Title: Benchmarking two recreational road running shoes using perceived comfort and running kinematics



STUDENT REPORT

Education: Master in Sports Technology Semester: 4th semester

Project period: 01/02-2022 - 01/06-2022

Supervisor: Pascal Madeleine

Project group: 4-04

Jens Toft Knudsen

Pages (Manuscript): 18

Pages (Worksheet): 44

Pages (Appendix 1&2): 5

Abstract:

Purpose: The purpose of this field study was to benchmark two recreational road running shoes with respect to comfort and running kinematics using a visual analogue scale (VAS) - questionnaire and inertial measurement units (IMU), respectively. It was hypothesized that higher running kinematic variability and lower peak resultant tibial acceleration (peak TA-R) would be seen for the shoe perceived the most comfortable.

Methods: Fifteen participants took part in this randomized cross-over study and ran 400 m in-field on an asphalt paved road in each of two shoe conditions. Running kinematics were measured using two IMUs. One fixated on the shoe upper and one on the medial distal tibia of the right leg to investigate relative variability (RV) of the late swing phase and peak TA-R, respectively. Five footwear comfort items were measured after the running procedure for each shoe using a 100 mm VAS-questionnaire.

Results: The statistical analysis did not reveal significant differences in comfort or running kinematics between the two shoes. Mean and median comfort scores were all > 50 mm

Conclusion: The current study showed that the two tested running shoes were characterized as equally comfortable with similar RV and peak TA-R.

Reading guide



Reference system

The American Psychological Association 7th ed. (APA) was used as reference system in this thesis. A reference will be placed immediately after a statement or in the end of a sentence, e.g.:

"Visual analogue scales (VAS) have been used to assess comfort with moderate to high reliability (Matthias et al., 2021; Menz and Bonanno, 2021)."

For the worksheet, if several lines refer to the same reference a reference will be placed after the full stop of the last sentence, e.g.:

"Regarding running injuries knee problems have shown to account for 42% including patellofemoral pain syndrome (PFPS) being the most frequent among 2002 runners. Other common injuries were; iliotibial band friction syndrome (ITBFS), meniscal injuries, tibial stress syndrome and plantar fasciitis. (Taunton, 2002)"

Report content

The content of this report is based on a manuscript and a worksheet and presented in that order. The worksheet was used to initiate the project work and document the problem background, methods, and results in depth. The manuscript is used to document the conducted study in a journal format.

Data availability

The data and custom MATLAB script supporting the findings of this report are available on request at: https://doi.org/10.5281/zenodo.6600697

Benchmarking two recreational road running shoes using perceived comfort and running kinematics

Jens T. Knudsen ^a

^aDepartment of Health Science and Technology, Sport Sciences-Performance and Technology, Aalborg University, Denmark

Academic supervisor: Professor, Pascal Madeleine

Abstract

Purpose: The purpose of this field study was to benchmark two recreational road running shoes with respect to comfort and running kinematics using a visual analogue scale (VAS) - questionnaire and inertial measurement units (IMU), respectively. It was hypothesized that higher running kinematic variability and lower peak resultant tibial acceleration (peak TA-R) would be seen for the shoe perceived the most comfortable. **Methods:** Fifteen participants took part in this randomized cross-over study and ran 400 m in-field on an asphalt paved road in each of two shoe conditions. Running kinematics were measured using two IMUs. One fixated on the shoe upper and one on the medial distal tibia of the right leg to investigate relative variability (RV) of the late swing phase and peak TA-R, respectively. Five footwear comfort items were measured after the running procedure for each shoe using a 100 mm VAS-questionnaire. **Results:** The statistical analysis did not reveal significant differences in comfort or running kinematics between the two shoes. Mean and median comfort scores were all > 50 mm **Conclusion:** The current study showed that the two tested running shoes were characterized as equally comfortable with similar RV and peak TA-R.

Keywords: Relative variability, Tibial acceleration, Visual analogue scale, Inertial measurement unit

1. Introduction

Since the 1970's recreational running has evolved into becoming the most common type of physical activity. However, overuse injuries have shown to be the main reason for runners to quit running accounting for 31 % of reported reasoning (Koplan et al., 1995; Hennig, 2011). Furthermore, a systematic review by van Gent et al. (2007) has found that 19.4 % to 79.3 % of runners sustain a lower extremity running injury every year, including knee problems (patellofemoral pain syndrome, meniscal injuries), iliotibial band friction syndrome, tibial stress syndrome and plantar fasciitis, with knee problems being most frequent (Taunton, 2002). This has resulted in a tremendous development in running footwear over the past 50 years (Sun et al., 2020) as running footwear have been suggested as an important mechanism to influence running biomechanics related to the aetiology of running injuries (Sinclair et al., 2021). Running shoe development has further focused on enhancing running performance and improving shoe comfort (Sun et al., 2020). Shoe comfort has shown to be an important factor regarding running shoes as it may enhance running performance (Wan Alsenoy et al., 2021; Menz and Bonanno, 2021), reduce lower extremity injury incidence (Mündermann et al., 2001) and is considered the most important factor regarding customers willingness to

May 31, 2022

buy a shoe (Martínez-Martínez et al., 2017). In general the trend of shoe development has twice moved back and forth between initially minimal supportive shoes to highly cushioned and supportive shoes with varying levels of midsole thickness and hardness (Sun et al., 2020). Of the different footwear characteristics, shoes that are well-fitted to the user and feature low mass, a curved rocker-sole and a softer midsole have shown to be perceived as generally more comfortable (Menz and Bonanno, 2021). Footwear characteristics can overall be categorised as minimal, traditional, and maximal running shoes (Sinclair et al., 2021), though there is little consensus of the definitions of different types of running shoes (Marchena-Rodriguez et al., 2020). Traditional shoes feature a thicker and less flexible midsole compared to minimal shoes, often with an elevated heel providing a heel-toe drop up to 12 mm (Sun et al., 2020; Marchena-Rodriguez et al., 2020). Maximal running shoes incorporate a much thicker midsole along the full length of the shoe to provide a higher level of cushioning, while not increasing the heel-toe drop (Sinclair et al., 2021).

A systematic review based on 63 biomechanical studies on running shoes, identified that running shoes with thicker midsoles can attenuate impact shocks and therefore provide a better cushioning effect though a reduction in plantar sensation may be seen compared to minimal shoes (Sun et al., 2020). Whereas minimal shoes can increase the running economy though also increase the loading of the ankle and metatarsophalangeal joints. In addition, a reduction of impact shocks has been shown for midsoles with lower hardness which might minimize the risk of running injuries related to impact (Sun et al., 2020). Studies have often included vastly different footwear conditions, including varying cushioning technologies from minimal to maximal shoes, for comparison in effort to investigate the effect of comfort or footwear conditions on different biomechanical parameters (Jordan et al., 1997; Luo et al., 2009; Dinato et al., 2015; Lam et al., 2018b; Meyer et al., 2018; Horvais et al., 2019; Yang et al., 2019). Therefore, it could prove important to assess comfort in comparison of shoes with similar characteristics, to investigate if differences can be found for more similar shoes.

Comfort as a concept is complex and multifactorial as it is affected by feelings of wellbeing, relaxation and design aesthetics (De Looze et al., 2003) but also physiological and anatomical differences (Menz and Bonanno, 2021). Therefore in order to benchmark running shoes, and thereby provide shoe recommendations, comfort and running biomechanical parameters could be used as these have shown to be important in relation to running shoes and possibly also the incidence of running related injuries. Visual analogue scales (VAS) have been used to assess comfort with moderate to high reliability (Matthias et al., 2021; Menz and Bonanno, 2021). Further, kinematic variability might be a potential way to complement comfort assessments as the measured deviations between steps in a repeating acceleration signal during gait allows exploration of biological signals (Svenningsen et al., 2020). This might be used to identify relations to injury risk (Svenningsen et al., 2020) and to give further information as to how the individual shoes affect the runner. Additionally, Meyer et al. (2018) have shown that kinematic variability is affected by perceived comfort showing a higher relative variability (RV) when wearing the most comfortable shoe compared to the least comfortable. Further, impact shocks might be a relevant parameter to include in shoe benchmarking as impact shocks during running have often been investigated due to its potential link to running injuries (Sheerin et al., 2019) and, as above mentioned, may be

reduced by thicker and softer midsoles which are generally perceived to be more comfortable.

Assessing shoes, regarding comfort and running kinematics, should preferably be performed with self-selected running speeds in-field to ensure natural running patterns and rhythmicity which might not be supported to the same extend during inside laboratory studies (Svenningsen et al., 2020; Aristizábal Pla et al., 2021). To accomplish this, inertial measurement units (IMU) can be used as these make it possible to implement biomechanical recordings over long distances and have shown to reliably estimate locomotion variations on both regular and irregular surfaces (Svenningsen et al., 2020; Aristizábal Pla et al., 2021). Further, IMUs have shown to adequately measure both kinematic variability (Svenningsen et al., 2020; Meyer et al., 2018) and peak resultant tibial accelerations (peak TA-R) as a proxy measure for impact shocks received during running (Sheerin et al., 2019).

1.1. Purpose

The purpose of this field study was to benchmark two recreational road running shoes with respect to comfort and running kinematics using a VAS-questionnaire and IMUs, respectively. It was hypothesized that higher running kinematic variability and lower peak resultant tibial acceleration would be seen for the shoe perceived the most comfortable.

2. Method

2.1. Participants

Fifteen recreational runners participated in this study including seven males and eight females (Mean \pm SD, age: 34.7 \pm 6.6 years, weight: 70.8 \pm 12.8 kg, height: 172.1 \pm 9.0 cm) with a self-reported weekly running mileage of 16.9 \pm 9.1 km. The recruitment was carried out through word of mouth and electronic flyers to local running groups and social media. All participants had a weekly running volume of between 5 to 25 km and used shoe EU size 43 or 39 for males and females respectively. Further exclusion criteria were; outside the range of 18-50 years of age, being injured in lower extremity within 3 months prior to the study and pain in lower extremities or lower back when running. All participants gave informed written consent prior to the experimental protocol. The study was performed in accordance with the Helsinki Declaration.

2.2. Shoe conditions

Two shoe conditions (Shoe 1: Nike Pegasus 38, Shoe 2: Icebug Aura RB9X) were tested in a balanced randomized order for each participant (Figure 1(A)). Shoe 1 by the brand Nike (Nike, Inc., Oregon, USA) was chosen as a benchmark shoe being rated as the best overall running shoe 2022 in the RunRepeat online running shoe guide (www.runrepeat.com) whereas Shoe 2 was a new road running shoe prototype by the brand Icebug (Icebug AB, Västra Götaland, Sweden). Both shoes were road running shoes with a neutral designed footstep and had similar shoe constructions. Further, the shoes had similar weight, midsole hardness (shore), heel-toe drop and inside dimensions in length and width (Table 1). Though the Icebug Aura RB9X (EU size 43) was 42 g (14.3 %) heavier and had a measured 3.8 shore (7.5 %) higher midsole hardness compared to the Nike Pegasus 38 (EU size 43). No further mechanical tests were performed on the two shoe conditions.

