

Bone & Joint Research



Supplementary Material

10.1302/2046-3758.126.BJR-2022-0280.R1

Supplementary Material 1. Search blocks, filters, and search results identifying validated outcome measurement instruments for each core outcome in the open lower limb fracture population.

| Search blocks and filters |
|---|
| 1. The Biomedische Informatie Group (BMI). A group of Dutch medical information specialists have compiled a series of open access search strategy building blocks for common constructs, including for walking or gait, return to work, pain and quality of life. Use of BMI search blocks is recommended by COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN). ¹ |
| 2. The University of Oxford, PROM Group search filter for identifying patient-reported outcome measures (PROMs). ² |
| 3. The COSMIN highly sensitive and validated search filter for identifying studies on measurement properties. ³ |

Search blocks and filters were translated for use on the Ovid search platform where necessary.

Search 1. Walking, gait and mobility

Database(s): Embase 1974 to present, Medline (Ovid MEDLINE Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE Daily and Ovid MEDLINE) 1946 to present.

Date searched: 17 July 2019

| # | Searches | Results |
|----|---|-----------|
| 1 | (instrumentation or methods).fs. | 3,915,143 |
| 2 | (Validation Studies or Comparative Study).pt. | 1,911,035 |
| 3 | exp Psychometrics/ | 157,647 |
| 4 | psychometr*.ti,ab. | 94,968 |
| 5 | (clinimetr* or clinometr*).tw. | 2,429 |
| 6 | outcome assessment.ti,ab. | 8,425 |
| 7 | outcome measure*.tw. | 480,235 |
| 8 | exp Observer Variation/ | 60,439 |
| 9 | observer variation.ti,ab. | 2,503 |
| 10 | exp Health Status Indicators/ | 313,168 |
| 11 | exp Reproducibility of Results/ | 585,617 |

| | | |
|----|--|------------|
| 12 | reproducib*.ti,ab. | 344,581 |
| 13 | exp Discriminant Analysis/ | 27,755 |
| 14 | (reliab* or unreliab* or valid* or coefficient or homogeneity or homogeneous or internal consistency).ti,ab. | 3,001,053 |
| 15 | (cronbach* and (alpha or alphas)).ti,ab. | 45,698 |
| 16 | (item and (correlation* or selection* or reduction*)).ti,ab. | 48,619 |
| 17 | (agreement or precision or imprecision or precise values or test-retest).ti,ab. | 857,841 |
| 18 | (test and retest).ti,ab. | 55,926 |
| 19 | (reliab* and (test or retest)).ti,ab. | 195,247 |
| 20 | (stability or interrater or inter-rater or intrarater or intra-rater or intertester or inter-tester or intratester or intra-tester or interobserver or inter-observer or intraobserver or intraobserver or intertechnician or inter-technician or intratechnician or intra-technician or interexaminer or inter-examiner or intraexaminer or intra-examiner or interassay or interassay or intraassay or intra-assay or interindividual or inter-individual or intraindividual or intra-individual or interparticipant or inter-participant or intraparticipant or intra-participant or kappa or kappas or repeatab*).ti,ab. | 1,226,959 |
| 21 | ((replicab* or repeated) and (measure or measures or findings or result or results or test or tests)).ti,ab. | 455,125 |
| 22 | (generaliza* or generalisa* or concordance).ti,ab. | 192,092 |
| 23 | (intraclass and correlation*).ti,ab. | 50,165 |
| 24 | (discriminative or known group or factor analysis or factor analyses or dimension* or subscale*).ti,ab. | 1,280,941 |
| 25 | (multitrait and scaling and (analysis or analyses)).ti,ab. | 288 |
| 26 | (item discriminant or interscale correlation* or error or errors or individual variability).ti,ab. | 647,123 |
| 27 | (variability and (analysis or values)).ti,ab. | 221,225 |
| 28 | (uncertainty and (measurement or measuring)).ti,ab. | 15,307 |
| 29 | (standard error of measurement or sensitiv* or responsive*).ti,ab. | 3,350,311 |
| 30 | ((minimal or minimally or clinical or clinically) and (important or significant or detectable) and (change or difference)).ti,ab. | 588,727 |
| 31 | (small* and (real or detectable) and (change or difference)).ti,ab. | 17,827 |
| 32 | (meaningful change or ceiling effect or floor effect or Item response model or IRT or Rasch or Differential item functioning or DIF or computer adaptive testing or item bank or cross-cultural equivalence).ti,ab. | 29,251 |
| 33 | 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 | 14,426,731 |
| 34 | Gait Analysis/ or Gait Ataxia/ or Gait/ or Gait Apraxia/ or Gait Disorders, Neurologic/ or Gait.mp. | 142,866 |
| 35 | gait.ti,ab. | 110,728 |
| 36 | gaits.ti,ab. | 2,365 |
| 37 | Walking Speed/ or walking.mp. or Walking/ | 199,552 |
| 38 | walk*.ti,ab. | 259,251 |
| 39 | ambulation.mp. | 28,031 |
| 40 | ambulat*.ti,ab. | 207,958 |

| | | |
|----|---|-----------|
| 41 | mobility.ti,ab. | 284,115 |
| 42 | 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 | 833,530 |
| 43 | Fractures, Open/ | 9,000 |
| 44 | ((open or compound or severe* or mangle*) adj3 (fracture* or break*)).ti,ab. | 24,351 |
| 45 | 43 or 44 | 27,949 |
| 46 | exp lower extremity/ or exp buttocks/ or exp foot/ or exp hip/ or exp knee/ or exp leg/ or exp thigh/ | 525,089 |
| 47 | "lower extremit*".ti,ab. | 120,015 |
| 48 | "lower limb*".ti,ab. | 113,174 |
| 49 | (leg or legs).ti,ab. | 253,419 |
| 50 | (foot or feet).ti,ab. | 247,754 |
| 51 | thigh*.ti,ab. | 66,553 |
| 52 | ankle*.ti,ab. | 130,977 |
| 53 | (hip or hips).ti,ab. | 308,054 |
| 54 | exp Leg Bones/ | 192,599 |
| 55 | knee*.ti,ab. | 317,204 |
| 56 | femur*.ti,ab. | 113,358 |
| 57 | tibia*.ti,ab. | 181,496 |
| 58 | patella*.ti,ab. | 42,515 |
| 59 | talus*.ti,ab. | 9,624 |
| 60 | fibula*.ti,ab. | 25,859 |
| 61 | calcaneus*.ti,ab. | 11,331 |
| 62 | navicular*.ti,ab. | 4,913 |
| 63 | cuneiform*.ti,ab. | 2,978 |
| 64 | cuboid*.ti,ab. | 11,134 |
| 65 | metatarsal*.ti,ab. | 18,968 |
| 66 | exp Foot Bones/ | 36,188 |
| 67 | phalan*.ti,ab. | 23,223 |
| 68 | Leg/ | 133,994 |
| 69 | (toe or toes).ti,ab. | 44,759 |
| 70 | exp Toes/ | 25,107 |
| 71 | pilon*.ti,ab. | 5,342 |
| 72 | 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 55 or 56 or 57 or 58 or 60 or 59 or 61 or 62 or 63 or 64 or 65 or 66 or 67 or 68 or 69 or 70 or 71 | 1,698,001 |
| 73 | 45 and 72 | 13,030 |
| 74 | 33 and 42 and 73 | 372 |

Search 2: Return to life roles

Database(s): Embase 1974 to present, Medline (Ovid MEDLINE Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE Daily and Ovid MEDLINE) 1946 to present.

Date searched: 17 July 2019

| # | Searches | Results |
|----|--|-----------|
| 1 | (instrumentation or methods).fs. | 3,917,553 |
| 2 | (Validation Studies or Comparative Study).pt. | 1,911,383 |
| 3 | exp Psychometrics/ | 157,717 |
| 4 | psychometr*.ti,ab. | 95,052 |
| 5 | (clinimetr* or clinometr*).tw. | 2,434 |
| 6 | outcome assessment.ti,ab. | 8,439 |
| 7 | outcome measure*.tw. | 480,713 |
| 8 | exp Observer Variation/ | 60,465 |
| 9 | observer variation.ti,ab. | 2,505 |
| 10 | exp Health Status Indicators/ | 313,309 |
| 11 | exp Reproducibility of Results/ | 585,822 |
| 12 | reproducib*.ti,ab. | 344,805 |
| 13 | exp Discriminant Analysis/ | 27,787 |
| 14 | (reliab* or unreliab* or valid* or coefficient or homogeneity or homogeneous or internal consistency).ti,ab. | 3,004,087 |
| 15 | (cronbach* and (alpha or alphas)).ti,ab. | 45,760 |
| 16 | (item and (correlation* or selection* or reduction*)).ti,ab. | 48,665 |
| 17 | (agreement or precision or imprecision or precise values or test-retest).ti,ab. | 858,613 |
| 18 | (test and retest).ti,ab. | 55,988 |
| 19 | (reliab* and (test or retest)).ti,ab. | 195,420 |
| 20 | (stability or interrater or inter-rater or intrarater or intra-rater or intertester or inter-tester or intratester or intra-tester or interobserver or inter-observer or intraobserver or intraobserver or intertechnician or inter-technician or intratechnician or intra-technician or interexaminer or inter-examiner or intraexaminer or intra-examiner or interassay or interassay or intraassay or intra-assay or interindividual or inter-individual or intraindividual or intra-individual or interparticipant or inter-participant or intraparticipant or intra-participant or kappa or kappas or repeatab*).ti,ab. | 1,228,152 |
| 21 | ((replicab* or repeated) and (measure or measures or findings or result or results or test or tests)).ti,ab. | 455,544 |
| 22 | (generaliza* or generalisa* or concordance).ti,ab. | 192,314 |
| 23 | (intraclass and correlation*).ti,ab. | 50,237 |
| 24 | (discriminative or known group or factor analysis or factor analyses or dimension* or subscale*).ti,ab. | 1,282,060 |
| 25 | (multitrait and scaling and (analysis or analyses)).ti,ab. | 288 |
| 26 | (item discriminant or interscale correlation* or error or errors or individual variability).ti,ab. | 647,775 |
| 27 | (variability and (analysis or values)).ti,ab. | 221,443 |
| 28 | (uncertainty and (measurement or measuring)).ti,ab. | 15,317 |
| 29 | (standard error of measurement or sensitiv* or responsive*).ti,ab. | 3,352,813 |
| 30 | ((minimal or minimally or clinical or clinically) and (important or significant or detectable) and (change or difference)).ti,ab. | 589,426 |
| 31 | (small* and (real or detectable) and (change or difference)).ti,ab. | 17,845 |

| | | |
|----|--|------------|
| 32 | (meaningful change or ceiling effect or floor effect or Item response model or IRT or Rasch or Differential item functioning or DIF or computer adaptive testing or item bank or cross-cultural equivalence).ti,ab. | 29,320 |
| 33 | 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 | 14,437,782 |
| 34 | Musculoskeletal Pain/ or Pain Perception/ or Complex Regional Pain Syndromes/ or Pelvic Pain/ or Back Pain/ or Pain Insensitivity, Congenital/ or Pain, Postoperative/ or Acute Pain/ or Pain, Intractable/ or Abdominal Pain/ or Neck Pain/ or Pain Clinics/ or Pain, Procedural/ or pain*.mp. or Shoulder Pain/ or Patellofemoral Pain Syndrome/ or Chronic Pain/ or Visceral Pain/ or Breakthrough Pain/ or Labor Pain/ or Myofascial Pain Syndromes/ or Pain/ or Nociceptive Pain/ or Pain Measurement/ or Facial Pain/ or Cancer Pain/ or Pain Threshold/ or Low Back Pain/ or Pain Management/ or Chest Pain/ | 2,034,319 |
| 35 | pain*.ti,ab. | 1,588,108 |
| 36 | ache*.ti,ab. | 44,628 |
| 37 | pain measurement.mp. or Pain Measurement/ | 90,470 |
| 38 | Hyperalgesia.mp. or Hyperalgesia/ | 40,478 |
| 39 | Hyperalges*.ti,ab. | 30,912 |
| 40 | allodyni*.ti,ab. | 19,866 |
| 41 | pain perception.mp. or Pain Perception/ | 45,455 |
| 42 | nocicepti*.ti,ab. | 67,398 |
| 43 | vas.ti,ab. | 113,899 |
| 44 | visual analog scale*.ti,ab. | 57,639 |
| 45 | 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 | 2,152,107 |
| 46 | (HR-PRO or HRPRO or HRQL or HRQoL or QL or QoL).ti,ab. or quality of life.mp. or (health index* or health indices or health profile*).ti,ab. or health status.mp. or ((patient or self or child or parent or carer or proxy) adj (appraisal* or appraised or report or reported or reporting or rated or rating* or based or assessed or assessment*)).ti,ab. or ((disability or function or functional or functions or subjective or utility or utilities or wellbeing or well being) adj2 (index or indices or instrument or instruments or measure or measures or questionnaire* or profile or profiles or scale or scales or score or scores or status or survey or surveys)).ti,ab. | 1,716,830 |
| 47 | Fractures, Open/ | 9,003 |
| 48 | ((open or compound or severe* or mangle*) adj3 (fracture* or break*)).ti,ab. | 24,372 |
| 49 | 47 or 48 | 27,970 |
| 50 | exp lower extremity/ or exp buttocks/ or exp foot/ or exp hip/ or exp knee/ or exp leg/ or exp thigh/ | 525,454 |
| 51 | "lower extremit*".ti,ab. | 120,121 |
| 52 | "lower limb*".ti,ab. | 113,286 |
| 53 | (leg or legs).ti,ab. | 253,578 |
| 54 | (foot or feet).ti,ab. | 247,927 |
| 55 | thigh*.ti,ab. | 66,607 |
| 56 | ankle*.ti,ab. | 131,066 |
| 57 | (hip or hips).ti,ab. | 308,309 |
| 58 | exp Leg Bones/ | 192,712 |

