Foot deformity in diabetic neuropathy. A radiobiological and biomechanical analysis
Bus, S.A.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 3

Reproducibility of foot structure measurements in neuropathic diabetic patients using magnetic resonance imaging

Sicco A. Bus¹, Mario Maas², Robert Lindeboom³

Departments of ¹Internal Medicine, ²Radiology, and ³Clinical Epidemiology and Biostatistics, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands

Submitted
Chapter 3

Abstract

Structural changes in foot anatomy are an important contributing component of altered biomechanics and development of foot ulcers in diabetic patients with peripheral neuropathy. The aim of this study was to assess the intra- and inter-observer agreement of several commonly reported foot structure measurements in the diabetic foot using magnetic resonance imaging (MRI).

Twenty-three neuropathic diabetic patients and five age-matched healthy controls underwent MRI examination of the foot. From sagittal plane images, joint configuration and plantar fat-pad thickness in the second and third rays of the forefoot were assessed twice by the same observer and once by a different observer. The degree of intrinsic muscle atrophy was scored from coronal plane images on a 5-point scale by two observers following consensus agreement.

Intraclass correlation coefficients between occasions and between observers were larger than 0.94. The mean differences (bias) and the limits of agreement (LoA = mean ± 2 SDs) were small for metatarsal-phalangeal joint angle, toe angle, and plantar fat-pad thickness (bias ≤ 0.8 degrees or 0.2 mm, LoA ≤ 3.8 degrees or 0.8 mm), but larger for the inter-phalangeal joint angles (bias ≤ 3.4 degrees, LoA ≤ 8.8 degrees). The weighted kappa between assessments of intrinsic muscle atrophy was 0.94.

The results show that plantar fat-pad thickness, intrinsic muscles atrophy, and joint configuration in the forefoot (with the exception of the inter-phalangeal joints), can be assessed reliably using magnetic resonance imaging. Magnetic resonance imaging is a very useful technique for obtaining static foot structure data of neuropathic diabetic patients, which may be used to predict biomechanical outcomes and ulceration.
Introduction

Diabetic foot ulceration is a frequently reported and serious complication in people with diabetes and peripheral neuropathy, which precedes the vast majority (85%) of lower-extremity amputations. Elevated dynamic plantar pressure has been established as a major risk factor for plantar ulceration in the diabetic foot with lack of protective sensation.

Structural changes and deformities of the foot such as claw/hammer toe deformity and thinning of soft tissue or fat pads underneath the metatarsal heads are associated with increased plantar pressures in diabetic subjects. Additionally, intrinsic muscle atrophy secondary to peripheral neuropathy may significantly affect foot function and lead to altered gait biomechanics.

Owing to its multi-planar imaging capability and inherent superiority in tissue contrast, magnetic resonance imaging (MRI) is increasingly used as tool to assess structural pathology in the foot or lower leg of diabetic patients with peripheral neuropathy. These assessments are considered important for improving our understanding of the relationship between structure and function in the foot and the pathogenesis of foot ulceration.

However, little is known about the reproducibility of the MRI assessments for this purpose. Therefore, the objective of this study was to evaluate the intra-observer and inter-observer agreement of (semi-)quantitative measurements of foot structure in neuropathic diabetic patients performed using MRI.

Methods

Subjects

Twenty-eight subjects participated in this study and underwent MRI foot examination. The study group consisted of 23 diabetic patients with peripheral sensory neuropathy (15 men, 8 women, mean age 56.8 years) and 5 age-matched healthy subjects who were included for reference purposes in the assessment of intrinsic muscle atrophy (3 men, 2 women, mean age 58.0 years). The presence of sensory neuropathy in the patients with diabetes was confirmed by (1) the inability to feel the pressure of a 10-grams Semmes-Weinstein monofilament on at least one of six plantar foot sites (chapter 5) and (2) abnormal vibration perception thresholds measured at the dorsal hallux using a Biothesiometer (Bio-Medical Instrument Company, Newbury, OH). As this research was part of a larger study.
investigating the structural and functional implications of claw/hammer toe deformity in the diabetic foot, a wide range of joint configurations were present for reproducibility assessment. Subjects with current foot ulceration, edema, or fracture, or with conditions precluding MRI examination were excluded. The medical ethics committee of the Academic Medical Center of the University of Amsterdam approved this study and written informed consent was obtained from each subject.

