RESEARCH   **Open Access**

Comparison of the responsiveness of the Foot Health Status Questionnaire and the Manchester Foot Pain and Disability Index in older people

Hylton B Menz[1,2*], Maria Auhl[1,2], Sonja Ristevski[1,2], Nicoletta Frescos[2] and Shannon E Munteanu[1,2]

Abstract

Background: In recent years, several questionnaires have been developed for the assessment of foot health and its impact on quality of life. In order for these tools to be useful outcome measures in clinical trials, their ability to detect change over time (responsiveness) needs to be determined. Therefore, the aim of this study was to assess the responsiveness of two commonly-used questionnaires in older people with foot pain.

Methods: Participants (n = 59; 24 women and 35 men, mean age [SD] 82.3 [7.8] years) allocated to the intervention arm of a randomised controlled trial assessing the effectiveness of extra-depth footwear compared to usual care completed the Foot Health Status Questionnaire (FHSQ) and Manchester Foot Pain and Disability Index (MFPDI) at baseline and 16 weeks. Responsiveness of the FHSQ subscales (pain, function, footwear and general foot health) and MFPDI subscales (pain, functional limitation and concern about appearance) was determined using (i) paired t-tests, (ii) Cohen's d, (iii) the standardised response mean (SRM), and (iv) the Guyatt index.

Results: Overall, the FHSQ pain subscale exhibited the highest responsiveness, as evidenced by a highly significant paired t-test (p <0.001), Cohen's d = 0.63 (medium effect size), SRM = 0.50 (medium effect size) and Guyatt index = 1.70 (huge effect size). The next most responsive measure was the FHSQ function subscale, as evidenced by a borderline paired t-test (p = 0.050), Cohen's d = 0.37 (small effect size), SRM = 0.26 (small effect size) and GI = 1.22 (very large effect size). The FHSQ footwear, FHSQ general foot health and MFPDI pain, functional limitation and concern about appearance subscales demonstrated lower responsiveness, with negligible to medium effect sizes.

Conclusion: The FHSQ pain and function subscales were most responsive to change in older people with foot pain receiving a footwear intervention. These findings provide useful information to guide researchers in selecting appropriate outcome measures for use in future clinical trials of foot disorders.

Keywords: Foot, Health status, Questionnaires, Responsiveness

Background

Foot pain affects one in four people over the age of 45 years [1] and has a significant impact on mobility [2,3] and quality of life [4,5]. In recent years, several patient-reported outcome measures have been developed to assess the severity and impact of foot pain in clinical practice, epidemiological studies and clinical trials [6-11]. In order to provide useful information, these questionnaires need to be both valid (i.e. they actually

measure what they are intended to measure) and reproducible (i.e. they are able to produce the same scores in identical conditions on different occasions). In addition, to be useful outcome measures in clinical trials, such tools need to be capable of detecting changes in health status over time, a construct commonly referred to as 'responsiveness' [12-16].

Broadly speaking, there are two main approaches for assessing the responsiveness of an outcome measure, most commonly referred to as *anchor-based* and *distribution-based* [17]. *Anchor-based* approaches compare interval changes in outcome measure scores to a dichotomised 'global' rating of change score (using a question

* Correspondence: h.menz@latrobe.edu.au
[1]Lower Extremity and Gait Studies Program, School of Allied Health, La Trobe University, Bundoora 3086, Victoria, Australia
[2]Department of Podiatry, School of Allied Health, La Trobe University, Bundoora 3086, Victoria, Australia

such as "Overall, how has your condition changed as a result of your treatment?"). The outcome measure score that corresponds to a meaningful change (generally defined as a response of "somewhat better" or above) is considered to be the smallest difference which participants perceive as beneficial, and is termed the minimal important difference, or MID [12]. The MID can then be used as a benchmark for interpreting the effectiveness of an intervention. *Distribution-based* approaches involve examining the statistical distribution of change scores for interval outcome measures (for either between-group or within-group change over time comparisons), and applying a range of statistics to calculate effect sizes [18]. The smaller the MID or larger the effect size, the more responsive an outcome measure is considered to be.