| | | |
|----|--|-----------|
| 59 | knee*.ti,ab. | 317,512 |
| 60 | femur*.ti,ab. | 113,436 |
| 61 | tibia*.ti,ab. | 181,615 |
| 62 | patella*.ti,ab. | 42,544 |
| 63 | talus*.ti,ab. | 9,625 |
| 64 | fibula*.ti,ab. | 25,875 |
| 65 | calcaneus*.ti,ab. | 11,336 |
| 66 | navicular*.ti,ab. | 4,914 |
| 67 | cuneiform*.ti,ab. | 2,979 |
| 68 | cuboid*.ti,ab. | 11,139 |
| 69 | metatarsal*.ti,ab. | 18,973 |
| 70 | exp Foot Bones/ | 36,198 |
| 71 | phalan*.ti,ab. | 23,235 |
| 72 | Leg/ | 134,032 |
| 73 | (toe or toes).ti,ab. | 44,792 |
| 74 | exp Toes/ | 25,123 |
| 75 | pilon*.ti,ab. | 5,346 |
| 76 | 50 or 51 or 52 or 53 or 54 or 55 or 56 or 57 or 58 or 59 or 60 or 61 or 62 or 64 or 63 or 65 or 66 or 67 or 68 or 69 or 70 or 71 or 72 or 73 or 74 or 75 | 1,699,234 |
| 77 | 49 and 76 | 13,043 |
| 78 | 33 and 45 and 46 and 77 | 111 |

Search 3. Pain of discomfort

Database(s): Embase 1974 to present, Medline (Ovid MEDLINE Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE Daily and Ovid MEDLINE) 1946 to present.

Date searched: 17 July 2019

| # | Searches | Results |
|----|--|-----------|
| 1 | (instrumentation or methods).fs. | 3,917,553 |
| 2 | (Validation Studies or Comparative Study).pt. | 1,911,383 |
| 3 | exp Psychometrics/ | 157,717 |
| 4 | psychometr*.ti,ab. | 95,052 |
| 5 | (clinimetr* or clinometr*).tw. | 2,434 |
| 6 | outcome assessment.ti,ab. | 8,439 |
| 7 | outcome measure*.tw. | 480,713 |
| 8 | exp Observer Variation/ | 60,465 |
| 9 | observer variation.ti,ab. | 2,505 |
| 10 | exp Health Status Indicators/ | 313,309 |
| 11 | exp Reproducibility of Results/ | 585,822 |
| 12 | reproducib*.ti,ab. | 344,805 |
| 13 | exp Discriminant Analysis/ | 27,787 |
| 14 | (reliab* or unreliab* or valid* or coefficient or homogeneity or homogeneous or internal consistency).ti,ab. | 3,004,087 |

| | | |
|----|--|------------|
| 15 | (cronbach* and (alpha or alphas)).ti,ab. | 45,760 |
| 16 | (item and (correlation* or selection* or reduction*)).ti,ab. | 48,665 |
| 17 | (agreement or precision or imprecision or precise values or test-retest).ti,ab. | 858,613 |
| 18 | (test and retest).ti,ab. | 55,988 |
| 19 | (reliab* and (test or retest)).ti,ab. | 195,420 |
| 20 | (stability or interrater or inter-rater or intrarater or intra-rater or intertester or inter-tester or intratester or intra-tester or interobserver or inter-observer or intraobserver or intraobserver or intertechnician or inter-technician or intratechnician or intra-technician or interexaminer or inter-examiner or intraexaminer or intra-examiner or interassay or interassay or intraassay or intra-assay or interindividual or inter-individual or intraindividual or intra-individual or interparticipant or inter-participant or intraparticipant or intra-participant or kappa or kappas or repeatab*).ti,ab. | 1,228,152 |
| 21 | ((replicab* or repeated) and (measure or measures or findings or result or results or test or tests)).ti,ab. | 455,544 |
| 22 | (generaliza* or generalisa* or concordance).ti,ab. | 192,314 |
| 23 | (intraclass and correlation*).ti,ab. | 50,237 |
| 24 | (discriminative or known group or factor analysis or factor analyses or dimension* or subscale*).ti,ab. | 1,282,060 |
| 25 | (multitrait and scaling and (analysis or analyses)).ti,ab. | 288 |
| 26 | (item discriminant or interscale correlation* or error or errors or individual variability).ti,ab. | 647,775 |
| 27 | (variability and (analysis or values)).ti,ab. | 221,443 |
| 28 | (uncertainty and (measurement or measuring)).ti,ab. | 15,317 |
| 29 | (standard error of measurement or sensitiv* or responsive*).ti,ab. | 335,2813 |
| 30 | ((minimal or minimally or clinical or clinically) and (important or significant or detectable) and (change or difference)).ti,ab. | 589,426 |
| 31 | (small* and (real or detectable) and (change or difference)).ti,ab. | 17,845 |
| 32 | (meaningful change or ceiling effect or floor effect or Item response model or IRT or Rasch or Differential item functioning or DIF or computer adaptive testing or item bank or cross-cultural equivalence).ti,ab. | 29,320 |
| 33 | 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 | 14,437,782 |
| 34 | Musculoskeletal Pain/ or Pain Perception/ or Complex Regional Pain Syndromes/ or Pelvic Pain/ or Back Pain/ or Pain Insensitivity, Congenital/ or Pain, Postoperative/ or Acute Pain/ or Pain, Intractable/ or Abdominal Pain/ or Neck Pain/ or Pain Clinics/ or Pain, Procedural/ or pain*.mp. or Shoulder Pain/ or Patellofemoral Pain Syndrome/ or Chronic Pain/ or Visceral Pain/ or Breakthrough Pain/ or Labor Pain/ or Myofascial Pain Syndromes/ or Pain/ or Nociceptive Pain/ or Pain Measurement/ or Facial Pain/ or Cancer Pain/ or Pain Threshold/ or Low Back Pain/ or Pain Management/ or Chest Pain/ | 2,034,319 |
| 35 | pain*.ti,ab. | 1,588,108 |
| 36 | ache*.ti,ab. | 44,628 |
| 37 | pain measurement.mp. or Pain Measurement/ | 90,470 |
| 38 | Hyperalgesia.mp. or Hyperalgesia/ | 40,478 |
| 39 | Hyperalges*.ti,ab. | 30,912 |

| | | |
|----|--|-----------|
| 40 | allodyni*.ti,ab. | 19,866 |
| 41 | pain perception.mp. or Pain Perception/ | 45,455 |
| 42 | nocicepti*.ti,ab. | 67,398 |
| 43 | vas.ti,ab. | 113,899 |
| 44 | visual analog scale*.ti,ab. | 57,639 |
| 45 | 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 | 2,152,107 |
| 46 | (HR-PRO or HRPRO or HRQL or HRQoL or QL or QoL).ti,ab. or quality of life.mp. or (health index* or health indices or health profile*).ti,ab. or health status.mp. or ((patient or self or child or parent or carer or proxy) adj (appraisal* or appraised or report or reported or reporting or rated or rating* or based or assessed or assessment*)).ti,ab. or ((disability or function or functional or functions or subjective or utility or utilities or wellbeing or well being) adj2 (index or indices or instrument or instruments or measure or measures or questionnaire* or profile or profiles or scale or scales or score or scores or status or survey or surveys)).ti,ab. | 1,716,830 |
| 47 | Fractures, Open/ | 9,003 |
| 48 | ((open or compound or severe* or mangle*) adj3 (fracture* or break*)).ti,ab. | 24,372 |
| 49 | 47 or 48 | 27,970 |
| 50 | exp lower extremity/ or exp buttocks/ or exp foot/ or exp hip/ or exp knee/ or exp leg/ or exp thigh/ | 525,454 |
| 51 | "lower extremit*".ti,ab. | 120,121 |
| 52 | "lower limb*".ti,ab. | 113,286 |
| 53 | (leg or legs).ti,ab. | 253,578 |
| 54 | (foot or feet).ti,ab. | 247,927 |
| 55 | thigh*.ti,ab. | 66,607 |
| 56 | ankle*.ti,ab. | 131,066 |
| 57 | (hip or hips).ti,ab. | 308,309 |
| 58 | exp Leg Bones/ | 192,712 |
| 59 | knee*.ti,ab. | 317,512 |
| 60 | femur*.ti,ab. | 113,436 |
| 61 | tibia*.ti,ab. | 181,615 |
| 62 | patella*.ti,ab. | 42,544 |
| 63 | talus*.ti,ab. | 9,625 |
| 64 | fibula*.ti,ab. | 25,875 |
| 65 | calcaneus*.ti,ab. | 11,336 |
| 66 | navicular*.ti,ab. | 4,914 |
| 67 | cuneiform*.ti,ab. | 2,979 |
| 68 | cuboid*.ti,ab. | 11,139 |
| 69 | metatarsal*.ti,ab. | 18,973 |
| 70 | exp Foot Bones/ | 36,198 |
| 71 | phalan*.ti,ab. | 23,235 |
| 72 | Leg/ | 134,032 |
| 73 | (toe or toes).ti,ab. | 44,792 |
| 74 | exp Toes/ | 25,123 |
| 75 | pilon*.ti,ab. | 5,346 |

| | | |
|----|--|-----------|
| 76 | 50 or 51 or 52 or 53 or 54 or 55 or 56 or 57 or 58 or 59 or 60 or 61 or 62 or 64 or 63 or 65 or 66 or 67 or 68 or 69 or 70 or 71 or 72 or 73 or 74 or 75 | 1,699,234 |
| 77 | 49 and 76 | 13,043 |
| 78 | 33 and 45 and 46 and 77 | 111 |

Search 4. Quality of life

Database(s): Embase 1974 to present, Medline (Ovid MEDLINE Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE Daily and Ovid MEDLINE) 1946 to present.