 Procedures

Detailed descriptions of the MRI procedures and all measurements performed, except for metatarsal-phalangeal and inter-phalangeal joint configuration, are reported elsewhere.\(^9;11\) (chapters 4 and 5) In summary, T1-weighted spin-echo sagittal and coronal plane images of the foot were collected non-weight bearing with a Siemens 1.5-Tesla Magnetom imager (Siemens, Erlangen, Germany). Sagittal plane imaging consisted of a series of 19 slices acquired between the first and fifth MTHs, with field of view of 256x256 mm and resolution matrix of 512x512 pixels. This resulted in pixel dimensions of 0.5x0.5 mm which was enhanced threefold for better visualization (i.e., 0.17x0.17 mm) using eFilm (Merge-eFilm, Milwaukee, WI). The coronal plane images were acquired between the proximal phalanx distally and the navicular bone proximally, with field of view 150x150 mm and resolution matrix of 256x256 pixels. The following foot structure measurements were performed:

![Figure 1. Sagittal plane image of the forefoot used for joint configuration assessment. Metatarsal-phalangeal (\(\alpha\)), proximal inter-phalangeal (\(\beta\)), and distal inter-phalangeal joint angles (\(\varphi\)) were measured from the bisectors of the bones located proximal and distal to these joints. Toe angle (\(\delta\)) was measured between a line parallel to the sole of the forefoot and the bisector of the proximal phalanx.](image)
1. Joint configuration of the second and/or third ray of the foot using Agfa IMPAX WEB1000 software (Agfa-Gevaert N.V., Mortsel, Belgium) (Figure 1). The metatarsal-phalangeal (MTP) joint angle was measured between the bisectors (= straight line dividing the bone into two sections with equal areas) of the metatarsal and proximal phalangeal bones. The proximal inter-phalangeal (PIP) joint angle was measured between the bisectors of the proximal phalanx and the middle phalanx. The distal inter-phalangeal (DIP) joint angle was measured between the bisectors of the middle and distal phalanges. The toe angle was defined as the angle between a line parallel to the sole of the forefoot and the bisector of the proximal phalanx.

2. Fat-pad thickness under the metatarsal heads and under the proximal phalangeal bone of the second and/or third ray using Scion Image (National Institutes of Health, Bethesda, MD). Both were measured at three proximal-to-distal locations to obtain a representative estimate of fat-pad thickness in the region (Figure 2). The average of the three measures per site was used for further analysis.

3. Intrinsic muscle atrophy in an anatomically referenced coronal plane slice through the fifth MTH was determined from printed images (Figure 3). The degree of atrophy was scored on a semi-quantitative 5-point scale with 0 representing healthy muscle or no

Figure 2. Sagittal plane image of the forefoot used for plantar fat-pad thickness assessment. Sub-MTH fat-pad thickness was measured perpendicular to the sole of the forefoot at three standardized locations between the proximal and distal borders of the metatarsal head. Sub-phalangeal fat-pad thickness was measured perpendicular to the bisector of the proximal phalanx, also at three standardized locations between the proximal border and the center of the proximal phalangeal bone.
atrophy; 1, mild atrophy; 2, moderate atrophy; 3, severe atrophy; and 4, almost no or no muscle tissue visible.

Two experienced observers (SB and MM) examined all foot images in a random order. Both observers independently assessed joint configuration and fat-pad thickness; intrinsic muscle atrophy was assessed by both observers together. The intra-observer agreement for joint configuration and fat-pad thickness was determined by comparing repeated assessments performed four weeks apart by the same observer (SB). The inter-observer agreement was determined by comparing the assessment from the second observer (MM) with the average of repeated assessments of the first observer (SB). Intrinsic muscle atrophy was assessed following consensus agreement for atrophy score between both observers meaning that for this variable only intra-observers agreement could be evaluated.

![Foot Images](image)

Figure 3. Assessment of degree of intrinsic muscle atrophy in an anatomically referenced cross-sectional image of the foot through the fifth MTH using a 5-point atrophy score. Shown are examples representing the 5 categories: score 0, healthy muscle tissue (no atrophy); score 1, mild atrophy; score 2, moderate atrophy; score 3, severe atrophy, and score 4, almost no muscle or no muscle visible.

**Statistical analysis**

Sample size determination: Five patients would be needed to detect an agreement between occasions or between observers with an anticipated acceptable correlation of 0.90 and a one-sided confidence interval of 0.05. We included more than five patients in this study to further increase the precision of the reproducibility estimates.