Specification	Shoe 1		Sho	be 2
Model and size (EU)	Pegasus 39	Pegasus 43	Aura 39	Aura 43
Weight (g)	252.0	294.0	258.0	336.0
Midsole hardness (shore)	51.0	49.9	50.5	54.3
Heel-toe drop (mm)	10.0	10.0	6.0	6.0
Inside length (cm)	25.1	28.1	25.1	28.1
Inside width (cm)	8.8	9.8	9.0	9.8

Table 1: Comparison of shoe characteristics of Shoe 1 (Nike Pegasus 38) and Shoe 2 (Icebug Aura RB9X) shown in size, weight, midsole hardness, heel-toe drop, inside length and inside width.

2.3. Study design

The experimental protocol of this study was carried out within 1.5 hours consisting of an initial inside laboratory protocol ($\approx 20 \text{ min}$) followed by an in-field test protocol ($\approx 70 \text{ min}$) (Figure 1(C)). Each participant gave informed written consent after receiving both written and verbal information of the purpose of the test prior to commencing the protocol.

During the inside laboratory protocol, a questionnaire regarding running history and the inclusion/exclusion criteria was answered followed by measurement of body height and mass, and foot width and length (Brannock device, The Brannock Device Company, USA). None of the participants were excluded as a result of the initial laboratory protocol.

During the in-field test protocol, the participant performed a 7-minute warmup procedure wearing own road running shoes to achieve familiarization of the running surface as per recommendations (Huang et al., 2022; Mohr et al., 2021). During the warmup the participant was instructed to find a self-selected steady and comfortable running pace. The warmup procedure was divided into two parts as the participant first ran 3.5 minutes to achieve a steady and comfortable pace. Then the last 3.5 minutes of the warmup were recorded (Garmin Forerunner 735XT) and the average pace (min/km) was used for the following running procedure. Prior to data collection the participant went through a familiarization of the calibration and running procedure, the Garmin Virtual Partner (Garmin Forerunner 735XT) as pace keeper, and the VAS-questionnaire. This included 200 m of running and assessment of own shoes. Data collected through the familiarization was discarded and not used for further analysis.

Two IMUs were used to measure the peak TA-R (Tibia IMU) and the RV of the running kinematics (Dorsum IMU) (Shimmer3, Weight: 23.6 g; Dimensions: 51 x 34 x 14 mm; Shimmer Research Ltd, Ireland). The participant was first instrumented with a firmly secured IMU on the distal medial tibia, proximal to the distal curve caused by the medial malleolus of the right leg (Tibia IMU), as per recommendations (Sheerin et al., 2019) using self-adhesive gauze (Fixomull®), see Figure 1(B1). The Tibia IMU was not removed until the protocol was completed for both shoe conditions. Thereafter the participant was blindfolded and equipped with one of two shoe conditions (balanced randomized order) with assistance from the test leader, while the participant tightened the laces as preferred. A second IMU was fixated to the upper of the right shoe (Dorsum IMU) with double-sided adhesive tape and further fixated using self-adhesive gauze on the same standardized laces for every participant (Figure 1(2a)). The Tibia IMU was configured to record accelerometer data (\pm 16 g) while the Dorsum IMU was configured to record both accelerometer (\pm 16 g) and gyroscope (\pm 2000 °/s) data. Both IMUs had a sampling frequency of 512 Hz in agreement with recommendations stating that tibial accelerations can be recorded at a sampling frequency ranging from 300 to 600 Hz (Sheerin et al., 2019). Before the blindfold was removed, the participant was instructed, and continuously reminded, to not look at the shoes during the entirety of the test. This aspect is important as blinding is reported to reduce the risk of bias of comfort assessments (Matthias et al., 2021).

For the IMU data collection of each shoe condition, the participant first completed an IMU calibration procedure followed by a 400 m run and a second IMU calibration procedure. The calibration procedure consisted of a 10-second quiet standing phase followed by three stamps with the right foot and a second 10-second quiet standing phase. This procedure was used for subsequent Dorsum IMU data frame rotation (Figure 1(2b)). The second quiet standing phase of the second calibration procedure had a duration of 60 seconds to ensure no data was lost on the IMUs. The whole data collection was performed on a predominantly flat asphalt paved track with the participant running the same 400 m for both shoe conditions using the Garmin Virtual Partner to help sustain the saved average pace from the warmup.

After the IMU data collection, the participant assessed the comfort of the current shoe condition by answering a 100 mm Visual Analogue Scale (VAS) questionnaire consisting of five footwear comfort items ('Overall comfort', 'Heel cushioning', 'forefoot cushioning', 'shoe stability' and 'forefoot flexibility'). The VAS for each item had the anchor words 'not comfortable at all' (0 mm) to 'most comfortable imaginable' (100 mm). The latter four items were derived by earlier work of Bishop et al. (2020) as meaningful in regard to footwear comfort assessment of running shoes, based on questionnaire answers from 282 recreational runners and further field-testing of 100 recreational runners.

The participant had a break of 5-7 minutes between shoe conditions, depending on the duration of equipment preparation.



Figure 1: (A): Shoe 1 (Nike Pegasus 38) and Shoe 2 (Icebug Aura RB9X) shown from side view. (B): Fixation and orientation of the Tibia IMU (1) on distal medial tibia using self-adhesive gauze and the Dorsum IMU (2a) on the shoe upper using double-sided adhesive tape on standardized laces followed by self-adhesive gauze. The z-axis for the Tibia IMU (1) points out of the page while the x-axis for the Dorsum IMU (2a) points into the page. The foot reference frame (2b) used for subsequent Dorsum IMU data rotation, following the standards of the International Society of Biomechanics (Wu and Cavanagh, 1995). (C): An overview of the experimental protocol divided into the inside laboratory protocol and the in-field protocol. The timeline and sequence are indicated by arrows.

2.4. Data processing

The data analysis and processing were conducted using MATLAB (version R2021b, Mathworks, Massachusetts, U.S). A custom script was developed for the data processing and analysis in this study. Before the data collection the IMUs were calibrated using the 9DOF Calibration Application software (version 2.10). Then prior to the following data processing, the minor deviations from 0 m/s² and 0 °/s still present after calibration, when

not affected by the gravitational pull and when stationary, were subtracted for each of the three IMU axes for the accelerometer and gyroscope data respectively.

2.4.1. Dorsum IMU

Every initial foot contact was identified within the time of the 400 m running phase following the detection method of Meyer et al. (2018), to divide the IMU data into gait cycles: 1. The z-axis acceleration signal was low pass filtered using a second order Butterworth filter with a cut-off frequency of 20 Hz for detection of peaks in the filtered signal present before foot contact. These were reflected as rapid increases in the acceleration signal (Meyer et al., 2018). 2. The raw z-axis acceleration was differentiated. 3. After each peak in the filtered signal the time index of the point where the filtered signal was first lower than the differentiated signal, indicated the initial foot contact.

As the z-axis direction of the Dorsum IMU, was opposite the one used by Meyer et al. (2018), the z-axis acceleration signal was first inverted.

Thereafter the coordinate system of the Dorsum IMU, was rotated and aligned with the anatomical coordinate system of the foot, according to the standards of the International Society of Biomechanics (Wu and Cavanagh, 1995) (Figure 1(2b)). The Dorsum IMU was rotated about its x-axis (Roll) and y-axis (Pitch) as these were most prone to tilt. The collected data from the first 10-second quiet standing phase was used to calculate the IMU tilt, Equation (1) which was subsequently used in Roll and Pitch transformation matrices, Equation (2) and (3).

$$\theta_x[rad] = \tan^{-1}\left(\frac{a_y}{a_z}\right) \; ; \; \theta_y[rad] = \tan^{-1}\left(\frac{a_x}{a_z}\right) \tag{1}$$

$$V_{rot} = (V \cdot Roll \ matrix) \cdot Pitch \ matrix \tag{2}$$

∜

$$V_{rot} = \left(\begin{bmatrix} v_1 \\ v_2 \\ v_{n\dots} \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_x) & -\sin(\theta_x) \\ 0 & \sin(\theta_x) & \cos(\theta_x) \end{bmatrix} \right) \cdot \begin{bmatrix} \cos(\theta_y) & 0 & \sin(\theta_y) \\ 0 & 1 & 0 \\ -\sin(\theta_y) & 0 & \cos(\theta_y) \end{bmatrix}$$
(3)

Of which θ_x and θ_y are the tilt angles in radians around the IMU x and y-axis, and a_x , a_y and a_z are the average measured acceleration in m/s² during the first quiet standing phase in each axis of the IMU coordinate system. Further, V_{rot} is the rotated data vector aligned with the anatomical coordinate system, V is the initial data vector and Roll and Pitch matrix are the two frame transformation matrices used for rotating the data. v_1, v_2 and v_n are elements of the initial data vector.

The following analysis and presentation of the results of the Dorsum IMU will refer to the anatomical foot coordinate system (Figure 1(2b)).

The middle part of the IMU signal from the identified foot contacts was divided into 50 gait cycles omitting the first and last gait cycles to ensure no data analysis was done on acceleration/deceleration phases during the start and end of the running phase. Then the identified gait cycles were time normalized from 0 % to 100 % (Meyer et al., 2018) consisting of the amount of data points equal to the median step duration in measured frames. The median duration was chosen to ensure the longest and shortest step duration were approximately equally affected by the normalization procedure. The mean and standard deviation (SD) were calculated for the late swing phase for each time point between 75 % - 95 % of the 50 normalized gait cycles (Meyer et al., 2018). Finally, the root mean square (RMS) of the mean values and SD values were found and used to calculate the relative variability (RV, %), Equation (4). This resulted in one RV-value for each of the rotated Dorsum IMU data vectors of each shoe condition per participant (Meyer et al., 2018).

$$Relative \ variability \ (RV)[\%] = \frac{RMS_{SD}}{RMS_{mean}}$$
(4)

Where RMS_{SD} is the RMS (m/s² or °/s) of the standard deviation values and RMS_{mean} is the RMS (m/s² or °/s) of the mean values.