Date searched: 17 July 2019

| # | Searches | Results |
|----|--|-----------|
| 1 | (instrumentation or methods).fs. | 3,917,553 |
| 2 | (Validation Studies or Comparative Study).pt. | 1,911,383 |
| 3 | exp Psychometrics/ | 157,717 |
| 4 | psychometr*.ti,ab. | 95,052 |
| 5 | (clinimetr* or clinometr*).tw. | 2,434 |
| 6 | outcome assessment.ti,ab. | 8,439 |
| 7 | outcome measure*.tw. | 480,713 |
| 8 | exp Observer Variation/ | 60,465 |
| 9 | observer variation.ti,ab. | 2,505 |
| 10 | exp Health Status Indicators/ | 313,309 |
| 11 | exp Reproducibility of Results/ | 585,822 |
| 12 | reproducib*.ti,ab. | 344,805 |
| 13 | exp Discriminant Analysis/ | 27,787 |
| 14 | (reliab* or unreliab* or valid* or coefficient or homogeneity or homogeneous or internal consistency).ti,ab. | 3,004,087 |
| 15 | (cronbach* and (alpha or alphas)).ti,ab. | 45,760 |
| 16 | (item and (correlation* or selection* or reduction*)).ti,ab. | 48,665 |
| 17 | (agreement or precision or imprecision or precise values or test-retest).ti,ab. | 858,613 |
| 18 | (test and retest).ti,ab. | 55,988 |
| 19 | (reliab* and (test or retest)).ti,ab. | 195,420 |
| 20 | (stability or interrater or inter-rater or intrarater or intra-rater or intertester or inter-tester or intratester or intra-tester or interobserver or inter-observer or intraobserver or intraobserver or intertechnician or inter-technician or intratechnician or intra-technician or interexaminer or inter-examiner or intraexaminer or intra-examiner or interassay or interassay or intraassay or intra-assay or interindividual or inter-individual or intraindividual or intra-individual or interparticipant or inter-participant or intraparticipant or intra-participant or kappa or kappas or repeatab*).ti,ab. | 1,228,152 |
| 21 | ((replicab* or repeated) and (measure or measures or findings or result or results or test or tests)).ti,ab. | 455,544 |
| 22 | (generaliza* or generalisa* or concordance).ti,ab. | 192,314 |
| 23 | (intraclass and correlation*).ti,ab. | 50,237 |
| 24 | (discriminative or known group or factor analysis or factor analyses or dimension* or subscale*).ti,ab. | 1,282,060 |

| | | |
|----|--|------------|
| 25 | (multitrait and scaling and (analysis or analyses)).ti,ab. | 288 |
| 26 | (item discriminant or interscale correlation* or error or errors or individual variability).ti,ab. | 647,775 |
| 27 | (variability and (analysis or values)).ti,ab. | 221,443 |
| 28 | (uncertainty and (measurement or measuring)).ti,ab. | 15,317 |
| 29 | (standard error of measurement or sensitiv* or responsive*).ti,ab. | 3,352,813 |
| 30 | ((minimal or minimally or clinical or clinically) and (important or significant or detectable) and (change or difference)).ti,ab. | 589,426 |
| 31 | (small* and (real or detectable) and (change or difference)).ti,ab. | 17,845 |
| 32 | (meaningful change or ceiling effect or floor effect or Item response model or IRT or Rasch or Differential item functioning or DIF or computer adaptive testing or item bank or cross-cultural equivalence).ti,ab. | 29,320 |
| 33 | 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 | 14,437,782 |
| 34 | (HR-PRO or HRPRO or HRQL or HRQoL or QL or QoL).ti,ab. or quality of life.mp. or (health index* or health indices or health profile*).ti,ab. or health status.mp. or ((patient or self or child or parent or carer or proxy) adj (appraisal* or appraised or report or reported or reporting or rated or rating* or based or assessed or assessment*)).ti,ab. or ((disability or function or functional or functions or subjective or utility or utilities or wellbeing or well being) adj2 (index or indices or instrument or instruments or measure or measures or questionnaire* or profile or profiles or scale or scales or score or scores or status or survey or surveys)).ti,ab. | 1,716,830 |
| 35 | Fractures, Open/ | 9,003 |
| 36 | ((open or compound or severe* or mangle*) adj3 (fracture* or break*)).ti,ab. | 24,372 |
| 37 | 35 or 36 | 27,970 |
| 38 | exp lower extremity/ or exp buttocks/ or exp foot/ or exp hip/ or exp knee/ or exp leg/ or exp thigh/ | 525,454 |
| 39 | "lower extremit*".ti,ab. | 120,121 |
| 40 | "lower limb*".ti,ab. | 113,286 |
| 41 | (leg or legs).ti,ab. | 253,578 |
| 42 | (foot or feet).ti,ab. | 247,927 |
| 43 | thigh*.ti,ab. | 66,607 |
| 44 | ankle*.ti,ab. | 131,066 |
| 45 | (hip or hips).ti,ab. | 308,309 |
| 46 | exp Leg Bones/ | 192,712 |
| 47 | knee*.ti,ab. | 317,512 |
| 48 | femur*.ti,ab. | 113,436 |
| 49 | tibia*.ti,ab. | 181,615 |
| 50 | patella*.ti,ab. | 42,544 |
| 51 | talus*.ti,ab. | 9,625 |
| 52 | fibula*.ti,ab. | 25,875 |
| 53 | calcaneus*.ti,ab. | 11,336 |
| 54 | navicular*.ti,ab. | 4,914 |
| 55 | cuneiform*.ti,ab. | 2,979 |

| | | |
|----|--|-----------|
| 56 | cuboid*.ti,ab. | 11,139 |
| 57 | metatarsal*.ti,ab. | 18,973 |
| 58 | exp Foot Bones/ | 36,198 |
| 59 | phalan*.ti,ab. | 23,235 |
| 60 | Leg/ | 134,032 |
| 61 | (toe or toes).ti,ab. | 44,792 |
| 62 | exp Toes/ | 25,123 |
| 63 | pilon*.ti,ab. | 5,346 |
| 64 | 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 52 or 51 or 53 or 54 or 55 or 56 or 57 or 58 or 59 or 60 or 61 or 62 or 63 | 1,699,234 |
| 65 | 37 and 64 | 13,043 |
| 66 | Quality of Life.mp. or "Quality of Life"/ | 859,704 |
| 67 | quality of life.ti,ab. | 650,001 |
| 68 | life qualit*.ti,ab. | 18,889 |
| 69 | living qualit*.ti,ab. | 749 |
| 70 | quality of living.ti,ab. | 493 |
| 71 | Activities of Daily Living.mp. or "Activities of Daily Living"/ | 153,637 |
| 72 | activities of daily living.ti,ab. | 57,829 |
| 73 | activity of daily living.ti,ab. | 4,663 |
| 74 | activities of daily life.ti,ab. | 2,984 |
| 75 | activity of daily life.ti,ab. | 1,082 |
| 76 | daily living activit*.ti,ab. | 3,094 |
| 77 | daily life activit*.ti,ab. | 3,698 |
| 78 | adl.ti,ab. | 24,357 |
| 79 | chronic limitation of activity.ti,ab. | 3 |
| 80 | self care*.ti,ab. | 38,986 |
| 81 | Health Status.mp. or Health Status/ | 296,535 |
| 82 | Health Status.ti,ab. | 122,055 |
| 83 | level of health.ti,ab. | 9,984 |
| 84 | health level*.ti,ab. | 2,575 |
| 85 | qol.ti,ab. | 100,560 |
| 86 | hrql.ti,ab. | 8,780 |
| 87 | hrqol.ti,ab. | 38,485 |
| 88 | activity of daily living.ti,ab. | 4,663 |
| 89 | activities of daily life.ti,ab. | 2,984 |
| 90 | activity of daily life.ti,ab. | 1,082 |
| 91 | daily life activit*.ti,ab. | 3,698 |
| 92 | iadl.ti,ab. | 6,685 |
| 93 | living qualit*.ti,ab. | 749 |
| 94 | quality of living.ti,ab. | 493 |
| 95 | 66 or 67 or 68 or 69 or 70 or 71 or 72 or 73 or 74 or 75 or 76 or 77 or 78 or 79 or 80 or 81 or 82 or 83 or 84 or 85 or 86 or 87 or 88 or 89 or 90 or 91 or 92 or 93 or 94 | 1,266,060 |
| 96 | 33 and 34 and 65 and 95 | 163 |

Supplementary Material 2. COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) quality assessment tables for the Wales Lower Limb Recovery (WaLLTR) scale.

Source articles assessed for measurement properties:

1. **Trickett RW, Mudge E, Price P, Pallister I.** The development of a novel patient-derived recovery scale for open tibial fractures. *Bone Joint J.* 2020;102-B(1):17-25.
2. **Trickett RW, Mudge E, Price P, Pallister I.** A qualitative approach to recovery after open tibial fracture: the road to a novel, patient-derived recovery scale. *Injury.* 2012;43(7):1071-1078.

Measurement properties assessed/considered in the development of the Wales Lower Limb Recovery (WaLLTR) scale.

| Assessed/considered | Measurement property |
|---------------------|---|
| | <i>Content validity</i> |
| Yes | Box 1. PROM development |
| Yes | Box 2. Content validity |
| | <i>Internal structure</i> |
| N/A | Box 3. Structural validity |
| Yes | Box 4. Internal consistency |
| No | Box 5. Cross-cultural validity/measurement invariance |
| | <i>Remaining measurement properties</i> |
| Yes | Box 6. Reliability |
| Yes | Box 7. Measurement error |
| N/A | Box 8. Criterion validity |
| Yes | Box 9. Hypotheses testing for construct validity |
| Yes | Box 10. Responsiveness |

N/A, not applicable; PROM, patient-reported outcome measure.

COSMIN box 1. Standards for evaluating the quality of Wales Lower Limb Recovery (WaLLTR) scale development.

| <i>Ratings: V = very good; A = adequate; D = doubtful; I = inadequate; N = not applicable</i> | | Wales Lower Limb Recovery (WaLLTR) scale | | |
|---|--|---|----------------|------------------|
| 1a. Patient-reported outcome measure (PROM) design | | References: Studies 1 and 2 | | |
| <i>General design requirements</i> | | Rater 1 | Rater 2 | Consensus |
| 1 | Is a clear description provided of the construct to be measured? | V | V | V |
| 2 | Is the origin of the construct clear: was a theory, conceptual framework or disease model used or clear rationale provided to define the construct to be measured? | V | V | V |
| 3 | Is a clear description provided of the target population for which the PROM was developed? | V | V | V |
| 4 | Is a clear description provided of the context of use (i.e. discriminative, evaluative purpose, and/or predictive)? | V | V | V |
| 5 | Was the PROM development study performed in a sample representing the target population for which the PROM was developed? | V | V | V |
| <i>Concept elicitation (relevance and comprehensiveness)</i> | | Rater 1 | Rater 2 | Consensus |
| 6 | Was an appropriate qualitative data collection method used to identify relevant items for a new PROM? | V | V | V |
| 7 | Were skilled group moderators/interviewers used? | V | V | V |
| 8 | Were the group meetings or interviews based on an appropriate topic or interview guide? | V | V | V |
| 9 | Were the group meetings or interviews recorded and transcribed verbatim? | V | V | V |
| 10 | Was an appropriate approach used to analyse the data? | V | V | V |
| 11 | Was at least part of the data coded independently? | V | V | V |

| | | | | |
|---|---|---|---|---|
| 12 | Was data collection continued until saturation was reached? | V | V | V |
| 13 | For quantitative studies: was the sample size appropriate? | N | N | N |
| SUBTOTAL QUALITY CONCEPT ELICITATION STUDY <i>Lowest score of items 6-13</i> | | V | V | V |
| | | | | |
| TOTAL QUALITY OF THE PROM DESIGN <i>Lowest score of items 1-13</i> | | V | V | V |

1b. Cognitive interview study or other pilot test

| | | | | |
|------------------------------------|---|----------------|----------------|------------------|
| | | Rater 1 | Rater 2 | Consensus |
| 14 | Was a cognitive interview study or other pilot test performed? <i>If NO skip items 15-35</i> | YES | YES | YES |
| | | | | |
| <i>General design requirements</i> | | Rater 1 | Rater 2 | Consensus |
| 15 | Was the cognitive interview study or other pilot test performed in a sample representing the target population? | V | V | V |
| | | | | |
| <i>Comprehensibility</i> | | Rater 1 | Rater 2 | Consensus |
| 16 | Were patients asked about the <u>comprehensibility</u> of the PROM? <i>If NO or not clear, skip items 17-25</i> | D | V | V |
| | | Rater 1 | Rater 2 | Consensus |
| 17 | Were all items tested in their final form? | V | D | V |

| | | | | |
|---|--|---|---|---|
| 18 | Was an appropriate qualitative method used to assess the <u>comprehensibility</u> of the PROM instructions, items, response options, and recall period? | D | A | D |
| 19 | Was each item tested in an appropriate number of patients? | V | V | V |
| 20 | Were skilled interviewers used? | N | N | N |
| 21 | Were the interviews based on an appropriate interview guide? | N | N | N |
| 22 | Were the interviews recorded and transcribed verbatim? | N | N | N |
| 23 | Was an appropriate approach used to analyse the data? | A | A | A |
| 24 | Were at least two researchers involved in the analysis? | V | V | V |
| 25 | Were problems regarding the comprehensibility of the PROM instructions, items, response options, and recall period appropriately addressed by adapting the PROM? | V | V | V |
| SUBTOTAL QUALITY OF COMPREHENSIBILITY STUDY <i>Lowest score of items 15-25</i> | | D | D | D |

| <i>Comprehensiveness</i> | | Rater 1 | Rater 2 | Consensus |
|--------------------------|---|----------------|----------------|------------------|
| 26 | Were patients asked about the <u>comprehensiveness</u> of the PROM? <i>If NO or not clear, skip items 27-35</i> | D | D | D |
| | | Rater 1 | Rater 2 | Consensus |
| 27 | Was the final set of items tested? | - | - | - |
| 28 | Was an appropriate method used for assessing the comprehensiveness of the PROM? | - | - | - |
| 29 | Was each item tested in an appropriate number of patients? | - | - | - |
| 30 | Were skilled interviewers used? | - | - | - |
| 31 | Were the interviews based on an appropriate interview guide? | - | - | - |

| | | | | |
|---|--|---|---|---|
| 32 | Were the interviews recorded and transcribed verbatim? | - | - | - |
| 33 | Was an appropriate approach used to analyse the data? | - | - | - |
| 34 | Were at least two researchers involved in the analysis? | - | - | - |
| 35 | Were problems regarding the <u>comprehensiveness</u> of the PROM appropriately addressed by adapting the PROM? | A | A | A |
| SUBTOTAL QUALITY OF COMPREHENSIVENESS STUDY <i>Lowest score of items 15, 26-35</i> | | D | D | D |
| | | | | |
| TOTAL QUALITY OF THE PILOT STUDY <i>Lowest score of items 14-35</i> | | D | D | D |
| | | | | |
| TOTAL QUALITY OF THE PROM DEVELOPMENT STUDY <i>Lowest score of items 1-35</i> | | D | D | D |

COSMIN box 2. Standards for evaluating the quality of content validity studies of the Wales Lower Limb Recovery (WaLLTR) scale.

