universitario Costa del Sol* - Enrique Guerado, Encarnacion Cruz, Juan Ramon Cano. *Hospital Dr. Josep Trueta* - Miguel Angel Froufe, Lluis Marull Serra, Samer Al-dirra, Cristina Martinez. *The Geelong Hospital* - Richard Page, David Bainbridge, Richard Angliss, Ben Miller, Andrew Thomson, Graeme Brown, Simon Williams, Kevin Eng, David Bowyer, John Skelley, Chatar Goyal, Sally Beattie. *Ratandeep Hospital & Research Center, Kanpur*- Pradeep Tripathi, Sandesh Katiyar, Preksha Shukla. *Hospital de la Ribera* – Francisco José Tarazona Santabalbina. Hospital Vall d'Hebron - Ernesto Guerra-Farfan, Jordi Teixidor Serra, Jordi Tomas Hernandez, Marc Aguilar Garcia, Vicente Molero Garcia, Sergi Barrera, Miriam Garrido.

FAITH trial:

Research grants were received from the following: Canadian Institutes of Health Research (CIHR) (PI: Mohit Bhandari); Stichting NutsOhra (PI: Martin J. Heetveld), The Netherlands Organisation for Health Research and Development (PI: Esther M.M. Van Lieshout); Physicians' Services Incorporated (PI: Mohit Bhandari). Funding for the pilot phase of FAITH was supported, in part, by Stryker Inc. Dr. Bhandari was also funded, in part, by a Canada Research Chair in Musculoskeletal Trauma which is unrelated to the present study (McMaster University, Hamilton, ON, Canada). We would also like to acknowledge the support of The County Durham & Tees Valley Comprehensive Local Research Network which operates as part of the National Institute for Health Research Comprehensive Clinical Research Network in England. The funding sources had no role in design or conduct of the study; the collection, management, analysis, or interpretation of the data; or the preparation, review, or approval of the manuscript. Dr. Mohit Bhandari had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Steering Committee: Mohit Bhandari (Chair), Marc Swiontkowski, PJ Devereaux, Gordon Guyatt, Martin J. Heetveld, Kyle Jeray, Susan Liew, Emil H. Schemitsch, Lehana Thabane, Stephen Walter

Global Methods Centre: Mohit Bhandari (Principal Investigator); Sheila Sprague (Research Program Manager)**;** Helena Viveiros, *Paula McKay*, Taryn Scott, Marilyn Swinton, (Research Coordination); Victoria Truong and Kaitlin Koo (Adjudication Coordination); Diane Heels-Ansdell, Qi Zhou (Statistical Analysis); Lisa Buckingham, Aravin Duraikannan (Data

Management); Deborah Maddock, Nicole Simunovic (Grant Management) **(McMaster University)**

US Methods Centre: Marc Swiontkowski (Principal Investigator); Julie Agel (Research Coordination) **(University of Minnesota)**

Netherlands Method Centre: Martin J. Heetveld (Principal Investigator); Esther M.M. Van Lieshout (Research Coordination); Stephanie M. Zielinski (Trial Coordination) **(Erasmus MC, University Medical Center Rotterdam)**

UK Method Centre: Amar Rangan (Principal Investigator), Birgit C. Hanusch, Lucksy Kottam, Rachel Clarkson (Research Coordination) **(The James Cook University Hospital)**

Central Adjudication Committee: Gregory J Della Rocca (Chair), Robert Haverlag, Susan Liew, Gerard Slobogean

Data Safety Monitoring Board: Jeffrey Katz (Chair), Brenda Gillespie, Gail A. Greendale, Pierre Guy, Curtis Hartman, Craig Rubin, James Waddell

FAITH Investigators

The following persons participated in the FAITH trial:

Canada:

Robert McCormack, Kelly Apostle, Dory Boyer, Farhad Moola, Bertrand Perey, Trevor Stone, Darius Viskontas and H. Michael Lemke, Mauri Zomar, Karyn Moon, Raely Moon, Amber Oatt **(University of British Columbia/Fraser Health Authority)**; Richard E. Buckley, Paul Duffy, Robert Korley, Shannon Puloski, Kelly Johnston, James Powell, Kimberly Carcary **(Foothills Medical Centre);** David Sanders, Abdel Lawendy, Christina Tieszer **(London Health Sciences Centre);** David Stephen, Hans Kreder, Richard Jenkinson, Markku Nousiainen, Terry Axelrod, John Murnaghan, Diane Nam, Veronica Wadey, Albert Yee, Katrine Milner, Monica Kunz **(Sunnybrook Health Sciences Centre);** Emil H. Schemitsch, Henry Ahn, Jeremy A. Hall, Michael D. McKee, Daniel B. Whelan, Aaron Nauth, Milena R. Vicente, Lisa M. Wild, Ryan M. Khan, Jennifer T. Hidy **(St. Michael's Hospital);** Chad Coles, Ross Leighton, Michael Biddulph, David Johnston, Mark Glazebrook, David Alexander, Cathy Coady, Michael Dunbar, David Amirault, Michael Gross, William Oxner, Gerald Reardon, Glen Richardson, Andrew Trenholm, Ivan Wong, Kelly Trask, Shelley MacDonald, Gwen Dobbin **(Queen Elizabeth II Health Sciences Centre);** Ryan Bicknell, Jeff Yach, Davide Bardana, Gavin Wood, Mark Harrison, David Yen, Sue Lambert, Fiona Howells, Angela Ward **(Human Mobility Research Centre, Queen's University and Kingston General Hospital);** Paul Zalzal, Heather Brien, V. Naumetz, Brad Weening, Nicole Simunovic **(Oakville Trafalgar Memorial Hospital);** Eugene K. Wai, Steve Papp, Wade T. Gofton, Allen Liew, Stephen P. Kingwell, Garth Johnson, Joseph

O'Neil, Darren M. Roffey, Vivian Borsella **(Ottawa Hospital);** Victoria Avram **(Juravinski Hospital and Cancer Centre)**

United States:

Todd M. Oliver, Vicki Jones **(Boone Hospital Center – Columbia Orthopaedic Group);** Clifford B. Jones, James R. Ringler, Terrence J. Endres, Samuel G. Agnew, Debra L. Sietsema, Jane E. Walker **(Orthopaedic Associates of Michigan);** Kyle J. Jeray, J. Scott Broderick, David R. Goetz, Thomas B. Pace, Thomas M. Schaller, Scott E. Porter, Stephanie L. Tanner, Rebecca G. Snider, Lauren A. Nastoff, Shea A. Bielby **(Greenville Hospital System);** Julie A. Switzer, Peter A. Cole, Sarah A. Anderson, Paul M. Lafferty, Mengnai Li, Thuan V. Ly, Scott B. Marston, Amy L. Foley, Sandy Vang, David M. Wright **(Regions Hospital-University of Minnesota);** Andrew J. Marcantonio, Michael S.H. Kain, Richard Iorio, Lawrence M. Specht, John F. Tilzey, Margaret J. Lobo, John S. Garfi **(Lahey Hospital & Medical Center);** Heather A. Vallier, Andrea Dolenc, Chalitha Robinson **(MetroHealth Medical Center);** Michael J. Prayson, Richard Laughlin, L. Joseph Rubino, Jedediah May, Geoffrey Ryan Rieser, Liz Dulaney-Cripe, Chris Gayton **(Miami Valley Hospital);** James Shaer, Tyson Schrickel, Barbara Hileman **(St. Elizabeth Health Center);** John T. Gorczyca, Jonathan M. Gross, Catherine A. Humphrey, Stephen Kates, Krista Noble, Allison W. McIntyre, Kaili Pecorella **(University of Rochester Medical Center);** Craig A. Davis, Stewart Weinerman, Peter Weingarten, Philip Stull, Stephen Lindenbaum, Michael Hewitt, John Schwappach, Janell K. Baker, Tori Rutherford, Heike Newman, Shane Lieberman, Erin Finn, Kristin Robbins, Meghan Hurley, Lindsey Lyle, Khalis Mitchell, Kieran Browner, Erica Whatley, Krystal Payton, Christina Reeves **(Colorado Orthopedic Consultants);** Lisa K. Cannada, David Karges, Leslie Hill **(St.**