2.4.2. Tibia IMU

The Tibia IMU acceleration data was band pass filtered using a second order Butterworth filter (10 - 60 Hz) as per recommendations (Sheerin et al., 2019). The method of Aubol and Milner (2020) was then used to identify initial foot contacts within the 400 m running phase. This method is recommended for field tests and if some acceleration peak values can be expected to be low, which may occur with varying self-selected running velocities (Aubol and Milner, 2020). Following the method, the resultant acceleration (modulus) was first calculated as the vector sum of the measured three axes, Equation (5).

$$a_{res}\left[\frac{m}{s^2}\right] = \sqrt{a_x^2 + a_y^2 + a_z^2} \tag{5}$$

Where a_{res} is the resultant acceleration in m/s^2 and a_x , a_y and a_z are the acceleration vectors for each of the IMU coordinate system axes.

Then the local minimum occurring within 75 ms prior to a maximal peak in the resultant acceleration, was saved, caused by rapid increases in the signal during the beginning of the stand phase. This local minimum was considered the time point of the initial foot contact (Aubol and Milner, 2020; Garcia et al., 2021). Following the same method as for the Dorsum IMU, the signal of the Tibia IMU was cut into 50 gait cycles from the middle part of the signal. Lastly, the peak resultant acceleration was identified within the first 40 % of each of the 50 gait cycles and the mean peak resultant acceleration was calculated (Garcia et al., 2021). This resulted in one mean peak TA-R value for each shoe condition per participant.

2.4.3. VAS-questionnaire

One 100 mm VAS-questionnaire was answered for each of the two shoe conditions. The VAS-score for each comfort item (see 2.3 Study design) was measured as the length in mm from the left word anchor 'not comfortable at all' (0 mm) to the marked assessment on the scale towards the right word anchor 'most comfortable imaginable' (100 mm). This resulted in one VAS-score for each of the five comfort items per shoe condition for each participant.

2.5. Data analysis

All statistical analysis was carried out in SPSS Statistics (v. 28.0, 64-bit edition, IBM corp., United States). Paired-samples t-tests were performed to investigate if the shoe conditions had an effect on the VAS comfort scores and the running kinematic variables of the Tibia and Dorsum IMU (Shoe 1 versus Shoe 2). Further, Paired-samples t-tests were performed to investigate if any carry-over effect was present between the first and second running procedure (running the first and second time, respectively; Test 1 versus Test 2). If the assumptions, of either no significant outliers or normal distribution of the mean differences, were violated a non-parametric Wilcoxon signed-rank test was performed on those variables instead. Furthermore, a Bonferroni correction was used to compensate for multiple comparisons resulting in adjusted α -levels for the comfort variables of $\alpha = .01$ (5 variables) and for the RV variables of $\alpha = .008$ (6 variables). Tendencies were reported for p-values < 0.1.

3. Results

The following section is divided into; Shoe 1 versus Shoe 2 and Test 1 versus Test 2, comparing the five comfort items, six RV variables and peak TA-R. Figure 2 shows which statistical test was performed on each variable as mean \pm SD (indicated as bar plots) is shown for variables included in the parametric analysis (Paired-samples t-test) and median and [25-75] interquartile range (IQR) (indicated as boxplots) is shown for variables included in the non-parametric analysis (Wilcoxon signed-rank test).

3.1. Shoe 1 versus Shoe 2

For the comfort items no significant differences were seen between the variable population means and medians of Shoe 1 and Shoe 2, see Table 2. A general trend towards higher mean VAS-scores was seen for Shoe 1 compared to Shoe 2 for four comfort items, see Figure 2. The biggest trend was seen for 'Overall comfort' (Shoe 1: 75.2 ± 17.1 mm, Shoe 2: 60.0 ± 24.0 mm). This resulted in a tendency of difference (p < 0.1) in 'Overall comfort', see Table 2, with a medium effect size between Shoe 1 and Shoe 2, as this result would have been a significant difference if the Bonferroni correction was not included. Further, a visually high variance of the individual data points is apparent for all five comfort items with a trend towards higher SD for Shoe 2 compared to Shoe 1. No systematic trends were seen between data points for males and females except a visual trend of higher 'Shoe stability' scores for males compared to females for Shoe 1. Lastly, 12 out of 15 participants rated Shoe 1 higher than Shoe 2 in terms of 'Overall comfort' and 11 out of 15 participants expressed Shoe 1 as

the preferred and more comfortable shoe compared to Shoe 2 during verbal feedback after the completion of the test.

The results of the RV variables of the Dorsum IMU showed no significant differences between the variable population means and medians of Shoe 1 and Shoe 2, see Table 2. The highest RV was seen for the angular velocity around the x-axis (Gyro x) and the acceleration in the z-axis (Acc z) of up to a mean of 46.6 % and 37.8 % respectively followed by the angular velocity around the y-axis (Gyro y) of up to a median of 21.1 %. The remaining RV variables were all below a RV of 16 %. No systematic trends were seen between data points for males and females.

Peak TA-R showed no significant differences between Shoe 1 and Shoe 2 and no difference was seen in run duration. Lastly, all variables included in the Paired-samples t-test, except 'Overall comfort', had a low effect size (Cohen's d < 0.5).

3.2. Test 1 versus Test 2

For the comfort items no significant differences were seen between the variable population means and medians of Test 1 and Test 2, see Table 2. A general trend towards higher mean VAS-scores was seen for Test 2 compared to Test 1 for four comfort items, see Figure 2. The biggest trend was seen for 'Heel cushioning' (Test 1: 55.5 mm, Test 2: 70.7 mm). This resulted in a tendency of difference (p < 0.1) in 'Heel cushioning', see Table 2, with a small effect size between Test 1 and Test 2. No systematic trends were seen between data points for males and females except a visual trend of higher 'Shoe stability' scores for males compared to females for Test 2.

The results of the RV variables of the Dorsum IMU showed no significant differences between the mean or median of the six RV variables between Test 1 and Test 2. Further, no systematic trends between data points for males and females were seen.

Peak TA-R showed no significant differences between Test 1 and Test 2 and no difference was seen in run duration. Lastly, all variables included in the Paired-samples t-test, had a low effect size (Cohen's d < 0.5).



Figure 2: Mean values \pm standard deviation (SD) (indicated by bar plots and error bars) are shown for variables included in the parametric statistical tests while median values and [25-75] interquartile range (IQR) (indicated by boxplots) are shown for variables included in the non-parametric tests, for comparison between both Shoe 1 and Shoe 2 and Test 1 and Test 2. Further, individual data points are shown for each of the five comfort items, six relative variability (RV) variables of the Dorsum IMU and peak resultant tibial acceleration (Peak TA-R) of the Tibia IMU. The individual data points are further divided into 'Males' and 'Females' indicated by blue and red dots respectively. The x-axis shows the different variables while the y-axis shows the measured units; VAS-score, RV and acceleration for the comfort items, Dorsum IMU and Tibia IMU respectively. See Table 2 for statistical results.

Table 2: The results of the statistical analysis for Shoe 1 versus Shoe 2 and Test 1 versus Test 2 are shown for both the performed Paired-samples t-tests and Wilcoxon signed-rank tests. For the paired samples t-tests, the mean differences and its standard deviations (SD) are shown for each variable followed by the p-values and the computed Cohen's d. For the Wilcoxon signed-rank tests, the median differences and the computed p-values are shown. The variables consist of the five comfort items, six relative variability (RV) variables of the Dorsum IMU and peak resultant tibial acceleration (Peak TA-R) of the Tibia IMU. See Figure 2 for getting the mean \pm SD and median values and [25-75] interquartile range (IQR) values.

Statistical analysis: Shoe 1 vs Shoe 2							
Paired-samples t-test			Wilxo	ocon signed-rank tes	t		
Variable	Mean difference	SD	p-value	Cohen's d	Variable	Median difference	p-value
Overall comfort	15.2	27.0	0.047	0.563	Shoe stability	7.0	0.221
Heel cushioning	10.2	34.7	0.274	0.294	RV ax	-0.3	0.910
Forefoot cushioning	9.0	32.9	0.307	0.274	RV gy	1.2	0.173
Forefoot flexibility	5.7	20.6	0.305	0.275	RV gz	0.1	0.776
RV ay	0.1	1.3	0.744	0.086			
RV az	2.4	15.5	0.553	0.157			
RV gx	-2.0	15.3	0.617	-0.132			
Peak TA-R	6.5	18.5	0.197	0.349			
Run duration	-0.3	2.3	0.666	-0.114			

Statistical analysis: Test 1 vs Test 2

	Paired-samples	s t-test			Wilxoco	on signed-rank tes	t
Variable	Mean difference	SD	p-value	Cohen's d	Variable	Median difference	p-value
Overall comfort	-2.7	31.1	0.745	-0.086	Shoe stability	3.0	0.977
Heel cushioning	-15.1	32.7	0.095	-0.462	RV gx	-0.7	0.955
Forefoot cushioning	-6.1	33.6	0.495	-0.181	RV gy	-1.0	0.609
Forefoot flexibility	-6.7	20.3	0.219	-0.332	RV gz	-0.3	0.191
RV ax	-0.2	4.3	0.866	-0.044			
RV ay	-0.5	1.2	0.137	-0.407			
RV az	4.1	15.1	0.314	0.270			
Peak TA-R	-0.1	19.6	0.989	-0.004			
Run duration	0.5	2.3	0.384	0.232			

4. Discussion

This balanced randomized cross-over study investigated the effect on perceived footwear comfort, running kinematic variability and peak TA-R using two different recreational road running shoes with similar characteristics. It was hypothesized that higher running kinematic RV and lower peak TA-R would be seen for the shoe perceived the most comfortable. Contrary to the hypothesis there were no statistical differences in any of the five measured comfort items nor in running kinematics between the two shoe conditions.

4.1. Shoe conditions

Two shoe conditions were tested in this study (Shoe 1: Nike Pegasus 38, Shoe 2: Icebug Aura RB9X). Both conditions lie within the category of traditional running shoes intended for recreational road running, of the three described categories of running shoes; minimal, traditional and maximal (see section 1.). Both shoes had a thick midsole for cushioning purposes though not as extensive as the Hoka OneOne (HOKA, California, USA) often considered a maximal cushioned shoe (Sinclair et al., 2021). Further, both shoes had rockersole shapes and did not differ much in terms of weight, midsole hardness, heel-toe drop and inside dimensions, except Shoe 2 in size 43 with slightly higher midsole hardness and weight (see section 2.2). Therefore, the two shoe conditions investigated in this study were considered similar in comparison. Thus, the choice of two similar running shoes enabled to benchmark the one shoe to the other in conditions where the subjects' perception and the running kinematics may not be too different.