Trickett et al⁴ surveyed an alpha and beta scale on 35 and 228 participants, respectively, each of which contained additional free test space after each item for participant comments. This method could be assumed to have been used by the authors to evaluate the content validity of questionnaire items, i.e. allow patients to comment on the relevance, comprehensiveness, or compressibility of the WaLLTR items. However, this intention was not directly stated.

| | | Wales Lower Limb Recovery (WaLLTR) scale | | |
|--|--|--|----------------|------------------|
| <i>Score: V = very good; A = adequate; D = doubtful; I = inadequate; N= not applicable</i> | | References: Study 1 | | |
| 2a. Asking patient about relevance | | rater 1 | rater 2 | Consensus |
| 1 | Was an appropriate method used to ask patients whether each item is <u>relevant</u> for their experience with the condition? | A | D | A |
| 2 | Was each item tested in an appropriate number of patients? | V | D | V |
| 3 | Were skilled group moderators/interviewers used? | N | N | N |
| 4 | Were the group meetings or interviews based on an appropriate topic or interview guide? | N | N | N |
| 5 | Were the group meetings or interviews recorded and transcribed verbatim? | N | N | N |
| 6 | Was an appropriate approach used to analyse the data? | A | D | A |
| 7 | Were at least two researchers involved in the analysis? | A | A | A |
| SUBTOTAL QUALITY OF RELEVANCE STUDY <i>Lowest score of items 1-7</i> | | A | D | A |
| 2b. Asking patients about comprehensiveness | | rater 1 | rater 2 | Consensus |
| 8 | Was an appropriate method used for assessing the <u>comprehensiveness</u> of the PROM? | A | D | A |
| 9 | Was each item tested in an appropriate number of patients? | V | D | V |

| | | | | |
|--|---|----------|----------|----------|
| 10 | Were skilled group moderators/interviewers used? | N | N | N |
| 11 | Were the group meetings or interviews based on an appropriate topic or interview guide? | N | N | N |
| 12 | Were the group meetings or interviews recorded and transcribed verbatim? | N | N | N |
| 13 | Was an appropriate approach used to analyse the data? | A | D | A |
| 14 | Were at least two researchers involved in the analysis? | V | D | D |
| SUBTOTAL QUALITY OF COMPREHENSIVENESS STUDY <i>Lowest score of items 8-14</i> | | A | D | A |

| 2c. Asking patients about comprehensibility | | rater 1 | rater 2 | Consensus |
|---|---|----------------|----------------|------------------|
| 15 | Was an appropriate qualitative method used for assessing the <u>comprehensibility</u> of the PROM instructions, items, response options, and recall period? | I | D | I |
| 16 | Was each item tested in an appropriate number of patients? | V | V | V |
| 17 | Were skilled group moderators/interviewers used? | N | N | N |
| 18 | Were the group meetings or interviews based on an appropriate topic or interview guide? | N | N | N |
| 19 | Were the group meetings or interviews recorded and transcribed verbatim? | N | N | N |
| 20 | Was an appropriate approach used to analyse the data? | D | D | D |
| 21 | Were at least two researchers involved in the analysis? | V | V | V |
| SUBTOTAL QUALITY OF COMPREHENSIBILITY STUDY <i>Lowest score of items 15-21</i> | | I | D | I |

| 2d. Asking professionals about relevance | | rater 1 | rater 2 | Consensus |
|---|---|----------------|----------------|------------------|
| 22 | Was an appropriate method used to ask professionals whether each item is <u>relevant</u> for the construct of interest? | I | D | I |
| 23 | Were professionals from all relevant disciplines included? | D | D | D |
| 24 | Was each item tested in an appropriate number of professionals? | D | D | D |
| 25 | Was an appropriate approach used to analyse the data? | D | D | D |
| 26 | Were at least two researchers involved in the analysis? | D | D | D |
| SUBTOTAL QUALITY OF RELEVANCE STUDY <i>Lowest score of items 22-26</i> | | I | D | I |

| 2e. Asking professionals about comprehensiveness | | rater 1 | rater 2 | Consensus |
|---|--|----------------|----------------|------------------|
| 27 | Was an appropriate method used for assessing the <u>comprehensiveness</u> of the PROM? | I | D | I |
| 28 | Were professionals from all relevant disciplines included? | D | D | D |
| 29 | Was each item tested in an appropriate number of professionals? | D | D | D |
| 30 | Was an appropriate approach used to analyse the data? | D | D | D |
| 31 | Were at least two researchers involved in the analysis? | D | D | D |
| SUBTOTAL QUALITY OF COMPREHENSIVENESS STUDY <i>Lowest score of items 27-31</i> | | I | D | I |

Evaluation of patient-reported outcome measure (PROM) development and content validity studies on the Wales Lower Limb Recovery (WaLLTR) scale against the ten criteria for good measurement properties for content validity and grading for the quality of evidence using the modified Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach.

In the table below, the content validity of WaLLTR is rated based on the summary tables above.

The results of all available studies were not quantitatively summarized as only one PROM development study with an associated qualitative study has been published for the WaLLTR scale.^{4,5}

| Wales Lower Limb Recovery (WaLLTR) scale | | PROM development study and content validity study 1 | | | | | |
|---|---|---|---------|-----------|-------------------------------|----------------|----------------|
| | | RATING | | | QUALITY OF EVIDENCE | | |
| | | + / - / ? | | | High, moderate, low, very low | | |
| <i>Score: + = sufficient; - = insufficient; ? = indeterminate; ± = inconsistent</i> | | rater 1 | rater 2 | consensus | rater 1 | rater 2 | consensus |
| Relevance | | | | | | | |
| 1 | Are the included items relevant for the construct of interest? | + | + | + | | | |
| 2 | Are the included items relevant for the target population of interest? | + | + | + | | | |
| 3 | Are the included items relevant for the context of use of interest? | + | + | + | | | |
| 4 | Are the response options appropriate? | + | + | + | | | |
| 5 | Is the recall period appropriate? | ? | ? | ? | | | |
| | RELEVANCE RATING (+ / - / ± / ?) | + | + | + | moderate | moderate | moderate |
| Comprehensiveness | | | | | | | |
| 6 | Are all key concepts included? | ? | ? | ? | | | |
| | COMPREHENSIVENESS RATING (+ / - / ± / ?) | ? | ? | ? | Unable to rate | Unable to rate | Unable to rate |
| Comprehensibility | | | | | | | |
| 7 | Are the PROM instructions understood by the population of interest as intended? | - | ? | - | | | |
| 8 | Are the PROM items and response options understood by the population of interest as intended? | - | ? | - | | | |
| 9 | Are the PROM items appropriately worded? | - | + | - | | | |

| Wales Lower Limb Recovery (WaLLTR) scale | | PROM development study and content validity study 1 | | | | | |
|---|---|---|---------|-----------|-------------------------------|----------------|----------------|
| | | RATING | | | QUALITY OF EVIDENCE | | |
| | | + / - / ? | | | High, moderate, low, very low | | |
| <i>Score: + = sufficient; - = insufficient; ? = indeterminate; ± = inconsistent</i> | | rater 1 | rater 2 | consensus | rater 1 | rater 2 | consensus |
| 10 | Do the response options match the question? | - | + | - | | | |
| | COMPREHENSIBILITY RATING (+ / - / ± / ?) | - | ? | - | Unable to rate | Unable to rate | Unable to rate |
| | | | | | | | |
| | CONTENT VALIDITY RATING (+ / - / ± / ?) | ± | ? | ± | moderate | moderate | moderate |

COSMIN boxes 3 to 10. Evaluation of the internal structure of the Wales Lower Limb Recovery (WaLLTR) scale.

| | | Wales Lower Limb Recovery (WaLLTR) scale | | |
|--|---|--|----------------|------------------|
| <i>Score: V = very good; A = adequate; D = doubtful; I = inadequate; N= not applicable</i> | | References: Study 1 | | |
| 3. Structural validity | | rater 1 | rater 2 | Consensus |
| 1 | For CTT: Was exploratory or confirmatory factor analysis performed? | N | N | N |
| 2 | For IRT/Rasch: does the chosen model fit to the research question? | N | N | N |
| 3 | Was the sample size included in the analysis adequate? | N | N | N |
| 4 | Were there any other important flaws? | N | N | N |
| TOTAL <i>Lowest score of items 1-4</i> | | N | N | N |

The WaLLTR scale appears to be based on a formative model (items together form a construct), as such unidimensionality or structural validity is not relevant

| | | rater 1 | rater 2 | Consensus |
|--------------------------------|--|----------------|----------------|------------------|
| 4. Internal consistency | | | | |
| 1 | Was an internal consistency statistic calculated for each unidimensional (sub)scale separately? | V | V | V |
| 2 | For continuous scores: Was Cronbach's alpha or omega calculated? | V | N | V |
| 3 | For dichotomous scores: Was Cronbach's alpha or KR-20 calculated? | N | N | N |
| 4 | For IRT-based scores: Was standard error of the theta (SE (θ)) or reliability coefficient of estimated latent trait value (index of (subject or item) separation) calculated? | N | N | N |
| 5 | Were there any other important flaws? | V | V | V |

| | | | |
|---|---|---|---|
| TOTAL <i>Lowest score of items 1-5</i> | V | V | V |
|---|---|---|---|

| 5. Cross-cultural validity/measurement invariance | | rater 1 | rater 2 | Consensus |
|--|--|---------|---------|-----------|
| 1 | Were the samples similar for relevant characteristics except for the group variable? | D | D | D |
| 2 | Was an adequate approach used to analyse the data? | D | D | D |
| 3 | Was the sample size included in the analysis adequate? | D | D | D |
| 4 | Were there any other important flaws? | D | D | D |
| TOTAL <i>Lowest score of items 1-4</i> | | D | D | D |

| 6. Reliability | | rater 1 | rater 2 | Consensus |
|-----------------------|---|---------|---------|-----------|
| 1 | Were patients stable in the interim period on the construct to be measured? | A | D | A |
| 2 | Was the time interval appropriate? | V | D | V |
| 3 | Were the test conditions similar for the measurements? e.g. type of administration, environment, instructions | A | D | A |
| 4 | For continuous scores: Was an intraclass correlation coefficient (ICC) calculated? | A | A | A |
| 5 | For dichotomous/nominal/ordinal scores: Was kappa calculated? | N | N | N |
| 6 | For ordinal scores: Was a weighted kappa calculated? | N | N | N |

| | | | | |
|---|--|---|---|---|
| 7 | For ordinal scores: Was the weighting scheme described? e.g. linear, quadratic | N | N | N |
| 8 | Were there any other important flaws? | V | V | V |
| TOTAL <i>Lowest score of items 1-8</i> | | A | A | A |

| 7. Measurement error | | rater 1 | rater 2 | Consensus |
|---|---|---------|---------|-----------|
| 1 | Were patients stable in the interim period on the construct to be measured? | A | A | A |
| 2 | Was the time interval appropriate? | V | D | V |
| 3 | Were the test conditions similar for the measurements? e.g. type of administration, environment, instructions | A | A | A |
| 4 | For continuous scores: Was the Standard Error of Measurement (SEM), Smallest Detectable Change (SDC) or Limits of Agreement (LoA) calculated? | I | I | I |
| 5 | For dichotomous/nominal/ordinal scores: Was the percentage (positive and negative) agreement calculated? | N | N | N |
| 6 | Were there any other important flaws? | D | D | V |
| TOTAL <i>Lowest score of items 1-6</i> | | I | I | I |

| 8. Criterion validity | | rater 1 | rater 2 | Consensus |
|------------------------------|--|---------|---------|-----------|
| 1 | For continuous scores: Were correlations, or the area under the receiver operating curve calculated? | N | N | N |

| | | | | |
|---|--|---|---|---|
| 2 | For dichotomous scores: Were sensitivity and specificity determined? | N | N | N |
| 3 | Were there any other important flaws? | N | N | N |
| TOTAL <i>Lowest score of items 1-3</i> | | N | N | N |