Recent reviews of measures of foot function, foot health and foot pain indicate that few have undergone adequate evaluation of responsiveness, thereby limiting their use in clinical trials [6-11]. Two of the most commonly used and most extensively validated measures of foot pain and disability are the Foot Health Status Questionnaire (FHSQ) [19] and the Manchester Foot Pain and Disability Index (MFPDI) [20], but neither has undergone detailed analysis of responsiveness. In a study of people with plantar heel pain receiving foot orthoses, the FHSQ was shown to be more responsive than the Foot Function Index (FFI), based on the observation of significant improvements in all four subscales of the FHSQ but only two out of three subscales of the FFI [21]. No responsiveness data have been published for the English language version of the MFPDI, although a modification of the MFPDI – the Manchester-Oxford Foot Questionnaire – has been shown to be responsive to improvements in foot health status following hallux valgus surgery [22], and a recent study concluded that a Dutch version of the MFPDI demonstrated only moderate responsiveness [23].

Given the increasing use of the FHSQ and MFPDI in foot and ankle research, there is a need for a more detailed evaluation of responsiveness of these instruments in order to determine whether it is appropriate to employ them as outcome measures in clinical trials of interventions for foot disorders. Therefore, as part of a randomised controlled trial assessing the effectiveness of extra-depth footwear in older people with foot pain [24,25], we compared the responsiveness of the FHSQ and MFPDI subscales, using a range of recommended statistical approaches [12-16].

Methods
Study design
This study was undertaken as part of a larger randomised controlled trial, the details of which have been published previously [24,25]. Briefly, the trial was a parallel-group randomised controlled trial design with a 16-week follow-up period, with participants randomly allocated to either a 'usual care' control group or the intervention group. The intervention group was provided with off-the-shelf footwear at the baseline assessment, and the data obtained from this group form the basis of the current study.

Participants
Participants residing in Melbourne, Victoria, Australia were recruited from the Department of Veterans' Affairs (DVA) database between October 2012 and May 2013. To be eligible to be included in the study, veterans needed to: (i) be aged 65 years or over; (ii) be a current DVA Gold Card client not eligible for medical grade footwear; (iii) have received podiatry treatment on at least three occasions in the past five years; (iv) have disabling foot pain, using the case definition of the MFPDI [20] proposed by Roddy *et al.* [26]; (v) have persistent foot pain, defined as foot pain present for at least 12 weeks, and; (vi) be capable of understanding the English language in verbal and written form. Participants were deemed ineligible for inclusion if they: (i) were currently residing in a high level care residential aged care facility; (ii) had diabetes and a history of foot ulceration (or current foot ulceration) or diabetic peripheral neuropathy (diagnosed with the 5.07 Semmes-Weinstein monofilament, using the International Working Group on the Diabetic Foot protocol [27]); (iii) had a neurodegenerative disorder (e.g. Parkinson's disease); (iv) had a lower limb or partial foot amputation (although single toe amputations will be permitted); (v) had been prescribed contoured foot orthoses within the past 3 months; (vi) were currently wearing the intervention footwear, or; (vii) had cognitive impairment (defined as a score of <7 on the Short Portable Mental Status Questionnaire [28]).

The Australian DVA Human Research Ethics Committee provided ethical approval (approval number E012/005[5.1]) and the La Trobe University Human Ethics Committee formally accepted this approval (E012/004). All participants provided written informed consent prior to enrolment.

Intervention
The intervention group was provided with off-the-shelf, extra-depth footwear (Dr Comfort®, Vasyli Medical, Queensland, Australia). Men received the *Brian* style and women received the *Annie* style. Both styles were available in three width fittings (medium, wide, extra-wide) and featured a stretchable Lycra® upper with Velcro® closure and a choice of two removable insoles (a flat, 4 mm foam insole or a cushioning insole with a contoured heel cup, 7 mm thick under the forefoot and 15 mm thick under the heel). The shoes were lightweight (ranging from approximately 200 to 500 gm, depending on size) and meet

all commonly-used criteria for appropriate footwear for older people with foot problems [29-33]. The assessors measured participants' feet using a Brannock Device® (Brannock Device Co, Inc., Liverpool, New York, USA) to ensure appropriate length and width, using the fitting protocol recommended by the footwear manufacturer [34]. Intervention group participants who wore flat insoles (or had been wearing contoured foot orthoses for more than 3 months) in their current footwear were permitted to wear them in their study footwear, provided that the fit of the shoes was appropriate. The accuracy of the shoe fitting procedure was evaluated by comparing total length, forefoot width and forefoot girth measurements of the foot obtained with a three-dimensional scanner to the corresponding last measurements of the allocated shoe size provided by the manufacturer, and participants reported the shoes to be well fitted and comfortable [35].