Reproducibility of the measurements between and within observers for joint configuration and fat-pad thickness was evaluated using intraclass correlation coefficients. The weighted kappa coefficient was used to evaluate the agreement between the ordinal scores used for intrinsic muscle atrophy. The absolute magnitude of the differences between or within observers was evaluated using the graphical method of Bland and Altman. With this method, the differences between two measurements are plotted against their average. The
Reproducibility of foot structure measurements

Mean difference represents the amount of bias or systematic error between two measurements. The limits of agreement (= mean difference ± 2 SDs) indicate how far apart both measurements are for most subjects in the sample. Five percent of the differences are expected to fall outside these limits. A zero mean difference does not mean the absence of bias over the total measurement range. The Bland and Altman scatter plots' regression lines were therefore tested for a zero slope using linear regression analysis. A zero slope indicates no systematic differences between observers or between occasions over the total range of measurement, meaning that the two assessments can be used interchangeably.

All analyses were carried out with SPSS (SPSS, Chicago, IL). The weighted kappa coefficient was calculated using StatXact 3.0.

Table 1. Intraclass correlation coefficients (ICC) for repeated assessments by the same observer (intra-observer agreement) and for assessments between observers (inter-observer agreement).

<table>
<thead>
<tr>
<th>Variable</th>
<th>intra-observer agreement</th>
<th>inter-observer agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ICC</td>
<td>ICC</td>
</tr>
<tr>
<td>Joint configuration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTP joint angle</td>
<td>0.997</td>
<td>0.998</td>
</tr>
<tr>
<td>PIP joint angle</td>
<td>0.987</td>
<td>0.968</td>
</tr>
<tr>
<td>DIP joint angle</td>
<td>0.988</td>
<td>0.943</td>
</tr>
<tr>
<td>Toe angle</td>
<td>0.997</td>
<td>0.990</td>
</tr>
<tr>
<td>Fat-pad thickness</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sub-MTH</td>
<td>0.991</td>
<td>0.979</td>
</tr>
<tr>
<td>Sub-phalangeal</td>
<td>0.978</td>
<td>0.976</td>
</tr>
<tr>
<td>Intrinsic muscle atrophy</td>
<td>0.94</td>
<td>-</td>
</tr>
</tbody>
</table>

*Weighted kappa

Results

The reproducibility coefficients summarizing the intra- and inter-observer agreement are shown in Table 1. For the quantitative measurements, the lowest agreement was for DIP joint angle where the intraclass correlation coefficients indicating intra- and inter-observer agreement were 0.958 and 0.943, respectively. Scores for intrinsic muscle atrophy (assessed by the two observers together) between occasions agreed in 74% of cases. In 6 of 23 feet, intrinsic muscle atrophy was scored one category different at the second when compared to
the first assessment. Differences of more than one category between occasions were not present. The weighted kappa was 0.94.

Joint configuration and fat-pad thickness data are shown in Table 2 for repeated assessments by the same observer and in Table 3 for assessments between observers. The mean differences between occasions for joint configuration ranged from −0.9 to −0.2 degrees (Table 2). The limits of agreement were between −5.3 degrees and +5.0 degrees. Mean fat-pad thickness measurement error did not exceed +0.1 mm, with limits of agreement ranging between −0.5 mm and +0.6 mm. The mean differences between observers for joint configuration ranged between −1.6 and +3.4 degrees (Table 3). The limits of agreement were between −8.8 degrees and +8.7 degrees. Mean fat-pad thickness differences did not exceed +0.2 mm. Limits of agreement ranged between −0.7 and +0.8 mm.

Table 2. Data for joint configuration and fat-pad thickness from repeated assessments by the same observer (intra-observer agreement).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Assessment 1</th>
<th>Assessment 2</th>
<th>99% CI of mean difference</th>
<th>Limits of agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n* Mean SD</td>
<td>n* Mean SD</td>
<td>Mean difference SD lower upper</td>
<td>lower upper</td>
</tr>
<tr>
<td>Joint configuration (*)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTP joint angle</td>
<td>21 -40.1 14.0</td>
<td>21 -40.6 13.8</td>
<td>-0.5</td>
<td>0.9</td>
</tr>
<tr>
<td>PIP joint angle</td>
<td>22 26.0 13.8</td>
<td>22 25.2 13.4</td>
<td>-0.9</td>
<td>2.1</td>
</tr>
<tr>
<td>DIP joint angle</td>
<td>20 16.9 9.2</td>
<td>20 16.8 8.3</td>
<td>-0.2</td>
<td>2.6</td>
</tr>
<tr>
<td>Toe angle</td>
<td>24 -8.8 11.7</td>
<td>24 -9.2 11.5</td>
<td>-0.4</td>
<td>0.8</td>
</tr>
<tr>
<td>Fat-pad thickness (mm)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sub-MTH</td>
<td>20 5.4 1.7</td>
<td>20 5.4 1.8</td>
<td>0.0</td>
<td>0.2</td>
</tr>
<tr>
<td>Sub-phalangeal</td>
<td>20 7.5 1.2</td>
<td>20 7.6 1.2</td>
<td>0.1</td>
<td>0.2</td>
</tr>
</tbody>
</table>