Louis University Hospital); Samir Mehta, John Esterhai, Jaimo Ahn, Annamarie D. Horan, Kelly McGinnis, Christine A. Kaminski, Brynn N. Kowalski **(University of Pennsylvania);** Jonathan P. Keeve, Christopher G. Anderson, Michael D. McDonald, Jodi M. Hoffman **(Northwest Orthopaedic Specialists);** Ivan Tarkin, Peter Siska, Gary Gruen, Andrew Evans, Dana J. Farrell, James Irrgang, Arlene Luther **(University of Pittsburgh Medical Center);** William W. Cross III, Joseph R. Cass, Stephen A. Sems, Michael E. Torchia, Tyson Scrabeck **(Mayo Clinic);** Mark Jenkins, Jules Dumais, Amanda W. Romero **(Texas Tech University Health Sciences Center – Lubbock);** Carlos A. Sagebien, Mark S. Butler, James T. Monica, Patricia Seuffert **(University Orthopaedic Associates, LLC);** Joseph R. Hsu, James Ficke, Michael Charlton, Matthew Napierala, Mary Fan **(US Army Institute of Surgical Research);** Paul Tornetta III, Chadi Tannoury, Hope Carlisle, Heather Silva **(Boston University Medical Center);** Michael Archdeacon, Ryan Finnan, Toan Le, John Wyrick, Shelley Hess **(UC Health/University of Cincinnati Medical Center);** Michael L. Brennan, Robert Probe, Evelyn Kile, Kelli Mills, Lydia Clipper, Michelle Yu, Katie Erwin **(Scott and White Memorial Hospital);** Daniel Horwitz, Kent Strohecker, Teresa K. Swenson **(Geisinger Medical Center);** Andrew H. Schmidt, Jerald R. Westberg **(Hennepin County Medical Center);** Kamran Aurang, Gary Zohman, Brett Peterson, Roger B. Huff **(Kaiser Permanente);** Joseph Baele, Timothy Weber, Matt Edison **(OrthoIndy);** Jessica McBeth **(Santa Clara Valley Medical Center);** Karl Shively, Janos P. Ertl, Brian Mullis, J. Andrew Parr, Ripley Worman, Valda Frizzell, Molly M. Moore, Erin Tobias, Emily Thomas **(Indiana University – Wishard Health Services);** Charles J. DePaolo, Rachel Alosky, Leslie E. Shell, Lynne Hampton, Stephanie Shepard, Tracy Nanney, Claudine Cuento **(Mission Hospital Research Institute);** Robert V. Cantu, Eric R. Henderson, Linda S. Eickhoff **(Dartmouth-Hitchcock Medical Center);** E. Mark Hammerberg, Philip

Stahel, David Hak, Cyril Mauffrey, Douglas Gibula, Hannah Gissel, Corey Henderson **(Denver Health and Hospital Authority);** David P. Zamorano, Martin C. Tynan, Deeba Pourmand, Deanna Lawson **(University of California Irvine Medical Center);** Gregory J. Della Rocca, Brett D. Crist, Yvonne M. Murtha, Linda K. Anderson **(University of Missouri Health Care);** Colleen Linehan, Lindsey Pilling **(Covenant Healthcare of Saginaw);** Courtland G. Lewis, Stephanie Caminiti, Raymond J. Sullivan, Elizabeth Roper **(Hartford Hospital);** William Obremskey, Philip Kregor, Justin E. Richards, Kenya Stringfellow **(Vanderbilt University Medical Center);** Michael P. Dohm, Abby Zellar **(Western Slope Study Group)**

The Netherlands:

Michiel J.M. Segers, Jacco A.C. Zijl, Bart Verhoeven, Anke B. Smits, Jean Paul P.M. de Vries, Bram Fioole, Henk van der Hoeven, Evert B.M. Theunissen, Tammo S. de Vries Reilingh, Lonneke Govaert, Philippe Wittich, Maurits de Brauw, Jan Wille, Peter M.N.Y.M. Go, Ewan D. Ritchie, Ronald N. Wessel, Eric R. Hammacher **(St. Antonius Ziekenhuis);** Martin J. Heetveld, Gijs A. Visser, Heyn Stockmann, Rob Silvis, Jaap P. Snellen, Bram Rijbroek, Joris J.G. Scheepers, Erik G.J. Vermeulen, Michiel P.C. Siroen, Ronald Vuylsteke, Hans L.F. Brom, Herman Rijna **(Kennemer Gasthuis);** Piet A.R. de Rijcke, Cees L. Koppert, Steven E. Buijk, Richard P.R. Groenendijk, Imro Dawson, Geert W.M. Tetteroo, Milko M.M. Bruijninckx, Pascal G. Doornebosch, Eelco J.R. de Graaf **(IJsselland Ziekenhuis);** Maarten van der Elst, Carmen C. van der Pol, Martijne van ’t Riet, Tom M. Karsten, Mark R. de Vries, Laurents P.S. Stassen, Niels W.L. Schep, G. Ben Schmidt, W.H. Hoffman **(Reinier de Graaf Gasthuis);** Rudolf W. Poolman, Maarten P. Simons, Frank H.W.M. van der Heijden, W. Jaap Willems, Frank R.A.J. de Meulemeester, Cor P. van der Hart, Kahn Turckan, Sebastiaan Festen, Frank de Nies, Robert

Haverlag, Nico J.M. Out, Jan Bosma **(Onze Lieve Vrouwe Gasthuis);** Albert van Kampen, Jan Biert, Arie B. van Vugt, Michael J.R. Edwards, Taco J. Blokhuis, Jan Paul M. Frölke, Leo M.G. Geeraedts, Jean W.M. Gardeniers, Edward T.C.H. Tan, Lodewijk M.S.J. Poelhekke, Maarten C. de Waal Malefijt, Bart Schreurs **(University Medical Center St. Radboud);** Gert R Roukema, Hong A. Josaputra, Paul Keller, Peter D. de Rooij, Hans Kuiken, Han Boxma, Berry I. Cleffken, Ronald Liem **(Maasstad Ziekenhuis);** Steven J. Rhemrev, Coks H.R. Bosman, Alexander de Mol van Otterloo, Jochem Hoogendoorn, Alexander C. de Vries, Sven A.G. Meylaerts **(Medisch Centrum Haaglanden);** Michiel H.J. Verhofstad, Joost Meijer, Teun van Egmond, Frank H.W.M. van der Heijden, Igor van der Brand **(St. Elisabeth Ziekenhuis);** Peter Patka, Martin G. Eversdijk, Rolf Peters, Dennis Den Hartog, Oscar J.F. Van Waes, Pim Oprel **(Erasmus MC, University Medical Center Rotterdam);** Harm M van der Vis, Martin Campo, Ronald Verhagen, G.H. Robert Albers, Arthur W. Zurcher **(Tergooi Ziekenhuizen);** Rogier K.J. Simmermacher, Jeroen van Mulken, Karlijn van Wessem, Taco J. Blokhuis, Steven M. van Gaalen, Luke P.H. Leenen **(University Medical Center Utrecht);** Maarten W.G.A. Bronkhorst, Onno R. Guicherit **(Bronovo Ziekenhuis);** J. Carel Goslings, Robert Haverlag, Kees Jan Ponsen **(Academic Medical Center)**

International:

Mahesh Bhatia, Vinod Arora, Vivek Tyagi **(RLB Hospital and Research Centre, India);** Susan Liew, Harvinder Bedi, Ashley Carr, Hamish Curry, Andrew Chia, Steve Csongvay, Craig Donohue, Stephen Doig, Elton Edwards, Greg Etherington, Max Esser, Andrew Gong, Arvind Jain, Doug Li, Russell Miller, Ash Moaveni, Matthias Russ, Lu Ton, Otis Wang, Adam Dowrick, Zoe Murdoch, Claire Sage **(The Alfred, Australia);** Frede Frihagen, John Clarke-Jenssen, Geir

Hjorthaug, Torben Ianssen, Asgeir Amundsen, Jan Egil Brattgjerd, Tor Borch, Berthe Bøe, Bernhard Flatøy, Sondre Hasselund, Knut Jørgen Haug, Kim Hemlock, Tor Magne Hoseth, Geir Jomaas, Thomas Kibsgård, Tarjei Lona, Gilbert Moatshe, Oliver Müller, Marius Molund, Tor Nicolaisen, Fredrik Nilsen, Jonas Rydinge, Morten Smedsrud, Are Stødle, Axel Trommer, Stein Ugland, Anders Karlsten, Guri Ekås, Elise Berg Vesterhus, Anne Christine Brekke **(Oslo University Hospital, Norway);** Ajay Gupta, Neeraj Jain, Farah Khan **(Nirmal Hospital, India);** Ateet Sharma, Amir Sanghavi, Mittal Trivedi **(Satellite Orthopaedic Hospital and Research Centre, India);** Anil Rai, Subash, Kamal Rai **(Highway Hospital, India);** Vineet Yadav, Sanjay Singh, Kamal Rai **(Popular Hospital, India);** *Kevin Tetsworth,* Geoff Donald, Patrick Weinrauch, Paul Pincus, Steven Yang, Brett Halliday, Trevor Gervais, Michael Holt, Annette Flynn **(Royal Brisbane and Women's Hospital, Australia);** Amal Shankar Prasad, Vimlesh Mishra **(Madhuraj Nursing Home, India);** D.C. Sundaresh, Angshuman Khanna **(M.S. Rammaiah Medical College & Hospital, India);** Joe Joseph Cherian, Davy J Olakkengil, Gaurav Sharma **(St John's Medical College Hospital, India**); *Marinis Pirpiris,* David Love, Andrew Bucknill, Richard J Farrugia **(Royal Melbourne Hospital, Australia);** Hans-Christoph Pape, Matthias Knobe, Roman Pfeifer **(University of Aachen Medical Center, Germany)**; *Peter Hull, Sophie Lewis, Simone Evans* **(*Cambridge University Hospitals, England)*;** Rajesh Nanda, Rajanikanth Logishetty, Sanjeev Anand, Carol Bowler **(University Hospital of North Tees, England);** Akhil Dadi, Naveen Palla, Utsav Ganguly **(Sunshine Hospital, India);** B. Sachidananda Rai, Janakiraman Rajakumar **(Unity Health Complex, India);** Andrew Jennings, Graham Chuter, Glynis Rose, Gillian Horner **(University Hospital of North Durham and Darlington Memorial Hospital, England);** Callum Clark, Kate Eke **(Wexham Park**

Hospital, England); Mike Reed, Chris Herriott, Christine Dobb **(Northumbria Healthcare NHS Foundation Trust, England)**

Appendix 2. Hypotheses for evaluating the construct validity of the WOMAC Instrument

1. WOMAC Pain was expected to correlate strongly (r >0.6) with SF-12 pain and EQ-Pain since they were expected to measure the same construct.
2. A moderate correlation (r <0.6) was therefore predicted between WOMAC Pain and all other scores since they assess different or less specific constructs.
3. A weak (r <0.3) correlation was expected for the specific WOMAC Stiffness and all other outcomes.
4. WOMAC Function was predicted to correlate strongly with SF-12 PCS, SF-12 pain, SF-12 Total score, EQ pain, EQ-US, and EQ-VAS.
5. A moderately to weak correlation was expected for WOMAC Function with SF-12 MCS.
6. WOMAC Total score was predicted to have a moderate correlation with SF-12 MCS and a strong correlation all other outcomes.