Foot Health Status Questionnaire (FHSQ)

The FHSQ consists of 13 questions reflecting four foot health-related domains: pain (4 questions), function (4 questions), footwear (3 questions), and general foot health (2 questions) [19]. Each question (item) is scored on a 5-point Likert scale, and individual item scores are then re-coded, tabulated, and finally transformed to a scale ranging from 0 (indicating poorest foot health) to 100 (indicating best foot health) for each of the 4 domains. The FHSQ demonstrates a high degree of content, criterion, and construct validity and high retest reliability, and has been used as an outcome measure in clinical trials for a range of foot disorders [9]. The FHSQ was measured at baseline and at 4, 8, 12 and 16 weeks. For the purpose of the analysis in this study, only the baseline and 16 week scores were used.

Manchester Foot Pain and Disability Index (MFPDI)

The MFPDI (Isis Innovation Ltd., Oxford, UK) consists of 19 statements prefaced by the phrase 'Because of pain in my feet', formalised under three subscales: functional limitation (10 items), pain intensity (five items) and personal appearance (two items). The remaining two items are concerned with difficulties in performing work or leisure activities, which are omitted if the respondent is of retirement age. Each of the 17 statements has a three-category response: 'none of the time' (score = 0), 'on some days' (score = 1) and 'on most/every day(s)' (score = 2) [20]. The MFPDI has primarily been used in epidemiological studies, but has also been employed as an outcome measure in clinical trials [36,37]. The MFPDI was measured at baseline and at 4, 8, 12 and 16 weeks. For the purpose of the analysis in this study, only the baseline and 16 week scores were used. To express the MFPDI as a linear, interval scale, Rasch-transformed scores calculated from the original MFPDI dataset by Gijon-Nogueron *et al.* [38] were used.

Statistical analysis

All data were normally distributed based on observations of linear normal Q-Q plots and skewness values of between −1 to +1. To evaluate the responsiveness of the FHSQ and MFPDI subscales, four different effect size statistics were applied [12-16]:

Paired t-test

The paired *t*-test was used to test the null hypothesis that there was no change in the mean scores from baseline to the 16 week follow-up. This was calculated using the formula:

$$t_0 = \frac{\bar{D}_x}{SD(D_x)/\sqrt{n}}$$

...where $\bar{D}_x$ = mean change scores between baseline and follow-up, $SD(D_x)$ = standard deviation in change scores from baseline to 16 week follow-up, and n = sample size.

Cohen's d

Also referred to as the standardised effect size, Cohen's *d* was calculated as the mean change scores between baseline and 16 week follow-up, divided by the standard deviation of the baseline scores [39]:

$$d = \frac{\bar{D}_x}{SD(X)}$$

Standardised response mean

The standardised response mean (SRM), also known as the responsiveness treatment coefficient [40], was calculated as the mean change scores between baseline and 16 week follow-up divided by the standard deviation of the differences between the baseline and 16 week follow-up scores:

$$SRM = \frac{\bar{D}_x}{SD(D_x)}$$

Guyatt index

The Guyatt index (GI) [12] represents the magnitude and variability in change scores for an outcome measure relative to the minimal important difference of the measure, and was calculated as:

$$GI = \Delta_x / \sqrt{2 * MSE_x}$$

...where Δ_x = the minimal important difference of the measure and MSE_x = the standard deviation of the individual change scores. In order to determine the minimal important difference for each outcome measure subscale, participants' perceptions of the overall treatment effect were documented at the 16 week follow-up using

the question "Overall, how has your foot pain changed since the start of the study?", with a 5-point Likert scale response ("marked worsening", "moderate worsening", "same", "moderate improvement" or "marked improvement"). This scale was dichotomised, with a positive outcome defined as moderate or marked improvement [24]. The minimal important difference for each measure was then calculated as the mean change score in participants who improved minus the mean change score in participants who did not improve or whose symptoms worsened.