n, number of cases for comparison; CI, confidence interval

* Because of anatomical misalignment of the joints in the sampling plane in several subjects, not all foot images could be examined for joint configuration and fat-pad thickness. In some subjects, both the second and third ray was assessed.

Bland and Altman plots are shown in Figure 4 for repeated assessments and in Figure 5 for assessments between observers. The mean differences between occasions and between observers were normally distributed ($P > 0.05$, Shapiro-Wilk’s test of normality). All but
Reproducibility of foot structure measurements

one of the slopes of the scatter plots' regression lines varied between 0.01 and 0.34; the slope for sub-MTH fat pad measurement was relatively high (0.53). None of the slopes was significantly different from zero, indicating no systematic differences in scores over the total range of the measurements.

Table 3. Data for joint configuration and fat-pad thickness from assessments between observers (inter-observer agreement).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Observer 1</th>
<th>Observer 2</th>
<th>Mean difference</th>
<th>SD</th>
<th>99% CI of mean difference</th>
<th>Limits of agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Mean</td>
<td>Mean</td>
<td>SD</td>
<td>lower</td>
<td>upper</td>
</tr>
<tr>
<td>Joint configuration (*)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTP joint angle</td>
<td>12</td>
<td>-36.8</td>
<td>11.1</td>
<td>-0.3</td>
<td>1.3</td>
<td>-1.4 0.9</td>
</tr>
<tr>
<td>PIP joint angle</td>
<td>12</td>
<td>23.7</td>
<td>12.9</td>
<td>3.4*</td>
<td>2.6</td>
<td>1.0 5.7</td>
</tr>
<tr>
<td>DIP joint angle</td>
<td>12</td>
<td>17.2</td>
<td>8.2</td>
<td>-1.6</td>
<td>3.6</td>
<td>-4.9 1.8</td>
</tr>
<tr>
<td>Toe angle</td>
<td>12</td>
<td>-8.9</td>
<td>8.4</td>
<td>0.8</td>
<td>1.5</td>
<td>-0.5 2.2</td>
</tr>
<tr>
<td>Fat-pad thickness (mm)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sub-MTH</td>
<td>12</td>
<td>5.6</td>
<td>1.1</td>
<td>0.1</td>
<td>0.4</td>
<td>-0.3 0.4</td>
</tr>
<tr>
<td>Sub-phalangeal</td>
<td>12</td>
<td>7.8</td>
<td>1.2</td>
<td>0.2</td>
<td>0.3</td>
<td>-0.1 0.4</td>
</tr>
</tbody>
</table>

n, number of cases for comparison; CI, confidence interval. * P < 0.01, significantly different from zero

* A random sample of 12 cases was selected from the total pool of cases for assessment of between observer agreement.

Discussion

The results showed small mean differences with narrow confidence intervals between occasions and between observers for MTP joint and toe angle and for fat-pad thickness with no significant differences from zero present. These results imply a high degree of reproducibility for average results obtained in these assessments of forefoot anatomical structure. This means that the same observer on a different occasion or a different observer would obtain the same average results when assessing a group of toes for joint configuration or fat-pad thickness. This is an important finding for group comparative studies on forefoot structure in neuropathic diabetic subjects (chapters 5 and 7) because reproducible average estimates imply that conclusions drawn from differences found between groups are methodologically robust.
The limits of agreement give an indication of the expected variation due to measurement error and define how well different assessments are likely to agree for an individual patient, given that 95% of the observed individual differences are expected to be between the limits of agreement. For example, the results show that when measuring sub-MTH fat-pad thickness, one observer should be within 0.8 mm of the thickness measured by another observer in 95% of the cases. The limits of agreement calculated for MTP joint and toe angle and for both sub-MTH and sub-phalangeal fat-pad thickness were relatively small and we did not consider them clinically relevant. The intraclass correlation coefficients for these variables were high (>0.98), which demonstrates that these variables can be assessed in a reproducible way. Small limits of agreement are important for regression analyses or correlation coefficients computed between variables of foot structure (chapter 5) or between foot structure and function (chapter 7) because the outcomes from these analyses are determined by the scatter of individual cases.