9. Hypotheses testing for construct validity

9a. Comparison with other outcome measurement instruments (convergent validity)

| | | rater 1 | rater 2 | Consensus |
|---|---|---------|---------|-----------|
| 1 | Is it clear what the comparator instrument(s) measure(s)? | N | N | N |
| 2 | Were the measurement properties of the comparator instrument(s) adequate? | N | N | N |
| 3 | Was the statistical method appropriate for the hypotheses to be tested? | N | N | N |
| 4 | Were there any other important flaws? | N | N | N |
| TOTAL <i>Lowest score of items 1-4</i> | | N | N | N |

9b. Comparison between subgroups (discriminative or known-groups validity)

| | | rater 1 | rater 2 | Consensus |
|---|---|---------|---------|-----------|
| 5 | Was an adequate description provided of important characteristics of the subgroups? | A | A | A |
| 6 | Was the statistical method appropriate for the hypotheses to be tested? | A | A | A |
| 7 | Were there any other important flaws? | V | V | V |
| TOTAL <i>Lowest score of items 5-7</i> | | A | A | A |

10. Responsiveness

10a. Criterion approach (i.e. comparison to a gold standard)

1 For continuous scores: Were correlations between change scores, or the area under the Receiver Operator Curve (ROC) curve calculated?

2 For dichotomous scales: Were sensitivity and specificity (changed versus not changed) determined?

3 Were there any other important flaws?

TOTAL *Lowest score of items 1-3*

10b. Construct approach (i.e. hypotheses testing; comparison with other outcome measurement instruments)

4 Is it clear what the comparator instrument(s) measure(s)?

5 Were the measurement properties of the comparator instrument(s) adequate?

6 Was the statistical method appropriate for the hypotheses to be tested?

7 Were there any other important flaws?

TOTAL *Lowest score of items 4-7*

10c. Construct approach: (i.e. hypotheses testing: comparison between subgroups)

| rater 1 | rater 2 | Consensus |
|---------|---------|-----------|
| N | N | N |
| N | N | N |
| N | N | N |
| N | N | N |
| | | |
| rater 1 | rater 2 | Consensus |
| N | N | N |
| N | N | N |
| N | N | N |
| N | N | N |
| N | N | N |
| N | N | N |
| | | |
| rater 1 | rater 2 | Consensus |

| | | | | |
|----|--|----------------|----------------|------------------|
| 8 | Was an adequate description provided of important characteristics of the subgroups? | V | V | V |
| 9 | Was the statistical method appropriate for the hypotheses to be tested? | V | V | V |
| 10 | Were there any other important flaws? | V | V | V |
| | TOTAL <i>Lowest score of items 8-10</i> | V | V | V |
| | | | | |
| | 10d. Construct approach: (i.e. hypotheses testing: before and after intervention) | rater 1 | rater 2 | Consensus |
| 11 | Was an adequate description provided of the intervention given? | N | N | N |
| 12 | Was the statistical method appropriate for the hypotheses to be tested? | N | N | N |
| 13 | Were there any other important flaws? | N | N | N |
| | TOTAL <i>Lowest score of items 11-13</i> | N | N | N |

Evaluation of patient-reported outcome measure (PROM) development and content validity studies on the Wales Lower Limb Recovery (WaLLTR) scale against the updated criteria for good measurement properties.

| Wales Lower Limb Recovery (WaLLTR) scale | PROM development study and content validity study 1 | | | | | |
|---|---|---------|-----------|-------------------------------|----------|-----------|
| | RATING | | | QUALITY OF EVIDENCE | | |
| <i>Score: + = sufficient; - = insufficient; ? = indeterminate; ± = inconsistent</i> | + / - / ? | | | High, moderate, low, very low | | |
| | rater 1 | rater 2 | consensus | rater 1 | rater 2 | consensus |
| Structural validity | Not applicable | | | | | |
| Internal consistency | + | + | + | High | High | High |
| Cross-cultural validity/measurement invariance | ? | ? | ? | Very low | Very low | Very low |
| Reliability | + | + | + | Moderate | Moderate | Moderate |
| Measurement error | - | - | - | Very low | Very low | Very low |
| Criterion validity | Not undertaken | | | | | |
| Hypothesis testing for construct validity | + | + | + | Moderate | Moderate | Moderate |
| Responsiveness | + | + | + | High | High | High |

Supplementary Material 3. Outcome measurement instruments (OMIs) not validated in target population assessed against pragmatic inclusion criteria (see Figure 1) for inclusion at the consensus meeting.