To aid interpretation of the effect size statistics, the following benchmarks were used: negligible effect size (<0.15), small effect size (≥0.15 and <0.40), medium effect size (≥0.40 and <0.75), large effect size (≥0.76 and <1.10), very large effect size (≥1.10 and <1.45) and huge effect size (>1.45) [41].

Results

Participant characteristics

Postal invitations were sent to 2,457 DVA clients and 341 were screened for eligibility by telephone. Of these, 121 underwent baseline screening and were randomised: 61 into the control group and 60 into the intervention group. Of the 60 participants allocated to the intervention group, one participant withdrew consent, leaving 59 participants with complete data. Characteristics of the sample are provided in Table 1.

Responsiveness

Means and standard deviations for the FHSQ and MFPDI subscales are shown in Table 2. The four responsiveness statistics for the FHSQ and MFPDI subscales are shown in Table 3. Overall, the FHSQ pain subscale

Table 1 Participant characteristics (n = 59)

Age, mean (SD) years	82.3 (7.8)
Sex, n (%) female	24 (40.7)
Body mass index, mean (SD) kg/m^2	29.2 (5.4)
Major medical conditions, n (%)	
Diabetes	8 (13.6)
Heart disease	19 (32.2)
Osteoarthritis	48 (81.4)
Rheumatoid arthritis	3 (5.1)
Stroke	8 (13.6)
Use of >4 medications	50 (84.7)
Foot problems	
Foot pain duration, mean (SD) years	13.1 (14.6)
Hallux valgus, n (%)	20 (33.9)
Lesser toe deformity, n (%)	52 (88.1)
Keratotic lesions on toes, n (%)	13 (22.0)
Plantar keratotic lesions, n (%)	37 (62.7)

Table 2 Mean (SD) scores for each outcome measure subscale at baseline and 16 week follow-up

	Baseline mean (SD)	16 week follow-up mean (SD)
FHSQ		
Pain	54.8 (20.1)	67.4 (23.1)
Function	54.1 (22.8)	62.6 (27.9)
Footwear	30.5 (25.4)	28.2 (23.5)
General foot health	34.4 (24.4)	41.4 (26.4)
MFPDI		
Pain	5.0 (1.6)	4.6 (2.1)
Functional limitation	10.5 (2.3)	9.4 (3.3)
Concern about appearance	1.1 (1.2)	1.0 (1.3)

FHSQ: Foot Health Status Questionnaire, MFPDI: Manchester Foot Pain and Disability Index.

exhibited the highest responsiveness, as evidenced by a highly significant paired *t*-test (p <0.001), Cohen's d = 0.63 (medium effect size), SRM = 0.50 (medium effect size) and GI = 1.70 (huge effect size). The next most responsive measure was the FHSQ function subscale, as evidenced by a borderline paired *t*-test (p = 0.050), Cohen's d = 0.37 (small effect size), SRM = 0.26 (small effect size) and Guyatt index = 1.22 (very large effect size). The FHSQ footwear, FHSQ general foot health and MFPDI pain, functional limitation and concern about appearance subscales demonstrated lower responsiveness, with negligible to medium effect sizes.

Discussion

The aim of this study was to evaluate the responsiveness of two commonly used measures of foot pain and disability: the Foot Health Status Questionnaire (FHSQ) [19] and the Manchester Foot Pain and Disability Index (MFPDI) [20]. To do this, we applied four of the most widely used responsiveness statistics to FHSQ and MFPDI subscale data obtained at baseline and at 16 weeks of follow-up from a clinical trial of off-the-shelf footwear for reducing foot pain in older people [24,25]. Overall, the FHSQ pain subscale exhibited the highest responsiveness, as evidenced by a highly significant paired *t*-test and effect sizes ranging from medium to huge. The next most responsive measure was the FHSQ function subscale (borderline paired *t*-test and effect sizes ranging from small to very large). Based on these findings, it would appear that the FHSQ is preferable to the MFPDI as an outcome measure in clinical trials evaluating the effectiveness of interventions in reducing foot pain and improving foot function in older people.