4.2. Shoe 1 versus Shoe 2

As reported in Table 2, there were no differences in comfort. This finding is contradictory to a range of earlier studies assessing perceived comfort for different footwear conditions in effort to investigate potential relations between footwear comfort and biomechanical measures (Dinato et al., 2015; Wegener et al., 2008; Horvais et al., 2019; Yang et al., 2019). Further, previous studies have found a relationship between comfort and biomechanical or physiological parameters (plantar and dorsal peak pressure and maximum force, oxygen consumption, kinematic RV, peak anterior-posterior plantar center of pressure velocity), while not reporting any significant comfort differences between footwear conditions, by comparing the least and most comfortable conditions (Jordan et al., 1997; Luo et al., 2009; Lam et al., 2018a; Meyer et al., 2018). In general, the majority of these studies have compared vastly different footwear conditions with varying cushioning technologies, weight and purpose, e.g., from minimalist to stability shoes. This is in contrast with the shoe conditions of this study as these were fairly similar (See section 4.1). Comfort is considered complex to assess as it is related to multiple factors such as well-being and relaxation but also aesthetic design (De Looze et al., 2003; Helander and Zhang, 1997) and physiological and anatomical differences making comfort difficult to quantify (Menz and Bonanno, 2021). In this study, comfort was assessed using a 100 mm VAS-questionnaire with four categories in line with Bishop et al. (2020). As shown on Figure 2, the mean and median comfort scores were all >50 mm. The same trend of high comfort scores has been shown by previous studies (Dinato et al., 2015; Wegener et al., 2008; Yang et al., 2019) which might indicate that participants in general perceive modern running footwear highly comfortable. This could be explained by the extensive attention and development of running footwear the past 50 years (see section 1.). This might also contribute to the fact that no comfort differences were found.

The lack of differences in comfort is most likely explained by the similar characteristics of the two tested shoes despite the observed tendency towards a higher overall comfort for Shoe 1 compared with Shoe 2. A higher comfort for Shoe 1 was supported by verbal feedback where 11 out of 15 participants expressed Shoe 1 as more comfortable than Shoe 2, when questioned after completion of the experimental protocol. This makes ground for speculation as to if the participants needed further familiarization to reliably assess the shoes using the VAS-questionnaire. Furthermore, the participants of this study assessed each shoe condition once based on a 400 m run after one familiarization with the VAS-questionnaire on their own shoes. Here, Meyer et al. (2018) averaged two comfort assessments from two different test days, each based on 200 m of running per condition, which might be a potential approach for increasing reliability.

The results based on IMU showed no differences in RV or peak TA-R between the two shoes contrary to Meyer et al. (2018). The authors have found differences in kinematic RV between the most and least comfortable shoe for each participant with higher RV for the most comfortable shoe. Kinematic variability allows exploration of biological signals, as one can investigate the deviations from an average step in e.g., a repetitive acceleration signal during running gait cycles (Svenningsen et al., 2020). In addition, it can be normalized which is important as it enables to compare participants with different running speed and IMU output (Meyer et al., 2018; Svenningsen et al., 2020). Meyer et al. (2018) have shown the largest kinematic differences in approximated kinematic variability for rotation in the transverse plane (external/internal ankle rotation) and acceleration in the frontal plane (medio-lateral translation) of the shoe in the late swing phase (75-95 % of gait cycle). When comparing the six RV variables, the results of this study resemble the patterns seen in Meyer et al. (2018). As shown in Figure 2, the largest mean and median RV values were seen for acceleration in z-axis and angular velocity around the x and y-axis (from 21.1 % to 46.6 %) and in the previous research the same corresponding variables was shown to be highest (\approx 35% to 50%). With respect to peak TA-R, studies have both shown lower (Sinclair et al., 2017) and higher (Lam et al., 2018b) peak tibial accelerations for more cushioned shoes. Oppositely, no differences have also been seen between shoe conditions featuring different midsole thickness and hardness levels (Hardin and Hamill, 2002; Chambon et al., 2014). These findings reflect scattered conclusions of the effect of different midsole types though it has been suggested that softer and thicker midsoles can provide better shock absorption (Sun et al., 2020). Comparing the peak TA-R, the mean result of this study lies between 93.2 m/s^2 and 99.7 m/s^2 which complies with the above-mentioned previous research which have shown peak tibial accelerations during running between 62.5 m/s^2 and 146.3 m/s^2 (originally reported in g).

As mentioned above, the similar characteristics of Shoe 1 and Shoe 2, as well as the lack of difference in comfort most likely explain why no differences were seen in RV and peak TA-R. Meyer et al. (2018) have proposed that their findings might help simplify footwear recommendations as the effect of comfort on RV, found by the authors, provides a potential objective measure for comfort. The current results of this study did not confirm this proposal. Therefore it can be discussed whether RV of running kinematics as an objective measure is relevant for personalized footwear recommendations in a shoe store or as a surrogate measure for perceived comfort as of yet. In addition, Meyer et al. (2018) have analyzed the RV based on the comparison of the least and most comfortable rated shoe for each participant, from a selection of 5 running shoes. Such an approach does not enable to distinguish one shoe model from the other.

In conclusion, the two tested shoes were characterized as equally comfortable and with similar RV and peak TA-R.

4.3. Methodological considerations

The running distance of 400 m for each shoe condition was chosen based on a range of studies including running distances of 160 m - 450 m (Lam et al., 2018a; Meyer et al., 2018; Mündermann et al., 2002). Though, it can be discussed whether the comfort assessed exclusively is a matter of initial comfort as complete familiarization of a new shoe condition might not have been present within 400 m of running. In addition, the two consecutive 400 m running procedures did not show any sign of a carry-over effect as no differences were

seen between Test 1 and Test 2 (see Table 2). This also indicates that the results were not influenced by the running procedures being executed during the same session. However, a tendency towards higher 'Heel cushioning' for Test 2 was reported.

This study supported assessment of comfort in the natural environment the shoe conditions were intended for as the experimental protocol was performed in-field on asphalt. Combined with a self-selected running velocity for each participant it ensured natural running patterns and rhythmicity during IMU data collection (Svenningsen et al., 2020). Being an in-field study, the reproducibility is affected by environmental parameters such as weather conditions and temperature (Svenningsen et al., 2020). However, this most likely did not affect the results as within-subject comparisons were performed in this study design and no carry-over effect was seen. Furthermore, IMUs were used to measure running kinematics as these enable in-field data collections over long distances (Svenningsen et al., 2020; Aristizábal Pla et al., 2021). With respect to measuring peak tibial accelerations IMUs are less accurate due to their higher mass (Shimmer3 IMU, Weight: 23.6 g) compared with a standalone tri-axial accelerometer (Sheerin et al., 2019). Again, this most likely did not affect the results as within-subject comparisons were performed. Additionally, saturation of the acceleration signal was seen for the Dorsum IMU for 5 out of 15 participants for Shoe 1 and for 7 out of 15 participants for Shoe 2. Further, saturation of the tibial IMU was seen for one participant for both Shoe 1 and Shoe 2. This might be explained by different running styles as running speed have shown to increase measured accelerations (Lam et al., 2018b). The y-axis was the most affected axis for both IMUs. Therefore, the results of the peak TA-R and RV were affected by the saturation. Of note, saturation occurred for both shoes suggesting that the acceleration data were mostly equally affected. However, no saturation was seen during pilot tests and accelerometers configured to a measurement range of $\pm~16~{\rm g}$ have been used in previous research (Lucas-Cuevas et al., 2017). This calls for use of IMUs with configurations higher than \pm 16 g for future studies. In addition, the RV results were most likely not affected by the fixation of the Dorsum IMU on two different running shoes for each participant. This, because the data was rotated to align with the foot coordinate system based on its measured tilt during the calibration procedure (see section 2.3).

As Shoe 2 was only provided in two prototype sizes, EU 43 for males and EU 39 for females, it can be discussed if the shoe fit was optimal for every participant and therefore if shoe fit affected the comfort assessments as well-fitted shoes are perceived more comfortable (Menz and Bonanno, 2021). Furthermore, I was not allowed by Icebug AB to blind Shoe 2 with tape, spray paint etc. as these were intended for further testing, why it is uncertain if the blindfold and instruction of not looking at the shoes was sufficient to reduce the risk of bias of comfort assessments (Matthias et al., 2021).

Lastly, a small effect size (Cohen's d < 0.5) was seen for all comfort and kinematic variables except 'Overall comfort' for Shoe 1 versus Shoe 2 (Cohen's d > 0.5). This indicates that there was a risk of not detecting a significant result even if there was a practical significance. The small effect sizes might have been caused by the sample size of 15 participants as small sample sizes negatively affects the effect size. The minimum sample size for future similar studies should consist of at least 23 participants to ensure a large effect size, based on a power analysis using the 'Overall comfort' mean difference and SD for Shoe 1 versus

Shoe 2.

Further studies are needed to investigate if differences between similar footwear conditions can be seen in comfort assessments and IMU running kinematics to gain a better understanding of the use of such methods on e.g., benchmarking of running shoes.

5. Conclusion

Two recreational road running shoes were benchmarked in-field with respect to perceived comfort and IMU running kinematics. In conclusion, no differences were found between the two shoes in terms of comfort, RV and peak TA-R during a 400 m run on asphalt paved road. Sensor miniaturization makes it possible to instrument running shoes with embedded systems that could include IMU and communication protocol similar to the Altra Torin IQ running shoe (Altra Running, Utah, USA). Such an approach can be valuable for both runners and running shoe companies to investigate worn-out footwear, biomechanical characteristics of runners as well as comfort and incidence rate of running injuries on a large scale in a prospective study design.

6. Data availability

The data and custom MATLAB script supporting the findings of this study are available on request at: https://doi.org/10.5281/zenodo.6600697

7. Disclosure

The author conducted the study during an internship at Icebug AB. Icebug AB commented on the experimental protocol but had no influence on the results and their interpretation.

8. Acknowledgements

The author would like to thank Icebug AB for a good and productive collaboration and for providing test shoes and accessories. Furthermore, the author would like to thank Aktivitus AB (Göteborg, Sweden) for providing a laboratory setting for conducting the experimental protocol.

References

- Aristizábal Pla, G., Hollville, E., Schütte, K., and Vanwanseele, B. (2021). The Use of a Single Trunk-Mounted Accelerometer to Detect Changes in Center of Mass Motion Linked to Lower-Leg Overuse Injuries: A Prospective Study. Sensors, 21(21):7385.
- Aubol, K. G. and Milner, C. E. (2020). Foot contact identification using a single triaxial accelerometer during running. *Journal of Biomechanics*, 105:109768.
- Bishop, C., Buckley, J., Esterman, A., and Arnold, J. (2020). The running shoe comfort assessment tool (RUN-CAT): Development and evaluation of a new multi-item assessment tool for evaluating the comfort of running footwear. *Journal of Sports Sciences*, 38(18):2100–2107.