| Key | | | | | | | | |
|--|--|--|------|----------------------------|------------------------------------|-----------------|---------------------------|--|
| Included | | Shaded green | | | | | | |
| Rejected | | Shaded red | | | | | | |
| OMI items assessed by ALA and HC for having face validity in measuring core outcomes | | | | | | | | |
| Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection | |
| Adverse events outcome measurement instruments | | Includes outcomes broadly labelled as some form of unintended consequence (e.g. adverse events/effects, adverse reactions, safety, harm, negative effects, toxicity, complications, sequelae). Specifically named adverse events should be classified within the appropriate taxonomy domain above with an additional level of categorization which identifies that this outcome is being considered as an adverse event. | | | | | | |
| 1 | Definitions of complications | 18 | | | | | Core outcome not measured | |
| 2 | Definitions of flap complications | 15 | | | | | Core outcome not measured | |
| Death outcome measurement instruments | | Includes overall (all-cause) survival/mortality and cause-specific survival/mortality, as well as composite survival outcomes that include death (e.g. disease-free survival, progression-free survival, amputation-free survival). | | | | | | |
| 3 | Charlson co-morbidity index ⁶ | Physician-reported outcome measurement instrument used to predict ten-year survival in patients with multiple comorbidities. | 1 | | | | Core outcome not measured | |
| 4 | Definitions of mortality | | 3 | | | | Core outcome not measured | |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection |
|--|--|------------------|------|----------------------------|------------------------------------|-----------------|---------------------------|
| Life impact outcome measurement instruments | Includes outcomes relating to the delivery of care, including - adherence/compliance - patient preference - tolerability/acceptability of intervention - withdrawal from intervention (e.g. time to treatment failure, reason for stopping therapy) - appropriateness of intervention - accessibility, quality, and adequacy of intervention - patient/carer satisfaction (emotional rather than financial burden) - process, implementation, and service outcomes | | | | | | |
| Satisfaction outcome measurement instruments | | | | | | | |
| 5 | Definitions of satisfaction | 1 | | | | | Core outcome not measured |
| 6 | The Patient Satisfaction Questionnaire Short Form (PSQ-18) ⁷ | 1 | | | | | Core outcome not measured |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| | Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection |
|---|--|---|------------------|----------|----------------------------|------------------------------------|-----------------|---------------------------|
| 7 | Visual analogue scale for satisfaction ⁸ | Patient-reported outcome measurement instruments assessing satisfaction on a continuous scale (0 to 100 mm). Novel score only used in citing study. | 1 | | | | | Core outcome not measured |
| | Global quality of life outcome measurement instruments | Includes only implicit composite outcomes measuring global quality of life. | | | | | | |
| 8 | Health-related quality of life outcome measurement instruments | | | | | | | |
| 9 | EuroQol five-dimension three-level (EQ-5D-3L) ⁹ | <p>The EQ-5D-3L descriptive system comprises the following five dimensions:</p> <ol style="list-style-type: none"> 1. mobility 2. self-care 3. usual activities 4. pain/discomfort 5. anxiety/depression <p>Each dimension has three levels: no problems, some problems, and extreme problems. The patient is asked to indicate his/her health state by ticking the box next to the most appropriate statement in each of the five dimensions.</p> | 2 | XX XX | XX XX | XX XX | XX XX | Included |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| | Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection |
|----|--|--|------------------|----------|----------------------------|------------------------------------|-----------------|----------------------|
| 10 | EuroQol five-dimension five-level (EQ-5D-5L) ¹⁰ | <p>Patient-reported outcome measurement instrument. The descriptive system comprises five dimensions:</p> <ol style="list-style-type: none"> 1. mobility 2. self-care 3. usual activities 4. pain/discomfort 5. anxiety/depression <p>Each dimension has five levels: no problems, slight problems, moderate problems, severe problems, and extreme problems. The patient is asked to indicate his/her health state by ticking the box next to the most appropriate statement in each of the five dimensions.</p> | 6 | XX XX | XX XX | XX XX | XX XX | Included |
| 11 | Short Form-12 (SF-12) ¹¹ | <p>Patient-reported outcome measurement instrument. Assessing the following domains:</p> <ol style="list-style-type: none"> 1. physical functioning 2. role-physical 3. bodily pain 4. general health 5. vitality 6. social functioning | 5 | XX XX | XX | XX XX | XX XX | Included |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection |
|------|---|------------------|--------------|----------------------------|------------------------------------|-----------------|----------------------------------|
| | 7. role-emotional 8. mental health | | | | | | |
| 12 | Short Form-36 (SF-36) ¹² Patient-reported outcome measurement instrument. Assessing the following domains: 1. physical functioning 2. role-physical 3. bodily pain 4. general health 5. vitality 6. social functioning 7. role-emotional 8. mental health | 17 | XX XX | XX XX | XX XX | XX XX | Included |
| 13 | Short Form-6 Dimensions (SF-6D) ¹³ Patient-reported outcome measurement instrument. Assessing the following domains: 1. physical functioning 2. role limitations 3. social functioning 4. pain 5. mental health | 1 | AA HC | HC | AA HC | AA HC | Used only once in the literature |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection | |
|------|--|--|------|----------------------------|------------------------------------|-----------------|----------------------|--|
| | | 6. vitality | | | | | | |
| 14 | <p>Sickness Impact Profile (SIP)^{14,15}</p> <p>Patient-reported outcome measurement instrument to assess perceived health status. The full-length SIP consists of 136 items within 12 domains:</p> <ol style="list-style-type: none"> 1. sleep and rest 2. eating 3. work 4. home management 5. recreation and pastimes 6. ambulation 7. mobility 8. body care and movement 9. social interaction 10. alertness behaviour 11. emotional behaviour 12. communication | 5 | XX | XX XX | XX XX | XX XX | Included | |
| 15 | Veterans Affairs System of Rating Disabilities ¹⁶ | Physician/technician reported outcome measurement instrument used to assess the level of disability for calculation of disability compensation. Each body system is assessed | 2 | N/A | N/A | N/A | N/A | Unfeasible – physician/technician reported |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection |
|---|---|------------------|----------|----------------------------|------------------------------------|-----------------|--|
| | separately for the level of disability, and a combined score is generated. | | | | | | |
| 16 Veterans RAND 12 Item Health Survey (VR-12) ¹⁷ | Patient-reported outcome measurement instrument assesses domains: 1. general health perceptions 2. physical functioning 3. role physical 4. role emotional 5. bodily pain 6. vitality/mental health 7. social functioning 8. change physical 9. change emotional | 1 | XX XX | | XX XX | XX XX | Not applicable to whole of UK open fracture population (i.e. developed for veterans) |
| Physiological or clinical outcome measurement instruments | Physiological/clinical outcomes include measures of physiological function, signs, and symptoms, as well as laboratory (and other scientific) measures relating to physiology and are categorized according to the underlying cause/body system. | | | | | | |
| Musculoskeletal and connective tissue outcome measurement instruments | | | | | | | |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| | Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection |
|----|--|---|------------------|----------|----------------------------|------------------------------------|-----------------|---|
| 17 | American Orthopedics Foot and Ankle Society (AOFAS) ankle-hindfoot scale ¹⁸ | <p>Patient and clinically reported outcome measurement instrument. Response domains include:</p> <ol style="list-style-type: none"> 1. pain 2. functional limitation on recreational and daily activities 3. maximum walking distance in blocks 4. walking ability on different surfaces 5. Gait abnormality 6. Sagittal motion 7. hindfoot motion 8. ankle -hindfoot stability 9. Alignment | 23 | XX XX | XX XX | | | <p>scale specific to ankle and hindfoot</p> <p>Unfeasible – clinician-reported component to scale</p> |
| | Appearance outcome measurement instruments | | | | | | | |
| 18 | Cosmetic outcome score by O’Toole et al. ¹⁹ | <p>Patient-reported outcome measurement instrument to assess satisfaction. The patient’s cosmetic outcome score was based on the question, “How satisfied are you with the appearance of your injured leg or artificial leg?” Participants were asked to respond to each question by using a five-point scale of descriptors</p> | 1 | | | | | Core outcome not measured |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| | Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection |
|----|---|---|------------------|----------|----------------------------|------------------------------------|-----------------|--|
| 19 | Vancouver scar scale ²⁰ | <p>that included “not at all satisfied,” “slightly satisfied,” “moderately satisfied,” “very satisfied,” and “completely satisfied.” This was a novel instrument only cited in this study</p> <p>Physician-reported OMI to assess scars and burns. Assessment areas:</p> <ol style="list-style-type: none"> 1. Vascularity 2. Height/thickness 3. pliability 4. pigmentation | 1 | | | | | Core outcome not measured |
| 20 | Association for the Study and Application of the Methods of Ilizarov (ASAMI) criteria ²¹ | <p>Physician-reported outcome measurement instrument. The ASAMI criteria assess two domains:</p> <ol style="list-style-type: none"> 1. bone results (union, infection, deformity, re-fracture, and limb length discrepancy) 2. functional results (active walking, limp, stiffness, knee extension, ankle dorsiflexion, reflex sympathetic dystrophy, pain, amputation, employment, and return to activities of daily living) | 24 | XX XX | XX | XX XX | | Unfeasible – physician/technician reported |
| | Bone outcome measurement instruments | | | | | | | |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| | Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection |
|----|--|--|------------------|------|----------------------------|------------------------------------|-----------------|--|
| 21 | Definitions of bone union | | 50 | | | | | Core outcome not measured |
| 22 | Definitions of clinical union | | 6 | | | | | Core outcome not measured |
| 23 | Definitions of malunion | | 20 | | | | | Core outcome not measured |
| 24 | Definitions of nonunion | | 16 | | | | | Core outcome not measured |
| 25 | Definitions of radiographic bone union | | 33 | | | | | Core outcome not measured |
| 26 | Chen grading system ²² | <p>Physician-reported outcome measurement instrument to grade the functional status of the lower limb. Domains covered include:</p> <ol style="list-style-type: none"> 1. return to previous work 2. walking with a normal gait 3. range of movement of the knee and ankle 4. sensation 5. trophic ulcers | 1 | | XX XX | XX XX | | Unfeasible – physician/technician reported |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| | Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection |
|----|---|---|------------------|------|----------------------------|------------------------------------|-----------------|-----------------------------------|
| 27 | Definitions of amputation | | 5 | | | | | Core outcome not measured |
| 28 | Definitions of pin site loosening | | 1 | | | | | Core outcome not measured |
| 29 | Disability Rating Index (DRI) ^{23,24} | <p>Patient-reported outcome measurement instrument assessing the level of disability. It consists of three domains with the following 12 response items:</p> <ol style="list-style-type: none"> 1. basic activities of daily life: dressing, outdoor walks, climbing stairs, and sitting a long time. 2. daily physical activities: standing bent over a sink, carrying a bag, making a bed, and running. 3. work-related/more vigorous activities: light work, heavy work, lifting heavy objects, and participating in exercise/sports. | 3 | | XX XX | | | Included |
| 30 | Disabilities of the Arm, Shoulder and Hand (DASH) questionnaire ²⁵ | <p>Patient-reported outcome measurement instrument specific to the upper extremity. The DASH is a 30-item symptom scale delivered over three domains:</p> <ol style="list-style-type: none"> 1. degree of difficulty in performing different physical activities because of the arm, shoulder, or hand problem (21 items) | 3 | | | XX | | Scale specific to upper extremity |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection | |
|------|---|--|------|----------------------------|------------------------------------|-----------------|----------------------|--|
| | <p>2. the severity of each of the symptoms of pain, activity-related pain, tingling, weakness, and stiffness (five items)</p> <p>3. the problem's impact on social activities, work, sleep, and self-image (four items)</p> | | | | | | | |
| 31 | Enneking score ²⁶ | <p>Physician-reported outcome measurement instrument that includes functional assessment measures. It assesses:</p> <ol style="list-style-type: none"> 1. pain 2. function 3. emotional acceptance 4. supports (walking aids) 5. walking 6. gait | 5 | XX XX | XX XX | | | Unfeasible – physician/technician reported |
| 32 | External fixation index (EFI) ²⁷ | <p>Physician-reported outcome measurement instrument. The EFI was calculated by dividing the time (days) in the external fixator by the lengthening achieved (centimetres).</p> | 1 | | | | | Unfeasible – physician/technician reported |
| 33 | Frequency intensity time (FIT) index ¹⁷ | <p>Patient-reported outcome measurement instrument assessing a person's level of physical activity over three domains:</p> | 1 | | | | | Core outcome not measured |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection |
|------|---|------------------|----------|----------------------------|------------------------------------|-----------------|---|
| 34 | <p>Functional outcome criteria by Tu et al²⁸</p> <p>1. frequency of exercise 2. intensity of exercise 3. time spent on workout</p> <p>Physician-reported outcome measurement instrument assessing functional outcome following open lower limb fractures in the following domains:</p> <p>1. pain 2. range of motion 3. ability to return to normal work</p> | 1 | XX XX | | XX XX | | Unfeasible – physician/technician reported |
| 35 | <p>Hamlyn Mobility Score (HMS)²⁹</p> <p>Physical performance and patient-reported outcome measurement instrument to assess the physical performance of patients doing a timed six-minute walk test, timed up and downstairs test, and a timed up and go test. Standard test metrics are recorded, e.g. time and distance with the addition of estimated step variation using an ear-worn accelerometer. The HMS also consists of four questions assessing:</p> <p>1. use of mobility aids 2. participation in work and leisure activities 3. satisfaction with walking</p> | 1 | XX XX | XX XX | XX XX | | Unfeasible – physician/technician reported objective physical performance measure |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection | |
|------|---|--|------|----------------------------|------------------------------------|-----------------|----------------------|----------------------------|
| 36 | Hospital for Special Surgery (HSS) knee scoring system ^{30,31} | <p>4. pain</p> <p>Physician-reported outcome measurement instrument to assess outcome after knee replacement surgery. The HSS knee score assesses seven domains:</p> <p>1. pain (rest pain, support required because of pain)</p> <p>2. stability (measured as total varus-valgus arc, extension)</p> <p>3. motion (measured as total passive arc)</p> <p>4. quadriceps strength (measured as a percentage of normal for age and sex)</p> <p>5. subtractions (for extension lag, flexion contracture, fixed varus or valgus deformity)</p> | 1 | XX XX | | | | Scale specific to the knee |
| | Infection outcome measurement instruments | | | | | | | |
| 37 | Definitions of deep infection | | 31 | | | | | Core outcome not measured |
| 38 | Definitions of deep wound Infection | | 5 | | | | | Core outcome not measured |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| | Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection |
|----|---|-------------|------------------|------|----------------------------|------------------------------------|-----------------|---------------------------|
| 39 | Definitions of infected implant | | 1 | | | | | Core outcome not measured |
| 40 | Definitions of infection | | 45 | | | | | Core outcome not measured |
| 41 | Definitions of joint sepsis | | 3 | | | | | Core outcome not measured |
| 42 | Definitions of osteomyelitis | | 11 | | | | | Core outcome not measured |
| 43 | Definitions of pin site infection | | 4 | | | | | Core outcome not measured |
| 44 | Definitions of recurrent infection | | 1 | | | | | Core outcome not measured |
| 45 | Definitions of wound infection | | 11 | | | | | Core outcome not measured |
| | Injury severity outcome measurement instruments | 0 | | | | | | |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| | Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection |
|----|--|--|------------------|----------|----------------------------|------------------------------------|-----------------|-----------------------------|
| 46 | Injury Severity Score (ISS) ^{32,33} | <p>Physician-reported outcome measurement instrument to assess the severity of injury over six body domains:</p> <ol style="list-style-type: none"> 1. head and neck 2. face 3. chest 4. abdomen 5. extremity (including pelvis) 6. external | 2 | | | | | Core outcome not measured |
| 47 | Mangled Extremity Severity Score (MESS) ^{34,35} | <p>Physician-reported outcome measurement instrument. Used to estimate the viability of an extremity after trauma. It assesses the following domains:</p> <ol style="list-style-type: none"> 1. limb ischaemia for greater than six hours 2. limb ischaemia 3. patient age 4. shock 5. injury mechanism | 4 | | | | | Core outcome not measured |
| 48 | Iowa ankle-evaluation rating system ³⁶ | <p>Physician-reported outcome measurement instrument assessing the function of the ankle. Domains assessed include:</p> | 5 | XX XX | XX XX | XX | | Scale specific to the ankle |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection |
|------|---|------------------|----------|----------------------------|------------------------------------|-----------------|---|
| 49 | <p>Johner-Wruhs evaluation³⁷</p> <p>1. function (housework or job, stair climbing, carrying heavy objects, ability to run, participate in athletics or heavy labour, walking independently, able to do garden work, difficulty getting in and out of a car)</p> <p>2. pain</p> <p>3. gait</p> <p>4. range of motion</p> <p>Physician-reported outcome measurement instrument assessing surgical outcome following a tibial fracture. Domains assessed:</p> <p>1. bone union, osteitis, or amputation</p> <p>2. neurovascular disturbances</p> <p>3. bone deformity</p> <p>4. range of motion</p> <p>5. pain</p> <p>6. gait</p> <p>7. participation in strenuous activities</p> | 10 | XX XX | XX XX | | | <p>Unfeasible – physician/technician reported</p> <p>Unfeasible – physician/technician reported</p> |
| 50 | <p>Lower Extremity Functional Scale (LEFS)³⁸</p> <p>Patient-reported outcome measurement instrument to assess a broad range of lower-extremity orthopaedic conditions including the hip, knee, leg ankle, or foot. The LEFS contains 20</p> | 3 | | XX XX | XX XX | | Included |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection |
|------|--|------------------|--------------|----------------------------|------------------------------------|-----------------|--|
| | items specifically assessing the International Classification of Functioning, Disability, and Health model (ICF) domains of activity and participation. | | | | | | |
| 51 | Lysholm Knee Scale ³⁹ Patient-reported outcome measurement instrument to assess knee function. Eight domains assessed including: 1. limp 2. walking support 3. pain 4. instability 5. locking 6. swelling 7. stair-climbing 8. squatting | 1 | XX XX | XX XX | | | Scale specific to the knee |
| 52 | Maryland foot score (MFS) ⁴⁰ Patient- and physician-reported outcome measurement instrument to assess foot injuries. Domains include: 1. pain 2. function (gait, distance walked, stability, support, limp, and wearing shoes) | 3 | XX XX | XX XX | | | Unfeasible – physician/technician reported component Scale specific to the foot |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| | Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection |
|----|--|--|------------------|----------|----------------------------|------------------------------------|-----------------|--|
| 53 | Mazur ankle function evaluation ⁴¹ | <p>Patient- and physician-reported outcome measurement instrument to assess ankle function over 12 response items in two domains:</p> <ol style="list-style-type: none"> 1. pain 2. function (limp, walking distance, support, hills up, hills down, stairs up, stairs down, ability to rise on toes, running, range of motion, plantar flexion) | 1 | XX XX | XX XX | | | <p>Unfeasible – physician/technician reported component</p> <p>Scale specific to the ankle</p> |
| 54 | Musculoskeletal Function Assessment (MFA) ⁴²⁻⁴⁴ | <p>Patient-reported outcome measurement instrument assessing musculoskeletal function consisting of 100 response items over the following domains:</p> <ol style="list-style-type: none"> 1. mobility 2. hand and fine motor 3. housework 4. self-care 5. sleep and rest 6. leisure and recreation 7. family relationships 8. cognition and thinking 9. emotional adjustment and adaptation 10. employment | 1 | XX XX | XX XX | XX XX | | Used only once in the literature |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| | Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection |
|----|--|---|------------------|----------|----------------------------|------------------------------------|-----------------|--|
| 55 | Neer knee score ^{45,46} | <p>Physician-reported outcome measurement instrument to assess knee function following a supracondylar fracture. The following domains are assessed:</p> <ol style="list-style-type: none"> 1. pain 2. function (as before injury, mild restriction, restricted; stairs sideways, cane or severe restriction, crutches or brace) 3. range of motion 4. work 5. gross anatomy 6. roentgenogram | 2 | XX XX | XX | XX XX | | <p>Unfeasible: physician/technician reported component</p> <p>Scale specific to the knee</p> |
| | Objective physical performance outcome measurement instruments | | | | | | | |
| 56 | Six-minute walk test ⁴⁷ | Objective physical performance measurement instrument used to assess functional exercise capacity. The six-minute walk test measures the distance an individual is able to walk over a total of six minutes on a hard, flat surface. | 2 | | XX XX | | | Unfeasible – physician/technician reported physical objective measure |
| 57 | Definitions of range of movement outcome measurement instruments | | 6 | | | | | Unfeasible – physician/technician reported |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| | Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection |
|----|--|--|------------------|----------|----------------------------|------------------------------------|-----------------|---|
| 58 | Timed up and downstairs ^{48,49} | Objective physical performance measurement instrument to assess physical performance by measuring the time taken to ascend and descend a flight of stairs consisting of 14 steps. | 1 | | XX XX | | | Unfeasible – physician/technician reported physical objective measure |
| 59 | Timed up and go test ⁵⁰ | Objective physical performance measurement instrument to determine fall risk and measure the progress of balance, sit to stand and walking. An individual is timed standing up from a chair with an armrest, walking 3 m, turning around, walking back, and sitting down again | 1 | | XX XX | | | Unfeasible – physician/technician reported physical objective measure |
| 60 | Timed walk test ⁵¹ | Objective physical performance measurement instrument to assess mobility by timing a patient to walk 30.5 m | 1 | | XX XX | | | Unfeasible – physician/technician reported physical objective measure |
| | Pain outcome measurement instruments | | | | | | | |
| 61 | Brief Pain Inventory (BPI) ⁵² | Patient-reported outcome measurement instrument assessing the severity of a patient's pain and the impact of this pain on the patient's daily functioning. The BPI consists of nine items assessing the following: 1. pain other than everyday kind of pains 2. location of pain (shading diagram) | 2 | XX XX | | | | Included |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection | |
|------|--|---|------|----------------------------|------------------------------------|-----------------|----------------------|-----------------------------|
| 62 | <p>Visual analogue scale for pain (VAS)⁵³</p> | <p>3. worst pain in the last 24 hours 4. least worst pain in the last 24 hours 5. pain on average 6. pain right now 7. treatments/medications for pain 8. relief of pain from treatment/medication 9. pain interference in the last 24 hours (general activity, mood, walking ability, normal work, relations with other people, sleep, enjoyment of life)</p> <p>Patient-reported outcome measurement instrument to assess pain. The VAS is a unidimensional measure of pain intensity. Often assessed on a 0 to 100 mm scale.</p> | 6 | XX XX | | | | Included |
| 63 | <p>Olerud and Molander scoring system⁵⁴</p> | <p>Patient-reported outcome measurement instrument assessing symptoms after ankle fracture in nine domains:</p> <p>1. pain 2. stiffness 3. swelling 4. stair climbing</p> | 1 | XX XX | XX XX | XX XX | | Scale specific to the ankle |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection |
|------|---|------------------|--------------|----------------------------|------------------------------------|-----------------|--|
| | 5. running 6. jumping 7. squatting 8. supports 9. work and activities of daily living | | | | | | |
| 64 | Paley criteria ⁵⁵⁻⁵⁷ Physician-reported outcome measurement instrument assessing bone and functional outcomes of the lower leg. Domains assessed: 1. Bone (consolidation, absence of infection, axial defect, limb-length discrepancy, docking site and osteogenesis zone consolidation solid enough not to require protection, nonunion and bone infection) 2. Function (pain free, walking, joint stiffness, talocrural or subtalar motion, use of analgesia, and activities of daily living) | 12 | XX XX | XX XX | XX XX | | Unfeasible – physician/technician reported |
| 65 | Sanders knee score ^{58,59} Physician-reported outcome measurement instrument assessing knee function following distal femoral fracture assessing the following domains: 1. Range of movement 2. Pain | 1 | XX XX | XX XX | XX XX | | 1. Unfeasible – physician/technician reported component 2. Scale specific to the knee |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection |
|------|---|------------------|----------|----------------------------|------------------------------------|-----------------|--|
| 66 | <p>Severn scale scoring system by Puno et al⁶⁰</p> <p>Physician-reported outcome measurement instrument assessing limb salvaged limb function in the following domains:</p> <ol style="list-style-type: none"> 3. Deformity 4. walking ability (walking and stair climbing) 5. return to work (employment/return to pre-injury functioning) 1. pain 2. activities of daily living 3. range of motion at the ankle and knee 4. residual deformity 5. radiological examination of degenerative joint changes and alignment 6. muscle strength of the foot 7. sensation | 7 | XX XX | XX XX | | | Unfeasible – physician/technician reported component |
| 67 | <p>Short Musculoskeletal Function Assessment Questionnaire (SMFA)⁶¹</p> <p>Patient-reported outcome measurement instrument assessing musculoskeletal function consisting of 46 items reduced from 101 in the MFA. There are two parts to the SMFA:</p> <ol style="list-style-type: none"> 1. dysfunction index (34 items assessing patient perceptions of functional performance in four categories including daily activities, emotional | 3 | XX XX | XX XX | XX XX | | |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection |
|--|---|------------------|----------|----------------------------|------------------------------------|-----------------|--|
| | <p>status, function of the arm and hand and mobility)</p> <p>2. bother index (12 items assessing broad functional areas including recreation and leisure, sleep and rest, work, and family)</p> | | | | | | |
| 68 | <p>The Knee Society clinical rating system⁶²</p> <p>Physician- and patient-reported outcome measurement instrument assessing:</p> <ol style="list-style-type: none"> 1. pain (on walking and climbing stairs) 2. range of motion 3. stability (medial/lateral and anterior/posterior) 4. deductions for extension lag, flexion contracture, malalignment, and pain at rest | 4 | XX XX | XX XX | | | <ol style="list-style-type: none"> 1. Unfeasible – physician/technician reported component 2. Scale specific to the knee |
| 69 | <p>Visual Analogue Scale Foot and Ankle (VAS FA)^{63,64}</p> <p>Patient-reported outcome measurement instrument to assess the foot and ankle in the following domains:</p> <ol style="list-style-type: none"> 1. pain 2. function 3. other complaints | 1 | XX XX | XX XX | XX XX | | Scale specific to the foot and ankle |
| Nervous system outcome measurement instruments | | | | | | | |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| | Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection |
|----|--|---|------------------|------|----------------------------|------------------------------------|-----------------|---------------------------|
| 70 | Oxford Medical Research Council (MRC) score ⁶⁵ | Physician-reported OMI to assess muscle strength. Response items: 1. flicker of movement 2. through full range actively with gravity counter balanced 3. through full range actively against gravity 4. through full range actively against some resistance 5. through full range actively against strong resistance | 2 | | | | | Core outcome not measured |
| | Psychiatric outcome measurement instruments | | | | | | | |
| | Depression outcome measurement instruments | | | | | | | |
| 71 | Patient Health Questionnaire (PHQ-9) ⁶⁶ | Patient-reported outcome measurement instrument measuring depression against nine domains for depression taken from the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition. | 1 | | | | | Core outcome not measured |
| | Post-traumatic stress disorder outcome measurement instruments | | | | | | | |