The FHSQ footwear and MFPDI concern about appearance subscales performed poorly in this analysis, with negligible to small effect sizes. However, these subscales are not particularly useful in the context of a trial

Table 3 Responsiveness of the outcome measure subscales

	Paired *t*-test		Cohen's *d*		SRM		GI	
	t	*p*	Value	Interpretation	Value	Interpretation	Value	Interpretation
FHSQ								
Pain	−3.83	<0.001	0.63	Medium	0.50	Medium	1.70	Huge
Function	−2.01	0.050	0.37	Small	0.26	Small	1.22	Very large
Footwear	0.89	0.378	0.09	Negligible	0.12	Negligible	0.21	Small
General foot health	−2.06	0.044	0.29	Small	0.27	Small	0.68	Medium
MFPDI								
Pain	1.61	0.114	0.26	Small	0.21	Small	0.51	Medium
Functional limitation	2.62	0.011	0.46	Medium	0.34	Small	0.41	Medium
Concern about appearance	0.61	0.542	0.09	Negligible	0.08	Negligible	0.06	Negligible

FHSQ: Foot Health Status Questionnaire, MFPDI: Manchester Foot Pain and Disability Index, SRM: standardised response mean, GI: Guyatt index.

involving a standardised footwear intervention, as the FHSQ footwear subscale items reflect difficulty with obtaining suitable footwear, and the MFPDI concern about appearance subscale items focus on participants' self-consciousness regarding the appearance of their feet and shoes. These two subscales were included in the analysis for the sake of completeness, as the FHSQ and MFPDI are generally administered in their entirety rather than as selected subscales. The poor responsiveness reported here is therefore not a useful indicator of the potential value of these subscales when applied to other interventions. For example, Bennett *et al.* [42] have demonstrated significant improvements in the FHSQ footwear subscale following foot surgery, which is likely due to the surgery allowing a wider range of shoes to be worn postoperatively.

Our observation of the limited responsiveness of the MFPDI is consistent with van der Zwaard *et al.* [23], who reported that a Dutch translation of the MFPDI was only moderately responsive to change in people aged 50 years or over who were enrolled in a randomised trial of treatment for forefoot pain. The authors suggested several reasons for this, including: (i) the three level response options ('none of the time', on some days' and 'on most/every day/s') are too widely spaced, (ii) pain intensity is not directly addressed, and (iii) the concern about appearance subscale (consisting of only two items) has a large floor effect. In developing the Manchester-Oxford Foot Questionnaire (a modification of the MFPDI for use in foot surgery), Dawson *et al.* [43] addressed many of these issues by increasing the response categories from three to five, adding a pain severity item, and combining the concern about appearance and ability to undertake social, recreational and work activities items into a separate construct referred to as 'social interaction'. The high responsiveness of this amended scale in patients undergoing hallux valgus surgery [22] suggests that there may be some scope for improving the MFPDI as an outcome measure.

When interpreting these findings, it should be noted that there is currently no accepted gold standard approach for assessing responsiveness of outcome measures, and that each statistical approach has limitations (for a detailed discussion, see Husted *et al.* [13] and Revicki *et al.* [16]). Paired *t*-tests provide an indication of the statistical significance of the observed change in the outcome measure scores, but this is influenced not only by the magnitude of change, but also the sample size (i.e. larger sample sizes are more likely to detect statistically significant differences). Cohen's *d* and the standardised response mean are influenced by the variability of the denominator (baseline scores for Cohen's *d* and change scores for standardised response mean), so higher variability in the denominator will result in smaller effect sizes. Finally, although the Guyatt index is considered by some to be the most appropriate effect size statistic [13,44], it requires the calculation of a minimal important difference, which may vary across different populations, conditions and interventions [16]. For this reason, we determined the minimal important difference of the FHSQ subscales by dichotomising a 5-point Likert scale response of perceived overall improvement (i.e. a positive outcome defined as moderate or marked improvement) and calculating mean change scores from our data, rather than using minimal important difference scores calculated from people with heel pain reported by Landorf *et al.* [45]. As no minimally important difference scores have been reported for the MFPDI, we used the same approach for each of the MFPDI subscales, which also allowed us to make direct comparisons between the two outcome measures. However, there is currently no consensus regarding the most appropriate question or number of response levels in determining the anchor used to define the minimal important difference, and anchor-based approaches have limited discriminative ability in trials where most participants report improvement in their condition [16].