- Chambon, N., Delattre, N., Guéguen, N., Berton, E., and Rao, G. (2014). Is midsole thickness a key parameter for the running pattern? *Gait & Posture*, 40(1):58–63.
- De Looze, M. P., Kuijt-Evers, L. F. M., and Van Dieën, J. (2003). Sitting comfort and discomfort and the relationships with objective measures. *Ergonomics*, 46(10):985–997.
- Dinato, R., Ribeiro, A., Butugan, M., Pereira, I., Onodera, A., and Sacco, I. (2015). Biomechanical variables and perception of comfort in running shoes with different cushioning technologies. *Journal of Science* and Medicine in Sport, 18(1):93–97.
- Garcia, M. C., Gust, G., and Bazett-Jones, D. M. (2021). Tibial acceleration and shock attenuation while running over different surfaces in a trail environment. *Journal of Science and Medicine in Sport*, 24(11):1161–1165.
- Hardin, E. C. and Hamill, J. (2002). The Influence of Midsole Cushioning on Mechanical and Hematological Responses during a Prolonged Downhill Run. Research Quarterly for Exercise and Sport, 73(2):125–133.
- Helander, M. G. and Zhang, L. (1997). Field studies of comfort and discomfort in sitting. Ergonomics, 40(9):895–915.
- Hennig, E. (2011). Eighteen years of running shoe testing in Germany a series of biomechanical studies. Footwear Science, 3(2):71–81.
- Horvais, N., Samozino, P., Chiementin, X., Morin, J.-B., and Giandolini, M. (2019). Cushioning perception is associated with both tibia acceleration peak and vibration magnitude in heel-toe running. *Footwear Science*, 11(1):35–44.
- Huang, M., Mo, S., Pak-Kwan Chan, P., Chan, Z. Y., Zhang-Lea, J. H., and Cheung, R. T. (2022). The influence of running shoes on familiarization time for treadmill running biomechanics evaluation. *Sports Biomechanics*, pages 1–14.
- Jordan, C., Payton, C., and Bartlett, R. (1997). Perceived comfort and pressure distribution in casual footwear. *Clinical Biomechanics*, 12(3):S5.
- Koplan, J. P., Rothenberg, R. B., and Jones, E. L. (1995). The natural history of exercise: a 10-yr follow-up of a cohort of runners. *Medicine and Science in Sports and Exercise*, 27(8):1180–1184.
- Lam, C. K.-Y., Mohr, M., Nigg, S., and Nigg, B. (2018a). Definition and quantification of 'ride' during running. *Footwear Science*, 10(2):77–82.
- Lam, W.-K., Liebenberg, J., Woo, J., Park, S.-K., Yoon, S.-H., Cheung, R. T.-H., and Ryu, J. (2018b). Do running speed and shoe cushioning influence impact loading and tibial shock in basketball players? *PeerJ*, 6:e4753.
- Lucas-Cuevas, A. G., Encarnación-Martínez, A., Camacho-García, A., Llana-Belloch, S., and Pérez-Soriano, P. (2017). The location of the tibial accelerometer does influence impact acceleration parameters during running. *Journal of Sports Sciences*, 35(17):1734–1738.
- Luo, G., Stergiou, P., Worobets, J., Nigg, B., and Stefanyshyn, D. (2009). Improved footwear comfort reduces oxygen consumption during running. *Footwear Science*, 1(1):25–29.
- Marchena-Rodriguez, A., Ortega-Avila, A. B., Cervera-Garvi, P., Cabello-Manrique, D., and Gijon-Nogueron, G. (2020). Review of Terms and Definitions Used in Descriptions of Running Shoes. International Journal of Environmental Research and Public Health, 17(10):3562.
- Martínez-Martínez, J. M., Martín-Guerrero, J. D., Soria-Olivas, E., Bernabeu, J. A., Escandell-Montero, P., Stark, R. H., Serrano-López, A. J., and Montiel, E. (2017). Use of SOMs for footwear comfort evaluation. *Neural Computing and Applications*, 28(7):1763–1773.
- Matthias, E. C., Banwell, H. A., and Arnold, J. B. (2021). Methods for assessing footwear comfort: a systematic review. *Footwear Science*, 13(3):255–274.
- Menz, H. B. and Bonanno, D. R. (2021). Footwear comfort: a systematic search and narrative synthesis of the literature. *Journal of Foot and Ankle Research*, 14(1):63.
- Meyer, C., Mohr, M., Falbriard, M., Nigg, S., and Nigg, B. (2018). Influence of footwear comfort on the variability of running kinematics[†]. *Footwear Science*, 10(1):29–38.
- Mohr, M., von Tscharner, V., Nigg, S., and Nigg, B. M. (2021). Systematic reduction of leg muscle activity throughout a standard assessment of running footwear. *Journal of Sport and Health Science*, page S209525462100003X.

- Mündermann, A., Nigg, B., Stefanyshyn, D., and Humble, R. (2002). Development of a reliable method to assess footwear comfort during running. *Gait and Posture*, 16(1):38–45.
- Mündermann, A., Stefanyshyn, D. J., and Nigg, B. M. (2001). Relationship between footwear comfort of shoe inserts and anthropometric and sensory factors:. *Medicine and Science in Sports and Exercise*, 33(11):1939–1945.
- Sheerin, K. R., Reid, D., and Besier, T. F. (2019). The measurement of tibial acceleration in runners—A review of the factors that can affect tibial acceleration during running and evidence-based guidelines for its use. *Gait & Posture*, 67:12–24.
- Sinclair, J., Brooks, D., Taylor, P. J., and Liles, N. B. (2021). Effects of running in minimal, maximal and traditional running shoes: a musculoskeletal simulation exploration using statistical parametric mapping and Bayesian analyses. *Footwear Science*, 13(2):143–156.
- Sinclair, J., Rooney, E., Naemi, R., Atkins, S., and Chockalingam, N. (2017). EFFECTS OF FOOTWEAR VARIATIONS ON THREE-DIMENSIONAL KINEMATICS AND TIBIAL ACCELERATIONS OF SPE-CIFIC MOVEMENTS IN AMERICAN FOOTBALL. Journal of Mechanics in Medicine and Biology, 17(02):1750026.
- Sun, X., Wing-Kai, L., Zhang, X., Wang, J., and Fu, W. (2020). Systematic Review of the Role of Footwear Constructions in Running Biomechanics: Implications for Running-Related Injury and Performance. *OJournal of Sports Science and Medicine*, 19(*OJournal of Sports Science and Medicine*):20–37.
- Svenningsen, F. P., Pavailler, S., Giandolini, M., Horvais, N., and Madeleine, P. (2020). A narrative review of potential measures of dynamic stability to be used during outdoor locomotion on different surfaces. *Sports Biomechanics*, 19(1):120–140.
- Taunton, J. E. (2002). A retrospective case-control analysis of 2002 running injuries. British Journal of Sports Medicine, 36(2):95–101.
- Van Alsenoy, K., van der Linden, M. L., Girard, O., and Santos, D. (2021). Increased footwear comfort is associated with improved running economy – a systematic review and meta-analysis. *European Journal* of Sport Science, pages 1–13.
- van Gent, R. N., Siem, D., van Middelkoop, M., van Os, A. G., Bierma-Zeinstra, S. M. A., Koes, B. W., and Taunton, J. E. (2007). Incidence and determinants of lower extremity running injuries in long distance runners: a systematic review * COMMENTARY. British Journal of Sports Medicine, 41(8):469–480.
- Wegener, C., Burns, J., Penkala, S., and Sc, G. D. E. S. (2008). Effect of Neutral-Cushioned Running Shoes on Plantar Pressure Loading and Comfort in Athletes with Cavus Feet: A Crossover Randomized Controlled Trial. *The American Journal of Sports Medicine*, 36(11):2139–2146.
- Wu, G. and Cavanagh, P. R. (1995). ISB recommendations for standardization in the reporting of kinematic data. Journal of Biomechanics, 28(10):1257–1261.
- Yang, F., Liu, Z.-L., Ma, R.-Y., and Lam, W.-K. (2019). Using plantar pressure to predict insole comfort in running. *Footwear Science*, 11(sup1):S15–S17.

Worksheet

1.0 Early Problem Analysis (25.02.2022)	2
 2.0 Theory of Inertial measurement units 2.1 Accelerometers 2.2 Gyroscopes 2.3 Analog to digital conversion 	5 5 6
3.0 Method	8
3.1 Quantification of subjective comfort	8
3.1.1 Measurement tools for assessing perceived footwear comfort	8
3.1.2 Reliability and validity of comfort assessments	9
3.1.3 Initial or prolonged comfort	10
3.2 Biomechanical and physiological assessments and their relation to perceived footwear comfort	11
3.2.1 Kinetic recordings	11
3.2.2 Kinematic recordings	12
3.2.3 Physiological recordings	13
3.2.4 Choice of objective measurement method	14
3.3 Pilot test	15
3.4 Experimental Method	16
3.4.1 Participants	16
3.4.2 Shoe conditions	16
3.4.3 Study design	18
3.4.4 Measurement equipment	21
3.4.5 Data processing and analysis	22
3.4.5.1 Dorsum IMU - Relative variability	22
3.4.5.2 Tibia IMU - peak resultant tibial acceleration	26
3.4.5.3 VAS questionnaire	28
3.4.5.4 Statistical analysis	29
4.0 Results	32
4.1 Shoe 1 versus shoe 2	32
4.2 Test 1 versus Test 2	36
5.0 References	39

1.0 Early Problem Analysis (25.02.2022)

The following problem analysis reflects the process of this study as it shows the early focus of this master's thesis and was the current version as per 25.02.2022. In the bottom of this section is a brief description of the current purpose of this study.

Since the running boom started during the 1970's in the USA there has been an interest in running injuries and their prevention since many runners experienced overuse injuries (Hennig, 2011; Novacheck, 1998). Koplan et al. (1995) showed that 53% of 521 runners experienced at least one injury from 1981 to 1990 while Nigg et al. (2015) has found that relative running injury frequencies vary between 15% to 85% based on 17 studies. This might be explained by the use of different definitions of running injuries from requiring medical attention to symptoms of discomfort or pain (Nigg et al., 2015). Since the 1970's recreational running has evolved into becoming the most common type of physical activity, but overuse injuries have shown to be the main reason for practitioners to quit running accounting for 31% of reported reasoning (Hennig, 2011; Koplan et al., 1995). Regarding running injuries knee problems have shown to account for 42% including patellofemoral pain syndrome (PFPS) being the most frequent among 2002 runners. Other common injuries were; iliotibial band friction syndrome (ITBFS), meniscal injuries, tibial stress syndrome and plantar fasciitis. (Taunton, 2002)

Extrinsic and intrinsic factors influencing running injuries have been assessed through a wide range of studies (Nigg et al., 2015). Extrinsic factors such as prior injuries, weekly running distance, low aerobic fitness and training environment are associated with overuse musculoskeletal injuries (Molloy, 2016; Nigg et al., 2015). Intrinsic factors such as pronation of the foot and high impact forces during the heel-toe stance phase were believed to increase risk of running injuries during early research on running shoes. These factors were already believed to be the main factors increasing risk of injury before biomechanical analysis, of running and injuries related to it, was available and therefore it had a great impact on the future research of the same topic (Hennig, 2011; Nigg et al., 2015; Novacheck, 1998). There is clinical support to the statement that excessive pronation of the foot can lead to painful conditions remedied by shoes or insoles designed to reduce pronation though only little evidence supports such an abnormal biomechanics. For the impact forces it was believed that the impact of heel strike was reason for running injuries though higher forces over a longer period is seen during the last 25% of the stance phase. (Novacheck, 1998) In addition, a review by Nigg et al. (2015) argues that no evidence supports that either vertical impact peaks, vertical loading rates or foot pronation contribute to increased risk of running injuries. Novacheck (1998) emphasizes however that attenuation of the shock experienced during ground contact is not unimportant. Conversely according to a review by Hennig (2011) it appears that prospective studies have shown that preventing excessive rearfoot motion (reducing overpronation) and increasing shock absorption in shoes lowers the risk of overuse injuries.