OMI items assessed by ALA and HC for having face validity in measuring core outcomes

| | Name | Description | Frequency of use | Pain | Walking, gait and mobility | Being able to return to life roles | Quality of life | Reason for rejection |
|----|---|---|------------------|------|----------------------------|------------------------------------|-----------------|---------------------------|
| 72 | Post-traumatic stress disorder (PTSD) checklist ⁶⁷ | Patient-reported outcome measurement instrument assessing the severity of PTSD symptoms. The checklist has 17 response items based on the Diagnostic and Statistical Manual of Mental Disorders-IV criteria for PTSD. | 1 | | | | | Core outcome not measured |
| | Stress outcome measurement instruments | | | | | | | |
| 73 | Definitions of physical and mental stress | | 1 | | | | | Core outcome not measured |
| | Renal and urinary outcome measurement instruments | | | | | | | |
| | Acute kidney injury outcome measurement instruments | | | | | | | |
| 74 | Definitions of acute kidney injury | | 2 | | | | | Core outcome not measured |
| 75 | RIFLE criteria ⁶⁸ | Physician-reported OMI assessing acute kidney injury assessment areas: 1. Risk of kidney injury 2. Injury to kidney 3. Failure of kidney | 1 | | | | | Core outcome not measured |

Supplementary Material 4. Results of face validity checks for each outcome measurement instrument against core outcomes.

| Electronic voting scores for face validity for each shortlisted outcome measurement instrument against core outcomes (number of voters = 25) | | | | |
|---|---|--|--|--|
| Outcome measurement instrument | Quality of life Number of votes (%) | Being able to return to life roles Number of votes (%) | Walking, gait and mobility Number of votes (%) | Pain or discomfort Number of votes (%) |
| EuroQol five-dimension three-level (EQ-5D-3L) | 20 (80) | 11 (44) | 13 (52) | 16 (64) |
| EuroQol five-dimension five-level (EQ-5D-5L) | 16 (64) | 12 (48) | 20 (80) | 17 (68) |
| Short Form-12 (SF-12) | 22 (88) | 11 (44) | 10 (40) | 11 (44) |
| Short Form-36 (SF-36) | 21 (84) | 15 (60) | 17 (68) | 15 (60) |
| Sickness Impact Profile (SIP) | 12 (48) | 5 (20) | 9 (36) | 3 (12) |
| Disability Rating Index (DRI) | 1 (4) | 6 (24) | 18 (72) | 0 (0) |
| Lower Extremity Functional Scale (LEFS) | 3 (12) | 8 (32) | 24 (96) | 1 (4) |
| Short Musculoskeletal Function Assessment Questionnaire (SMFA) | 7 (28) | 10 (40) | 16 (64) | 2 (8) |
| Brief Pain Inventory (BPI) | 0 (0) | 3 (12) | 0 (0) | 23 (92) |
| Visual analogue scale for pain (VAS) | 0 (0) | 1 (4) | 0 (0) | 22 (88) |
| Wales Lower Limb Trauma Recovery (WaLTTR) scale | 15 (60) | 20 (80) | 10 (40) | 11 (44) |

References

1. **No authors listed.** Search Blocks. Biomedische Informatie Group (BMI). <https://www.bmi-online.nl/searchblocks/search-blocks-bmi/> (date last accessed 12 April 2023).
2. **No authors listed.** PROM Group Construct & Instrument Type Filters: February 2010. Patient-Reported Outcome Measurement Group, Department of Public Health, University of Oxford. 2010. <https://cosmin.nl/wp-content/uploads/prom-search-filter-oxford-2010.pdf> (date last accessed 5 June 2022).
3. **Terwee CB, Jansma EP, Riphagen II, de Vet HCW.** Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res.* 2009;18(8):1115–1123.
4. **Trickett RW, Mudge E, Price P, Pallister I.** The development of a novel patient-derived recovery scale for open tibial fractures. *Bone Joint J.* 2020;102-B(1):17–25.
5. **Trickett RW, Mudge E, Price P, Pallister I.** A qualitative approach to recovery after open tibial fracture: the road to a novel, patient-derived recovery scale. *Injury.* 2012;43(7):1071–1078.
6. **Charlson M, Szatrowski TP, Peterson J, Gold J.** Validation of a combined comorbidity index. *J Clin Epidemiol.* 1994;47(11):1245-1251.
7. **Thayaparan AJ, Mahdi E.** The Patient Satisfaction Questionnaire Short Form (PSQ-18) as an adaptable, reliable, and validated tool for use in various settings. *Med Educ Online.* 2013;18:21747.
8. **Frisvoll C, Clarke-Jenssen J, Madsen JE, et al.** Long-term outcomes after high-energy open tibial fractures: Is a salvaged limb superior to prosthesis in terms of physical function and quality of life? *Eur J Orthop Surg Traumatol.* 2019;29(4):899-906.
9. **van Hout B, Janssen MF, Feng YS, et al.** Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. *Value Health.* 2012;15(5):708-715.
10. **Herdman M, Gudex C, Lloyd A, et al.** Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res.* 2011;20(10):1727-1736.
11. **Jenkinson C, Layte R, Jenkinson D, et al.** A shorter form health survey: can the SF-12 replicate results from the SF-36 in longitudinal studies? *J Public Health Med.* 1997;19(2):179-186.
12. **Jenkinson C, Coulter A, Wright L.** Short form 36 (SF36) health survey questionnaire: normative data for adults of working age. *BMJ.* 1993;306(6890):1437-1440.
13. **Obradovic M, Lal A, Liedgens H.** Validity and responsiveness of EuroQol-5 dimension (EQ-5D) versus Short Form-6 dimension (SF-6D) questionnaire in chronic pain. *Health Qual Life Outcomes.* 2013;11:110.
14. **Bergner M, Bobbitt RA, Carter WB, Gilson BS.** The Sickness Impact Profile: development and final revision of a health status measure. *Med Care.* 1981;19(8):787-805.