Despite these limitations, and the fact that each test calculates the magnitude of the effect size in different ways, the four statistics we used resulted in a reasonably consistent pattern of responsiveness across the outcome measure subscales. We can therefore be more confident of the superiority of the FHSQ pain and function subscales using this combined approach rather than using one statistical test alone. Nevertheless, this approach only addresses 'internal' responsiveness (the ability of a measure to change over time), not 'external' responsiveness, which Husted *et al.* [13] have defined as the extent to which changes in a measure over time relate to a corresponding change in an established reference measure of health status. We were unable to evaluate external responsiveness in this study due to the absence of an appropriate reference measure for comparison. Although we collected Short Form 12 data from this sample, generic health outcome measures are generally less responsive than condition-specific measures and are therefore not considered to be suitable reference standards [6,22]. Outcome measures more directly related to mobility and physical function in older people, such as the disability index of the Health Assessment Questionnaire [46], may be more appropriate reference standards for future evaluation of external responsiveness.

Our findings provide further support for the continued use of the FHSQ as an outcome measure in clinical trials of foot disorders. However, although the recoding of FHSQ response from a Likert scale to a 100-point scale enables the FHSQ subscales to be expressed as interval data, the FHSQ has so far not undergone Rasch analysis – a statistical technique which evaluates whether overall scores summed from ordinal items can be considered to be linear, interval-level variables [47]. Such an analysis is necessary to confirm whether it is indeed appropriate to analyse FHSQ subscales using parametric statistical approaches.

In summary, this study has shown that the FHSQ pain and function subscales are most responsive to change over time in older people receiving a footwear intervention to alleviate foot pain. The FHSQ footwear and general foot health subscales and the three subscales of the MFPDI exhibited lower responsiveness, so may not be appropriate outcome measures in this population. Further research is required to determine the internal responsiveness of the FHSQ as an outcome measure in trials of other foot disorders, interventions and clinical populations, and to evaluate external responsiveness against an appropriate reference standard.

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
HBM and NF conceived the idea and obtained funding for the study. HBM, NF and SEM designed the study protocol. MA and SR collected and entered the data. HBM conducted the statistical analysis and drafted the manuscript.

All authors assisted with the writing of the manuscript, and read and approved the final manuscript.