With foot pronation and impact forces during stance phase being the most commonly studied variables related to running injuries (Nigg et al., 2015), rearfoot motion and shock absorption have become the primary objectives for biomechanical evaluation of running shoes (Hennig, 2011). Therefore the running shoe industry uses arguments like good shock attenuation and

overpronation prevention to sell their shoes, which has led to the development of a series of different technologies in order to try to achieve this. This includes a higher medial midsole than lateral at the heel or using a more stiff midsole material at the medial side compared to the lateral side of the heel in order to try to reduce overpronation. (Hennig, 2011) Shoe construction differences are seen in midsole hardness and thickness, heel-toe drop and different designs of crash pads in order to try to achieve good shock attenuation. Further running shoe technologies and designs vary within; shoe lacing, insole design, heel cup, shoe upper, bending stiffness and minimalist vs. conventional design. (Nigg et al., 2015; Sun et al., 2020) Of these different approaches to sell footwear and possibly reduce risk of injury, shock absorption and overpronation reduction has as aforementioned been suggested to reduce risk of injury (Hennig, 2011). Shock absorption technologies such as a softer or thicker midsole have proven to reduce loading rates and impact forces (Sun et al., 2020). Contrary to the more discussed and uncertain effect of the shock attenuation softer shoe insoles in military boots have shown to reduce injuries (Nigg et al., 2015).

The above mentioned biomechanics and shoe technologies and their direct effect on the frequency of acquired running injuries seem to be conflicting between studies or with little or no evidence as to if one specific factor can reduce the risk of injuries. Though other factors have shown to be important in relation to injuries. "The preferred movement path" is introduced by Nigg et al. (2015), and states that the actual path of the movement does not change much between different conditions of barefoot vs shod running or varying shoe or insole conditions. The changes between the conditions were primarily found in the range of movement but still following the same path of movement. This is explained by the activation of the skeleton muscles ensuring that the skeleton stays within the correct path, though the amplitude of the path may vary. Therefore if the shoe does not support one's preferred path it will demand more muscle activity to keep the skeleton in the correct path. (Nigg et al., 2015) Further (Mündermann et al., 2001) tested the relationship between shoe inserts and footwear comfort, dividing 206 soldiers into an insert group (n=103) and a control group (n=103), where the insert group chose one of six varying insoles based on which scored highest of their individual perceived comfort. The study showed that the insert group wearing their individually most comfortable insole had 53% less lower-extremity injuries compared to the control group. In relation to the above mentioned Nigg et al. (2015) proposes that comfort is an overall important factor for reducing risk of injuries of the lower extremities during movement and introduces the "comfort filter" as a new paradigm. Meaning when a user is choosing a running shoe they will automatically avoid the uncomfortable shoes and thereby also avoid the potentially harmful effect of the shoe. (Nigg et al., 2015)

A systematic review by Van Alsenoy et al. (2021) also concludes that a small ($-2.06 mL \cdot kg^{-1} \cdot min^{-1}$) though statistically lower oxygen consumption is seen (P=0.01) when wearing most comfortable footwear versus least comfortable during submaximal speed resulting in a slightly better running economy. Furthermore, comfortable footwear might promote being more physically active (Shields et al., 2017) and from the commercial perspective comfort is of high priority as a user will only choose a shoe if it feels comfortable (Hennig, 2011; C. K.-Y. Lam et al., 2018; Nigg et al., 2015). Lastly comfort is very subjective and different individuals will choose different shoe constructions to be able to feel comfortable which is assumed to be influenced by factors like skeletal alignment, foot sensitivity and foot shape but also the individual's perception and aesthetic preferences given that comfort is a psychological and neurophysiological construct (Menz & Bonanno, 2021; Miller et al., 2000; Mündermann et al., 2001; Nigg et al., 2015). Therefore comfort can be difficult to quantify (Nigg et al., 2015) and a review by Menz & Bonanno (2021) reports that no studies of a total of 101 manuscripts provides a specific definition of the term; comfort. Hennig (2011) also emphasizes that test subjects are not very well capable of differentiating between specific important and different shoe construction features. More specifically if a subject likes a shoe it affects all the shoe properties positively and vice versa though subjective evaluation is still important to get an understanding of overall liking. For that reason both mechanical, objective and subjective measures are deemed important for evaluation of shoes. (Hennig, 2011)

For the above cited studies, perceived comfort has been measured in order to assess whether or not a user likes the running feeling of a shoe. Though comfort is associated with feelings of well-being, relaxation and even to the aesthetics of a design. On the other hand discomfort is mainly associated with pain, soreness, numbness and tiredness, which might be more relevant if the differences between the subjective feeling of a range of shoes is to say something about the actual physical feeling of running without being affected by other factors such as aesthetics. (De Looze et al., 2003; Helander & Zhang, 1997)

With a wish to develop a more sustainable recreational running shoe focusing on the discomfort aspect and to address the broadest spectrum of running shoe users it is important to get knowledge about the impact of the running shoe on the body. Such knowledge should be acquired through testing that resembles the requirements and purpose of the shoe given that comfort assessment is influenced by how and where a shoe is used (Agresta et al., 2021; Menz & Bonanno, 2021; Miller et al., 2000). Further to address the broadest spectrum of runners it would be advantageous to use rearfoot strikers as they represent 80% of distance runners (Novacheck, 1998). In order to test the prototype shoe the discussed approaches to assess running shoes on the body both objectively and subjectively can give valuable insight into potential weak spots, areas of improvements and to shed light upon differences between a running shoe prototype in development and already commercial recreational running shoes.

Based on the aforementioned, the purpose of this study is (a) to evaluate the discomfort of a sustainability focused recreational road running shoe prototype of the brand Icebug AB compared to other recreational running shoes by other commercial brands for heel-toe runners by getting insights from plantar pressure distribution, accelerometer kinematics and VAS/NRS during sub-maximal running (b) to suggest improvement areas to the Icebug prototype to both reduce perceived discomfort and most likely lower the risk of running injuries.

Approach after early problem analysis

On the basis of the above early problem analysis it was agreed during a supervisor meeting on 02/03/2022 with both the academic and company supervisor to focus the purpose of the study on; benchmarking two recreational road running shoes with respect to comfort and running kinematics using a visual analogue scale (VAS) - questionnaire and inertial measurement units (IMU), respectively. It was hypothesized that higher running kinematic relative variability (RV) and lower peak resultant tibial acceleration (Peak TA-R) would be seen for the shoe perceived the most comfortable.

2.0 Theory of Inertial measurement units

An inertial measurement unit (IMU) is an electronic device using the combination of accelerometers, gyroscopes and often magnetometers to measure linear acceleration, angular velocity and the geomagnetic field respectively (Arun Faisal et al., 2019). As only accelerometers and gyroscopes are used in this study, magnetometers are not described further.

2.1 Accelerometers

Accelerometers measure the linear acceleration often combined in three axes (tri-axial accelerometer) as a result of inertial movement and the gravitational pull. This acceleration can be transformed to velocity or spatial displacement for each measured time point by integration or double integration, respectively. The accelerometer can further be used to calculate inclination during static conditions, which oppositely becomes inaccurate during more dynamic conditions as it is not possible to accurately decompose the acceleration signal into gravitational and inertial components during such conditions. (Lambrecht & del-Ama, 2014).

Measuring the change in capacitance of a moving mass is a common method for measuring acceleration as the two are related, resultning in the accelerometer measuring the acceleration as mV/g (g \approx 9,82 m/s²) (Dadafshar, 2015). This method is associated with not being prone to variation or noise caused by changes in temperature and its high accuracy. One example of an accelerometer would be a single movable mass located and connected between two fixated electrodes with a mechanical spring in a closed system. A displacement of the mass, caused by acceleration, relative to the fixed electrodes will cause a change in the capacitances. This change in capacitance can be used to calculate both the size of the displacement and its direction by finding the difference between the capacitance of the two fixated electrodes (Arun Faisal et al., 2019; Dadafshar, 2015). With the known displacement, the force acting upon the mass can be derived with the known spring constant (Hooke's law):

$$F = k \cdot x$$

Where F is the force (N), k is the spring constant and x is the displacement (m). Thereafter, based on Newton's second law of motion, the linear acceleration can be calculated as both the force and mass is known: (Arun Faisal et al., 2019; Dadafshar, 2015)

$$a = F \cdot m$$

Where a is the acceleration (m/s^2) , F is the force (N) and m is the mass (kg).