The mean differences and the confidence intervals of the mean differences for DIP and PIP joint angles were relatively large, with a significant 3.4 degree mean difference from zero between observers for PIP joint angle. The limits of agreement for these parameters were relatively large up to about ±5 degrees between occasions and about ±9 degrees between observers. Despite high intraclass correlation coefficients (>0.94), measurement of inter-phalangeal joint angles may thus be 9 degrees different from one observer to the other. This may be unacceptable for research purposes. Whether this difference is clinically relevant requires further investigation as toe deformity is not normally classified using angular measurements, and, as a result, thresholds of inter-phalangeal joint angles that may predict further pathology are not known. The higher limits of agreement compared with the assessment of MTP joint and toe angles is presumably explained by the difficulty in accurately drawing bisectors in relatively small bones such as the middle and distal phalanges, which also seem to be more variable in shape across subjects than the metatarsal or proximal phalangeal bones. Additionally, drawing a tangent along the flat sole of a forefoot (because of minimal weight bearing) was experienced to be easier than drawing a bisector of a bone, which may explain the smaller variability for toe angle than for inter-phalangeal joint angle measurement. A larger acquisition matrix (larger than 512x512 pixels) and/or a smaller field of view used in MRI will improve visibility of the small bones and likely improve reliability in inter-phalangeal joint angle assessment. The data on the two most distal joints of the toes also demonstrates the usefulness of Bland and Altman plots in addition to intraclass correlation coefficients because they show the absolute measurement error between occasions or observers.
Intrinsic muscle atrophy scored by two observers was similar on different occasions. Only six of 23 feet were scored different during the second assessment, with only one category difference for these six cases. The weighted kappa was 0.94 showing very good intra-observer agreement. Remarkable degrees of intrinsic muscle atrophy have been shown in patients with long-standing diabetes and peripheral neuropathy, and has been suggested to play a significant role in altered gait biomechanics. (chapter 2) In these reported studies, intrinsic muscle atrophy was assessed quantitatively using special MRI imaging.
sequences. However, these techniques are complex, require special hardware and software,
and are time-consuming; an easier, more directly interpretable method of visual inspection
from standard imaging sequences is required. In this study we show that a semi-quantitative
5-point scoring system (Figure 3 may be used as guidance) can be a good alternative for
assessing intrinsic muscle atrophy in the neuropathic diabetic foot.

Figure 5. Bland-Altman plots of assessments between observers for joint configuration and fat-pad thickness. In these
plots, the differences between assessments from two observers (observer 2 – observer 1) is shown against the
average of assessments by these observers for (A) MTP joint angle, (B) toe angle, (C) PIP joint angle, (D) DIP joint
angle, (E) Sub-MTH fat-pad thickness, and (F) sub-phalangeal fat-pad thickness. The Beta coefficients of the scatter
plots’ regression lines are also shown.
Reproducibility of foot structure measurements

Measures of structural changes and deformity in the diabetic foot have been reported using different in-vivo imaging techniques, such as ultrasound\textsuperscript{11,15}, radiography\textsuperscript{13}, and computed tomography (CT).\textsuperscript{16,17,19,21} Saltzmann et al.\textsuperscript{20} determined the inter-observer variability of radiographic foot measurements in non-diabetic subjects and found a 95\% limit for MTP joint angle of 6 degrees, which was higher than the limit in the present study (<3.0 degrees). Using CT, Commean et al.\textsuperscript{14} assessed the intra-observer reliability of foot structure measurements in the diabetic patient, including MTP joint angle and plantar soft-tissue thickness. Although the bias for these parameters agreed well with our data (close to zero), the standard deviations (1 SD) of the mean differences (approximately 2 degrees and 1 mm) were higher than in our study reflecting more variability and thus lower agreement. The smaller variability found in the present study may be related to the type of imaging modality used. MRI has superior soft-tissue contrast over other imaging techniques, including radiography and CT. These properties improve the visibility of tissue boundaries and as a result may improve reproducibility of foot structure measurements.

In conclusion, anatomical foot structure related to joint configuration, fat-pad thickness and intrinsic muscle atrophy can be assessed in a reproducible manner using MRI. This shows that MRI is a useful method for obtaining structural data of the feet in diabetic patients with peripheral neuropathy, which could be used for establishing relationships between structural and functional parameters that have been shown or are suggested to be involved in the development of ulceration.
Chapter 3

References


58
Reproducibility of foot structure measurements