References
1. Thomas MJ, Roddy E, Zhang W, Menz HB, Hannan M, Peat G: **The population prevalence of foot and ankle pain in middle and old age: a systematic review.** *Pain* 2011, **152**:2870–2880.
2. Benvenuti F, Ferrucci L, Guralnik JM, Gangemi S, Baroni A: **Foot pain and disability in older persons: an epidemiologic survey.** *J Am Geriatr Soc* 1995, **43**:479–484.
3. Menz HB, Dufour AB, Casey VA, Riskowski JL, McLean RR, Katz P, Hannan MT: **Foot pain and mobility limitations in older adults: the Framingham foot study.** *J Gerontol A Biol Sci Med Sci* 2013, **68**:1281–1285.
4. Menz HB, Tiedemann A, Kwan MMS, Plumb K, Lord SR: **Foot pain in community-dwelling older people: an evaluation of the Manchester foot pain and disability index.** *Rheumatology* 2006, **45**:863–867.
5. Hill CL, Gill T, Menz HB, Taylor AW: **Prevalence and correlates of foot pain in a population-based study: the North West Adelaide health study.** *J Foot Ankle Res* 2008, **1**:2.
6. Button G, Pinney S: **A meta-analysis of outcome rating scales in foot and ankle surgery: is there a valid, reliable, and responsive system?** *Foot Ankle Int* 2004, **25**(8):521–525.
7. Martin RL, Irrgang JJ: **A survey of self-reported outcome instruments for the foot and ankle.** *J Orthop Sports Phys Ther* 2007, **37**:72–84.
8. Walmsley S, Williams AE, Ravey M, Graham A: **The rheumatoid foot: a systematic literature review of patient-reported outcome measures.** *J Foot Ankle Res* 2010, **3**:12.
9. Riskowski JL, Hagedorn TJ, Hannan MT: **Measures of foot function, foot health, and foot pain.** *Arthritis Care Res* 2011, **63**:S229–S239.
10. Hunt KJ, Hurwitt D: **Use of patient-reported outcome measures in foot and ankle research.** *TJ Bone Joint Surg Am* 2013, **95**:e118. 111–119.
11. Shultz SOA, Ramsey O, Schmitz M, Wyatt V, Cook C: **A systematic review of outcome tools used to measure lower leg conditions.** *Int J Sports Phys Ther* 2013, **8**:838–848.
12. Guyatt G, Walter S, Norman G: **Measuring change over time: assessing the usefulness of evaluative instruments.** *J Chronic Dis* 1987, **40**:171–178.
13. Husted JA, Cook RJ, Farewell VT, Gladman DD: **Methods for assessing responsiveness: a critical review and recommendations.** *J Clin Epidemiol* 2000, **53**:459–468.
14. Beaton DE: **Understanding the relevance of measured change through studies of responsiveness.** *Spine* 2000, **25**:3192–3199.
15. Crosby RD, Kolotkin RL, Williams GR: **Defining clinically meaningful change in health-related quality of life.** *J Clin Epidemiol* 2003, **56**:395–407.
16. Revicki D, Hays RD, Cella D, Sloan J: **Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes.** *J Clin Epidemiol* 2008, **61**:102–109.
17. Lydick E, Epstein RS: **Interpretation of quality of life changes.** *Qual Life Res* 1993, **2**:221–226.
18. Kazis LE, Anderson JJ, Meenan RF: **Effect sizes for interpreting change in health status.** *Med Care* 1989, **27**:S178–S189.
19. Bennett PJ, Patterson C, Wearing S, Baglioni T: **Development and validation of a questionnaire designed to measure foot-health status.** *J Am Podiatr Med Assoc* 1998, **88**:419–428.
20. Garrow AP, Papageorgiou AC, Silman AJ, Thomas E, Jayson MIV, Macfarlane GJ: **Development and validation of a questionnaire to assess disabling foot pain.** *Pain* 2000, **85**:107–113.