2.2 Gyroscopes

Gyroscopes measure the angular velocity around an axis often combined in three axes (tri-axial gyroscopes) which allows tracking changes in orientation. Opposite to the accelerometer the gyroscope can be applied in highly dynamic conditions but is prone to drift as a result of temperature when integrating the angular velocity to angular change. The drift

over time is caused by spikes in the integrated signal resulting in the calculated tilt being continuously further off compared to the true tilt angle. (Lambrecht & del-Ama, 2014)

Similar to the accelerometer the gyroscope uses changes in capacitance to calculate the experienced angular velocity, as these are related resulting in the gyroscope measuring angular velocity as mV/°/s. (Dadafshar, 2015) This calculation is based on the Coriolis effect which states that a mass (m) moving with a velocity (v) in a frame of reference that is rotating at an angular velocity (ω) will experience a force (F_c) (Arun Faisal et al., 2019):

$$F_c = -2 \cdot m(\omega \cdot v)$$

The gyroscope system contains fixated electrodes between which a mass is vibrating along the drive axis. When no rotation is experienced by the gyroscope the capacitance will remain constant as no displacement of the mass relative to the fixated electrodes is present. When the gyroscope rotates a secondary vibration will affect the mass along the perpendicular axis relative to the drive axis. This results in a displacement of the mass relative to the fixated electrodes which then causes a change in capacitance. As with the accelerometer the perpendicular force (F_c) experienced by the mass can be obtained by using Hooke's law. (Arun Faisal et al., 2019)

2.3 Analog to digital conversion

To transform the analogue electrical signal from the accelerometer and gyroscope to the digital domain a analog-to-digital converter (ADC) is used. Prior to that the analogue signal will be amplified in order to be able to store the signal digitally on a computer, as the electrical voltage in the IMU is weak. This is often performed using a differential amplifier which amplifies the difference of two input leads with the advantage of eliminating noise in the input signal. The Shimmer3 IMUs used in this study (see Section 3.4.4 'Measurement equipment'), has an integrated 12-bit ADC corresponding to an input range of 4095 intervals $(2^{12}-1 = 4095)$. (Samani, 2020) The accelerometer of the IMUs in this study was configured to ±16 g (1g \approx 9,82 m/s²) which results in a measurement sensitivity equal to 32 g / 4095 = 0.0078 g (7.8 mg) which translates to 7.8 mg per least significant bit (LSB) (7.8 mg/LSB). Further the gyroscope was configured to a measurement range of \pm 2000 degrees per second (°/s) corresponding to a digital signal sensitivity of 4000 °/s / 4095 = 0.98 °/s which translated to 0,98 °/s /LSB. This means that the analogue signal is transformed into a digital signal with intervals of 7.8 mg and 0.98 °/s for the accelerometer and gyroscope, respectively. The measured kinematic data contain both the true signal representing the movement of the body and signal noise.

Furthermore the sampling frequency was configured to 512 Hz which was within the capacity of the IMU both regarding data storage volume compared to the data collection time and data write speed to its SD-card. No lost data was seen at 512 Hz compared to a sampling frequency of 1024 Hz where loss in data points was seen. As signal noise is most often seen in the higher frequencies, low-pass filters between 40-100 Hz are often used for tibial acceleration measurements to filter out the signal noise (Sheerin et al., 2019). Here a band-pass filter (10-60 Hz) can be used to also filter out the voluntary leg movements to only

reveal the passive impacts received during running (Sheerin et al., 2019), which is used in this study. Therefore the configured sampling frequency of 512 Hz was high enough to comply with Nyquist criterion stating that a signal must be sampled with at least twice the sampling frequency as the maximum frequency component in the signal (Samani, 2020). This means that the sampling frequency has to be at least twice as high as the chosen cut-off frequency (Samani, 2020).

3.0 Method

This section will provide an overview of different approaches on quantifying subjective comfort and its relation to biomechanical and physiological parameters including argumentation for parameters of interest for this study. This is followed by a description of pilot laboratory tests and its findings prior to the experimental protocol. Thereafter the experimental method will describe the participants, shoe conditions and the study design including subsequent data processing and analysis.

3.1 Quantification of subjective comfort

A variety of subjective measurement tools have been used to evaluate and quantify the comfort of shoes, and to find comfort differences between both shoe construction features and objective biomechanical or physiological parameters (Matthias et al., 2021; Menz & Bonanno, 2021). The following will give an overview of these tools resulting in a choice of comfort measurement tool for this study.

3.1.1 Measurement tools for assessing perceived footwear comfort

A systematic review by Matthias et al. (2021), based on 99 articles, identified and synthesized methods for assessing perceived comfort of footwear and the corresponding reliability and validity of these methods. The authors identified four main categories of measurement tools; Visual Analogue Scales (VAS), Questionnaires, Likert-type scales and ranking scales. Of these tools VAS, a unidimensional scale, was the most frequent being used in 59% of articles while the most common length of the VAS used was 100mm. Further the most frequent word anchors used in 32 out of 45 studies was 'not comfortable at all' to 'most comfortable imaginable' when using VAS, see example in Figure 1. The frequency of the other main categories of measurement tools among the 99 articles was found to be; questionnaire (26%), Likert-type scale (9%) and ranking scale (6%) (Matthias et al., 2021). This is supported by a similar and more recent systematic review of 101 articles by Menz & Bonanno (2021) synthesizing literature regarding footwear comfort, subjective measurement scale and findings of factors affecting perceived comfort.



Figure 1 - An example of a 100mm Visual Analogue Scale (VAS) adapted from Bishop et al. (2020) to assess overall shoe comfort. The test subject will mark the line between the left word anchor 'Not comfortable at all' and the right word anchor "most comfortable shoe imaginable" to indicate his or her perceived comfort of the shoe.

Furthermore it was found that only two studies addressed discomfort by including it as anchor words in their measurement scales (Matthias et al., 2021; Menz & Bonanno, 2021). In earlier work it was found that comfort and discomfort are not based on the same factors. Contradictory to discomfort, comfort is associated with feelings of well-being, relaxation and even to the aesthetics of a design whereas discomfort is mainly associated with feelings of pain, soreness, numbness and tiredness. The feelings of discomfort increases over time in the setting of a workday and are assumed to be a result of physical factors such as tissue

pressure, circulation blockage and joint angles (De Looze et al., 2003; Helander & Zhang, 1997) Therefore one might argue that discomfort could be an appropriate measure as it may omit effects of factors like aesthetics and focus on the physical effect of a shoe though effects of shoe appearance may be reduced using different blinding approaches (Matthias et al., 2021). Still, in relation to shoe tests, comfort is an important aspect as it has been shown to be related to different biomechanical and physiological parameters (see section 3.2 'Biomechanical and physiological assessments...')

3.1.2 Reliability and validity of comfort assessments

Earlier studies have not extensively investigated the reliability of footwear comfort measurement tools and only few studies have assessed the validity during assessment tool development, resulting in the literature being fragmented and often inconsistent (Matthias et al., 2021; Menz & Bonanno, 2021). However 100 mm and 150 mm VAS have shown to have a moderate to high reliability and therefore may give reliable assessments of comfort (Matthias et al., 2021; Menz & Bonanno, 2021), while 5-point likert-type scales had a moderate reliability (Matthias et al., 2021). Intra-class correlation coeffients (ICC) of 0.67 and 0.705 (interday) and 0.889 (intraday) have been reported using the VAS (Bishop et al., 2020; Lindorfer et al., 2019). While an ICC of 0.63 has been reported for the likert-type scale (Lindorfer et al., 2019). Compared to the rating scales, ranking comfort of footwear based on preference is indicated to be a more reliable tool (Lindorfer et al., 2019; Mills et al., 2010), though only a limited number of studies have included ranking. Even though ranking offers a more reliable assessment it gives no indication to which degree comfort is experienced, and makes comparison between studies difficult (Matthias et al., 2021). In regard to self-report instruments; content, criterion and construct, are three types of validity that are considered (Brown, 2010). Content validity refers to a representative sample of features of interest included in the measure. Criterion validity refers to comparing variables considered a gold standard or more direct measure to the feature of interest. Lastly construct validity refers to scores behaving as theoretically expected which in the case of perceived comfort would be scores differing or distinguishing between shoes with theoretically different comfort levels (Matthias et al., 2021). Former studies have often based their comfort assessment items on the researchers themselves instead of the population of interest, and therefore may not accommodate the content validity (Matthias et al., 2021). Though one exception was found, a study by Bishop et al. (2020) identified items of interest (Heel cushioning, forefoot cushioning, shoe stability and forefoot flexibility) based on a questionnaire, regarding meaningful items of comfort, of 282 recreational runners (Matthias et al., 2021). This was done by the authors in order to ensure meaning and relevance of the included features (Bishop et al., 2020). Due to the lack of a gold standard measure of perceived comfort, criterion validity is more complicated. However, some comfort items as heel cushioning is shown to be associated with material properties as heel midsole stiffness, therefore giving some means of validating comfort items. Lastly construct validity, testing shoes with different material properties of for example midsole hardness, has been demonstrated for the VAS tool. (Matthias et al., 2021) This gives an indication that using VAS as a measurement tool, gives the best reliability of comfort assessment when the level of comfort is of interest. Furthermore, that using the identified comfort items of Bishop et al. (2020) can increase the level of meaning and relevance of the comfort assessment as his population of interest is corresponding to the population of interest in this study as described in the problem analysis. And lastly while getting information of comfort perception of more detailed aspects of the shoe as compared to only overall global perception of comfort using a standalone 100 mm VAS.

3.1.3 Initial or prolonged comfort

As this study will be assessing perceived comfort after a short distance run (400m) it can be assumed that the results are primarily a matter of initial comfort, which is the measure of interest for this study. This running distance is similar to several footwear comfort assessment studies. These have included shorter run distances between 160 m and 450 m for each footwear condition (C. K.-Y. Lam et al., 2018; Meyer et al., 2018; Mündermann et al., 2002) with a warm-up familiarization procedure of 900m (Mündermann et al., 2002) or 80 m to 200 m per footwear condition (C. K.-Y. Lam et al., 2018; Meyer et al., 2018). It has been proposed that test subjects should complete a warmup of 7 minutes in order to familiarize with the running environment (Huang et al., 2022; Mohr et al., 2021). The above-mentioned studies have used running lengths for warmup that might be completed in a shorter duration than 7 minutes, and therefore might result in the testperson not being familiarized with the running surface before change of shoe condition. Also, the movement pattern during running does not necessarily change when exposed to different footwear conditions (Mohr et al., 2021), which does not exclude that there might be a familiarization process as a result of being exposed to a new shoe condition. Due to the running lengths of this study and the above-mentioned studies, each in a new shoe or insole condition, one might argue that the comfort assessment is exclusively a matter of initial comfort. And not prolonged comfort as opposed to studies assessing footwear comfort in long-distance running as no steady state running is guaranteed. This indicates firstly that the warm-up procedure should be at least 7 minutes of running on the given test surface and that the comfort assessment is regarding initial comfort when running a short distance prior to assessment.

Therefore when comfort is mentioned in the following chapters it will imply the measure of initial comfort and not that of prolonged or continuous comfort. This is also the reasoning behind assessing comfort as opposed to discomfort, as discomfort increases over time.

Choice of measurement tool

On the basis of the above VAS was chosen for this study as a measurement tool of perceived footwear comfort, shown by former studies to provide a higher reliability than the opposing tools (Matthias et al., 2021; Menz & Bonanno, 2021). The items involved in the VAS questionnaire of this study will include 'overall shoe comfort' and 'Heel cushioning', 'forefoot cushioning', 'shoe stability' and 'forefoot flexibility' (Bishop et al., 2020) to increase the level of meaning and relevance of the comfort assessment (Matthias et al., 2021). The anchor words used will include comfort instead of discomfort as discomfort would be more interesting in relation to longer running procedures as discomfort increases over time (Helander & Zhang, 1997). Therefore the VAS will have the anchor words 'not comfortable at all' to 'most comfortable imaginable' and the length of 100 mm as used by the majority of former studies. (Matthias et al., 2021; Menz & Bonanno, 2021).