15. **Post MW, de Bruin A, de Witte L, Schrijvers A.** The SIP68: a measure of health-related functional status in rehabilitation medicine. *Arch Phys Med Rehabil.* 1996;77(5):440-445.
16. **McGeary M, Ford MA, McCutchen SR, Barnes DK.** *A 21st Century System for Evaluating Veterans for Disability Benefits.* Washington, DC: The National Academies Press. 2007.
17. **Selim AJ, Rogers W, Fleishman JA, et al.** Updated U.S. population standard for the Veterans RAND 12-item Health Survey (VR-12). *Qual Life Res.* 2009;18(1):43-52.
18. **Kitaoka HB, Alexander IJ, Adelaar RS, Nunley JA, Myerson MS, Sanders M.** Clinical rating systems for the ankle-hindfoot, midfoot, hallux, and lesser toes. *Foot Ankle Int.* 1994;15(7):349-353.
19. **O'Toole RV, Castillo RC, Pollak AN, MacKenzie EJ, Bosse MJ, LEAP Study Group.** Surgeons and their patients disagree regarding cosmetic and overall outcomes after surgery for high-energy lower extremity trauma. *J Orthop Trauma.* 2009;23(10):716-723.
20. **Baryza MJ, Baryza GA.** The Vancouver Scar Scale: an administration tool and its interrater reliability. *J Burn Care Rehabil.* 1995;16(5):535-538.
21. **Inam M, Saeed M, Khan I, Durrani A, Satar A, Arif M.** Outcome of ilizarov fixator in tibial non-union. *J Pak Med Assoc.* 2015;65(11 Suppl 3):S94-9.
22. **Parmaksizoglu F, Koprulu AS, Unal MB, Cansu E.** Early or delayed limb lengthening after acute shortening in the treatment of traumatic below-knee amputations and Gustilo and Anderson type IIC open tibial fractures: The results of a case series. *J Bone Joint Surg Br.* 2010;92-B(11):1563-1567.
23. **Salén BA, Spangfort EV, Nygren AL, Nordemar R.** The Disability Rating Index: an instrument for the assessment of disability in clinical settings. *J Clin Epidemiol.* 1994;47(12):1423-1435.
24. **Parsons H, Bruce J, Achten J, Costa ML, Parsons NR.** Measurement properties of the Disability Rating Index in patients undergoing hip replacement. *Rheumatology (Oxford).* 2015;54(1):64-71.
25. **Gummesson C, Atroshi I, Ekdahl C.** The disabilities of the arm, shoulder and hand (DASH) outcome questionnaire: longitudinal construct validity and measuring self-rated health change after surgery. *BMC Musculoskelet Disord.* 2003;4:11.
26. **Wada T, Kawai A, Ihara K, et al.** Construct validity of the Enneking score for measuring function in patients with malignant or aggressive benign tumours of the upper limb. *J Bone Joint Surg Br.* 2007;89-B(5):659-663.
27. **van Niekerk AH, Birkholtz FF, de Lange P, Tetsworth K, Hohmann E.** Circular external fixation and cemented PMMA spacers for the treatment of complex tibial fractures and infected nonunions with segmental bone loss. *J Orthop Surg (Hong Kong).* 2017;25(2):2309499017716242.
28. **Tu YK, Lin CH, Su JI, Hsu DT, Chen RJ.** Unreamed interlocking nail versus external fixator for open type III tibia fractures. *J Trauma.* 1995;39(2):361-367.
29. **Kwasnicki RM, Hettiaratchy S, Jarchi D, et al.** Assessing functional mobility after lower limb reconstruction: a psychometric evaluation of a sensor-based mobility score. *Ann Surg.* 2015;261(4):800-806.

30. **Hofmann AA, Tkach TK, Evanich CJ, Camargo MP, Zhang Y.** Patellar component medialization in total knee arthroplasty. *J Arthroplasty*. 1997;12(2):155-160.
31. **Evanich CJ, Tkach TK, von Glinski S, Camargo MP, Hofmann AA.** 6- to 10-year experience using countersunk metal-backed patellas. *J Arthroplasty*. 1997;12(2):149-154.
32. **Baker SP, O'Neill B, Haddon W Jr, Long WB.** The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care. *J Trauma*. 1974;14(3):187-196.
33. **Beverland DE, Rutherford WH.** An assessment of the validity of the injury severity score when applied to gunshot wounds. *Injury*. 1983;15(1):19-22.
34. **Johansen K, Daines M, Howey T, Helfet D, Hansen ST Jr.** Objective criteria accurately predict amputation following lower extremity trauma. *J Trauma*. 1990;30(5):568-72; discussion 572-573.
35. **Togawa S, Yamami N, Nakayama H, Mano Y, Ikegami K, Ozeki S.** The validity of the mangled extremity severity score in the assessment of upper limb injuries. *J Bone Joint Surg Br*. 2005;87-B(11):1516-1519.
36. **Merchant TC, Dietz FR.** Long-term follow-up after fractures of the tibial and fibular shafts. *J Bone Joint Surg Am*. 1989;71-A(4):599-606.
37. **Johner R, Wruhs O.** Classification of tibial shaft fractures and correlation with results after rigid internal fixation. *Clin Orthop Relat Res*. 1983;(178):7-25.
38. **Binkley JM, Stratford PW, Lott SA, Riddle DL.** The Lower Extremity Functional Scale (LEFS): scale development, measurement properties, and clinical application. North American Orthopaedic Rehabilitation Research Network. *Phys Ther*. 1999;79(4):371-383.
39. **Tegner Y, Lysholm J.** Rating systems in the evaluation of knee ligament injuries. *Clin Orthop Relat Res*. 1985;(198):43-49.
40. **Sanders R, Fortin P, DiPasquale T, Walling A.** Operative treatment in 120 displaced intraarticular calcaneal fractures. Results using a prognostic computed tomography scan classification. *Clin Orthop Relat Res*. 1993;(290):87-95.
41. **Mazur JM, Schwartz E, Simon SR.** Ankle arthrodesis. Long-term follow-up with gait analysis. *J Bone Joint Surg Am*. 1979;61-A(7):964-975.
42. **Martin DP, Engelberg R, Agel J, Snapp D, Swiontkowski MF.** Development of a musculoskeletal extremity health status instrument: the Musculoskeletal Function Assessment instrument. *J Orthop Res*. 1996;14(2):173-181.
43. **Engelberg R, Martin DP, Agel J, Swiontkowski MF.** Musculoskeletal function assessment: reference values for patient and non-patient samples. *J Orthop Res*. 1999;17(1):101-109.
44. **Engelberg R, Martin DP, Agel J, Obremsky W, Coronado G, Swiontkowski MF.** Musculoskeletal Function Assessment instrument: criterion and construct validity. *J Orthop Res*. 1996;14(2):182-192.
45. **Schmidt-Rohlfing B, Pfeifer R, Kaneshige J, et al.** Scoring systems for outcome after knee injuries. *Injury*. 2011;42(3):271-275.

46. **Neer CS 2nd, Grantham SA, Shelton ML.** Supracondylar fracture of the adult femur. A study of one hundred and ten cases. *J Bone Joint Surg Am.* 1967;49-A(4):591-613.
47. **Balke B.** A simple field test for the assessment of physical fitness. REP 63-6. *Rep Civ Aeromed Res Inst US.* 1963:1-8.
48. **Nightingale EJ, Pourkazemi F, Hiller CE.** Systematic review of timed stair tests. *J Rehabil Res Dev.* 2014;51(3):335-350.
49. **Zaino CA, Marchese VG, Westcott SL.** Timed up and down stairs test: preliminary reliability and validity of a new measure of functional mobility. *Pediatr Phys Ther.* 2004;16(2):90-98.
50. **Shumway-Cook A, Brauer S, Woollacott M.** Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test. *Phys Ther.* 2000;80(9):896-903.
51. **Hallstrand TS, Boitano LJ, Johnson WC, Spada CA, Hayes JG, Raghu G.** The timed walk test as a measure of severity and survival in idiopathic pulmonary fibrosis. *Eur Respir J.* 2005;25(1):96-103.
52. **Keller S, Bann CM, Dodd SL, Schein J, Mendoza TR, Cleeland CS.** Validity of the brief pain inventory for use in documenting the outcomes of patients with noncancer pain. *Clin J Pain.* 2004;20(5):309-318.
53. **Price DD, McGrath PA, Rafii A, Buckingham B.** The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. *Pain.* 1983;17(1):45-56.
54. **Olerud C, Molander H.** A scoring scale for symptom evaluation after ankle fracture. *Arch Orthop Trauma Surg (1978).* 1984;103(3):190-194.
55. **Paley D, Maar DC.** Ilizarov bone transport treatment for tibial defects. *J Orthop Trauma.* 2000;14(2):76-85.
56. **Paley D, Catagni MA, Argnani F, Villa A, Benedetti GB, Cattaneo R.** Ilizarov treatment of tibial nonunions with bone loss. *Clin Orthop Relat Res.* 1989;(241):146-165.
57. **Ferchaud F, Rony L, Ducellier F, et al.** Reconstruction of large diaphyseal bone defect by simplified bone transport over nail technique: A 7-case series. *Orthop Traumatol Surg Res.* 2017;103(7):1131-1136.
58. **Sanders R, Swiontkowski M, Rosen H, Helfet D.** Double-plating of comminuted, unstable fractures of the distal part of the femur. *J Bone Joint Surg Am.* 1991;73-A(3):341-346.
59. **Jain D, Arora R, Garg R, Mahindra P, Selhi HS.** Functional outcome of open distal femoral fractures managed with lateral locking plates. *Int Orthop.* 2020;44(4):725-733.
60. **Puno RM, Grossfeld SL, Henry SL, Seligson D, Harkess J, Tsai TM.** Functional outcome of patients with salvageable limbs with grades III-B and III-C open fractures of the tibia. *Microsurgery.* 1996;17(3):167-173.
61. **Swiontkowski MF, Engelberg R, Martin DP, Agel J.** Short musculoskeletal function assessment questionnaire: validity, reliability, and responsiveness. *J Bone Joint Surg Am.* 1999;81-A(9):1245-1260.

62. **Insall JN, Dorr LD, Scott RD, Scott WN.** Rationale of the Knee Society clinical rating system. *Clin Orthop Relat Res.* 1989;(248):13-14.
63. **Stüber J, Zech S, Bay R, Qazzaz A, Richter M.** Normative data of the Visual Analogue Scale Foot and Ankle (VAS FA) for pathological conditions. *Foot Ankle Surg.* 2011;17(3):166-172.
64. **Richter M, Zech S, Geerling J, Frink M, Knobloch K, Krettek C.** A new foot and ankle outcome score: Questionnaire based, subjective, Visual-Analogue-Scale, validated and computerized. *Foot Ankle Surg.* 2006;12(4):191-199.
65. **Compston A.** Aids to the investigation of peripheral nerve injuries. Medical Research Council: Nerve Injuries Research Committee. His Majesty's Stationery Office: 1942; pp. 48 (iii) and 74 figures and 7 diagrams; with aids to the examination of the peripheral nervous system. By Michael O'Brien for the Guarantors of Brain. Saunders Elsevier: 2010; pp. [8] 64 and 94 Figures. *Brain.* 2010;133(10):2838-2844.
66. **Kroenke K, Spitzer RL, Williams JB.** The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med.* 2001;16(9):606-613.
67. **Blanchard EB, Jones-Alexander J, Buckley TC, Forneris CA.** Psychometric properties of the PTSD Checklist (PCL). *Behav Res Ther.* 1996;34(8):669-673.
68. **Hocine A, Defrance P, Lalmand J, Delcour C, Biston P, Piagnerelli M.** Predictive value of the RIFLE urine output criteria on contrast-induced nephropathy in critically ill patients. *BMC Nephrol.* 2016;17:36.