21. Landorf KB, Keenan AM: **An evaluation of two foot-specific, health-related quality-of-life measuring instruments.** *Foot Ankle Int* 2002, **23**:538–546.
22. Dawson J, Doll H, Coffey J, Jenkinson C, Oxford, Birmingham F, Ankle Clinical Research G: **Responsiveness and minimally important change for the Manchester-Oxford foot questionnaire (MOXFQ) compared with AOFAS and SF-36 assessments following surgery for hallux valgus.** *Osteoarthritis Cartilage* 2007, **15**(8):918–931.
23. van der Zwaard BC, Terwee CB, Roddy E, Terluin B, van der Horst HE, Elders PJ: **Evaluation of the measurement properties of the Manchester foot pain and disability index.** *BMC Musculoskelet Disord* 2014, **15**:276.
24. Menz HB, Frescos N, Munteanu SE: **Effectiveness of off-the-shelf footwear in reducing foot pain in Australian Department of Veterans' Affairs recipients not eligible for medical grade footwear: study protocol for a randomized controlled trial.** *Trials* 2013, **14**:106.
25. Menz HB, Auhl M, Ristevski S, Frescos N, Munteanu SE: **Effectiveness of off-the-shelf, extra-depth footwear in reducing foot pain in older people: a randomized controlled trial.** *J Gerontol A Biol Sci Med Sci* 2014. doi:10.1093/gerona/glu1169.
26. Roddy E, Muller S, Thomas E: **Defining disabling foot pain in older adults: further examination of the Manchester foot pain and disability index.** *Rheumatology* 2009, **48**:992–996.
27. Apelqvist JBK, van Houtum WH, Nabuurs-Franssen MH, Schaper N: **International consensus and practical guidelines on the management and the prevention of the diabetic foot.** *Diabetes Metab Res Rev* 2000, **16**(Suppl 1):S84–S92.
28. Pfeiffer E: **A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients.** *J Am Geriatr Soc* 1975, **23**:433–441.
29. Nancarrow S: **Footwear suitability scale: a measure of shoe fit for people with diabetes.** *Australas J Podiatr Med* 1999, **33**:57–62.
30. Menz HB, Sherrington C: **The footwear assessment form: a reliable clinical tool for the evaluation of footwear characteristics of relevance to postural stability in older adults.** *Clin Rehabil* 2000, **14**:657–664.
31. Williams A: **Footwear assessment and management.** *Podiatry Now* 2006, **May:**S1–S8.
32. Healthy Footwear Guide [http://www.healthy-footwear-guide.com]
33. Barton CJ, Bonanno D, Menz HB: **Development and evaluation of a tool for the assessment of footwear characteristics.** *J Foot Ankle Res* 2009, **2**:10.
34. Dr Comfort Shoe Fitting Techniques [http://www.youtube.com/watch?v=XxkUZfZX7Ek]
35. Menz HB, Auhl M, Ristevski S, Frescos N, Munteanu SE: **Evaluation of the accuracy of shoe fitting in older people using three-dimensional foot scanning.** *J Foot Ankle Res* 2014, **7**:3.
36. Waxman R, Woodburn H, Powell M, Woodburn J, Blackburn S, Helliwell P: **FOOTSTEP: a randomized controlled trial investigating the clinical and cost effectiveness of a patient self-management program for basic foot care in the elderly.** *J Clin Epidemiol* 2003, **56**:1092–1099.
37. Spink MJ, Menz HB, Fotoohabadi MR, Wee E, Landorf KB, Hill KD, Lord SR: **Effectiveness of a multifaceted podiatry intervention to prevent falls in community dwelling older people with disabling foot pain: randomised controlled trial.** *BMJ* 2011, **342**:d3411.
38. Gijon-Nogueron G, Ndosi M, Luque-Suarez A, Alcacer-Pitarch B, Munuera PV, Garrow A, Redmond AC: **Cross-cultural adaptation and validation of the Manchester foot pain and disability index into Spanish.** *Qual Life Res* 2014, **23**(2):571–579.
39. Cohen J: *Statistical Power Analysis for the Behavioral Sciences.* 2nd edition. Hillsdale, NJ: Erlbaum; 1988.
40. Norman GR, Stratford P, Regehr G: **Methodological problems in the retrospective computation of responsiveness to change: the lesson of Cronbach.** *J Clin Epidemiol* 1997, **50**:869–879.
41. How to calculate effect sizes from published research articles: A simplified methodology [http://www.bwgriffin.com/gsu/courses/edur9131/content/Effect_Sizes_pdf5.pdf]
42. Bennett PJ, Patterson C, Dunne MP: **Health-related quality of life following podiatric surgery.** *J Am Podiatr Med Assoc* 2001, **91**:164–173.
43. Dawson J, Coffey J, Doll H, Lavis G, Cooke P, Herron M, Jenkinson C: **A patient-based questionnaire to assess outcomes of foot surgery: validation in the context of surgery for hallux valgus.** *Qual Life Res* 2006, **15**(7):1211–1222.
44. Deyo RA, Diehr P, Patrick DL: **Reproducibility and responsiveness of health status measures statistics and strategies for evaluation.** *Control Clin Trials* 1991, **12**:S142–S158.
45. Landorf KB, Radford JA, Hudson S: **Minimal Important Difference (MID) of two commonly used outcome measures for foot problems.** *J Foot Ankle Res* 2010, **3**:7.
46. Bruce B, Fries JF: **The Stanford health assessment questionnaire: dimensions and practical applications.** *Health Qual Life Outcomes* 2003, **1**:20.
47. Tennant A, Conaghan PG: **The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper?** *Arthritis Rheum* 2007, **57**:1358–1362.