3.2 Biomechanical and physiological assessments and their relation to perceived footwear comfort

In order to get insight into how footwear affects the body, several studies have tried to relate a broad selection of objective measures to perceived footwear comfort and running related injuries. These studies use an explorative approach within biomechanical, both kinetically and kinematically, and physiological effects of different shoes. The following will give an overview of these studies concluding in the choice of objective measures for this study.

3.2.1 Kinetic recordings

In regards to kinetic analysis, insole pressure systems have been used to explore its relation to comfort through investigating parameters such as peak pressure, maximum force, pressure-time integral and contact area often dividing the foot into several segments (Chen et al., 1995; Jordan et al., 1997; Wegener et al., 2008; Yang et al., 2019). Early studies such as Jordan et al. (1997) have suggested that peak pressure, maximum force and contact area are related to perceived comfort. In that latter study, 20 males volunteered and were tested during walking on a 10-meter walkway. Supporting this, lower peak pressure and mean plantar pressure has shown to be associated with higher total perceived comfort of a shoe during overground running (Yang et al., 2019). In contradiction a crossover randomized controlled study by Wegener et al. (2008) concluded that perceived footwear comfort did not relate to a reduction in plantar pressure. This was further supported by Dinato et al. (2015) which also did not find any significant relations between the aforementioned plantar pressure parameters and comfort. Another approach has been suggested by Lam et al. (2018) to quantify the 'ride' of the shoe defined as the peak anterior-posterior velocity of the center of pressure indicating the transition through the stance phase. The study showed that a lower peak CoP velocity was related to a significantly higher perceived shoe comfort among 13 subjects running with two different shoes (Soft midsole and hard midsole) (C. K.-Y. Lam et al., 2018). The above indicates that no clear or strong evidence is provided for the relation of one or more biomechanical insole pressure parameters to perceived comfort. This might be partly explained by the use of largely varying protocol designs often with little provided detail for acclimatization periods and amount of analyzed steps (Melvin et al., 2014) and partly explained by the difficulty of assessing comfort (De Looze et al., 2003; Helander & Zhang, 1997).

Further kinetic approaches include the use of force plates to measure ground reaction forces (GRF) to analyze peak impact force, time to peak impact force and loading rates (Dinato et al., 2015; Liu et al., 2021; Mündermann et al., 2003). Force plates in combination with motion capture systems are also used to investigate parameters as ankle- and knee moments (Liu et al., 2021; Mündermann et al., 2003). A study by Liu et al. (2021) investigating the influence of three different heel curvature designs (short-parallel, oblique and long-parallel) on 20 recreational runners and a study by Dinato et al. (2015) investigating the effect of four different cushioning conditions on 22 recreational runners, found no significant relationship between force plate kinetics and perceived comfort. The authors concluded that one cannot predict perceived comfort through impact. Additionally, Mündermann et al. (2003) tested three different insert conditions and a control condition (Posting, custom molding and posting + custom molding) on 21 recreational runners. This study concluded that 15 kinematic, kinetic and EMG variables could explain 34,9% of

differences in perceived comfort, though force plate variables did not show strong standardized regression coefficients (Vertical impact peak: 0.126; Maximum ankle plantarflexion moment: 0.258) (Mündermann et al., 2003).

The effort to link kinetic approaches to perceived comfort has not shown strong evidence for parameters measured with insole pressure sensors or force plates though suggested parameters such as the quantification of ride could be an interesting approach for further studies to explore.

3.2.2 Kinematic recordings

The kinematic approach to link objective measures to perceived comfort includes the use of motion capture systems to estimate joint angles of pelvis, hip, knee- and ankle and further moments in combination with force platforms as aforementioned (W.-K. Lam et al., 2018; Liu et al., 2021; Mündermann et al., 2003). The aforementioned study by Liu et al. (2021) also tested the effect of three different heel curvature conditions on lower extremity biomechanics and perceived comfort. It was found that kinematic parameters such as total eversion range of motion, maximum eversion angle and velocity was related to perceived rearfoot stability and that different heel curvatures resulted in change of knee joint moment though not related to perceived comfort parameters (Liu et al., 2021). Initial foot strike angle and perceived comfort measured with VAS was investigated in 18 basketball players during running at different speeds in three different shoe cushioning conditions (best cushioning, better cushioning and regular cushioning) concluding no change in initial foot strike across conditions and no reported relation to comfort (W.-K. Lam et al., 2018). Lastly, as aforementioned, kinematic, kinetic and EMG variables could explain 34,9% of changes in perceived comfort, though motion capture related variables as Delta tibia internal rotation (0.214), Maximum ankle plantarflexion (0.131), Maximum inversion velocity (0.131) is only shown to be vague predictors for changes in comfort revealed by their standardized regression coefficients (Mündermann et al., 2003).

More simple and less expensive equipment like accelerometers and inertial measurement units (IMU) has also been used to describe different aspects of running kinematics in relation to both injuries and perceived comfort (Aristizábal Pla et al., 2021; W.-K. Lam et al., 2018; Meyer et al., 2018; Svenningsen et al., 2020). Dynamic stability, has in later studies been investigated through the use of accelerometers, represented as the ability to recover from small perturbations or large or external perturbations during locomotion (Svenningsen et al., 2020) which decreases during fatique accumulation (Aristizabal Pla et al., 2021). Here, the maximum Lyapunov exponent has often been used (Crowley et al., 2021). Recently a study by Winter et al. (2020) showed that movement deviations prior to the occurrence of injuries could be identified using an accelerometer fixated to the trunk during a fatigue protocol. Prior to the onset of running related injuries (RRI) an increase in dynamic loading was seen in the posterior-anterior direction represented by the magnitude of impacts applied to the body during the running stance phase (Aristizabal Pla et al., 2021; Winter et al., 2020). Additionally a narrative review by Svenningsen et al. (2020), based on 17 articles assessing dynamic stability parameters, found that dynamic stability during locomotion can be adequately assessed using a tri-axial accelerometer when combining index of cycle repeatability and signal dispersion. Furthermore, a study by Meyer et al. (2018) investigated

kinematic variability on 36 recreational athletes running on an indoor-track using an IMU fixed on the dorsum of the foot. From calculating and dividing the RMS of the mean and standard deviation for both translational and angular tri-axial acceleration during the swing phase, the relative variability was found. Using the measure of relative variability, it was found that it was affected by both shoe conditions and perceived comfort (VAS), mainly affecting frontal and transverse joint rotations during late swing phase, suggesting that the use of an IMU can be used as a surrogate or complimentary measurement when assessing footwear comfort. The study compared the least and most comfortable shoe condition for each participant from five shoes of categories as minimalist, neutral and stability shoes differing in weight, midsole properties, pronation support and heel-to drop. (Meyer et al., 2018) A Parameter such as impact loading during running has been assessed through a wide range of studies, due to its potential link to overuse injuries mostly regarding tibial fatigue fractures (Sheerin et al., 2019). The authors showed that impact forces estimated through tibial acceleration are commonly assessed with segment mounted accelerometers as a proxy measurement for the forces affecting tibia during running. Here peak resultant tibial acceleration is often used as such a proxy measure for impact shocks. Further impact forces can be reduced using softer midsoles while thicker midsoles can attenuate shock during these impacts (Sun et al., 2020). Additionally, seven out of eight studies reported that shoes with softer midsoles were perceived to be more comfortable by participants (Menz & Bonanno, 2021). Therefore tibial acceleration acquired through an accelerometer can be seen as a complimentary measure related to perceived comfort, as lower impact forces is seen with the use of shoes with thicker and softer midsoles which seems to be prefered by the user in regards to comfort. Though the opposite has been shown by Dinato et al. (2015) stating that impact forces cannot be used to predict perception of comfort. This may be caused by the fact that only one of five shoe conditions showed a difference in overall comfort compared to the other shoes. The absolute average difference of the least and most overall comfortable shoe was 2.1 cm on a 100 mm VAS-scale. This might not be a big enough difference in comfort to affect the biomechanics considerably, calling for further studies investigating this issue. (Dinato et al., 2015)

Relating kinematic parameters to perceived comfort and RRI has, as the kinetic approach, not shown strong evidence for parameters measured with motion capture or accelerometry. Nevertheless, ankle joint kinematics measured through motion capture (Liu et al., 2021) and kinematic variability through IMU measurements (Meyer et al., 2018) has indicated a relation to perceived comfort while impact forces estimated via tibial accelerations using an accelerometer might be a potential measure linked to footwear comfort.

3.2.3 Physiological recordings

The exploration of more indirect measures to link the effect of footwear on the human body includes surface electromyography (sEMG) and oxygen consumption protocols (Giandolini et al., 2020; Mündermann et al., 2003; Van Alsenoy et al., 2021). During running the body experiences impact shocks that initiates tissue vibrations which can cause discomfort if exposed to soft tissue vibrations over longer periods. To dampen the soft tissue vibrations it has been suggested that the muscle activity is tuned prior to landing, called "muscle tuning". To maintain comfort an increase in muscle activity during muscle tuning might be causing higher muscular demands while lower damping coefficients is associated with prefered foot strike pattern during what might be a more efficient and optimal running pattern. It is

therefore speculated that an earlier onset of fatigue can be caused by the muscle tuning and its increase in muscle activity. (Giandolini et al., 2020) Testing two midsole conditions on 12 male runners on a 20 m indoor track a study by Giandolini et al. (2020) found that a more viscous midsole was related to a reduced amplitude of soft tissue vibrations and the intensity of muscle activity had an effect on the damping of soft tissue vibrations. This was assessed by the authors using sEMG of the lower limbs in combination with tri-axial accelerometers placed on the gastrocnemius medialis and vastus medialis. Further the most important predictors for changes in perceived comfort was found by Mündermann et al. (2003) to be the changes in the intensity of tibialis anterior sEMG signal for the low frequency band (25 -82 Hz) during pre heel strike (-0,437) and 30% - 100% of the stance phase (0.510) revealed by the standardized regression coefficients. A review and meta-analysis by Van Alsenoy et al. (2021) based on 6 articles concludes that when using the most comfortable shoe compared to the least comfortable shoe a significantly decreased VO2 is measured when reviewing the absolute difference between mean scores of the studies of $-2.06 \, mL \cdot kg^{-1} \cdot min^{-1}$.

All in all, this provides an indication of the use of sEMG and oxygen consumption as potential objective measures to evaluate the effect of footwear on the body and might be an indicator of if the footwear supports a prefered running pattern and therefore not providing unnecessary discomfort to the user.

3.2.4 Choice of objective measurement method

Some of the above-mentioned kinetic, kinematic and physiological approaches have shown to suggest a relation to perceived footwear comfort and might be good additional indicators to accompany measures of comfort. Though no clear or strong evidence is provided for the relation of one or more parameters to perceived comfort which might be explained by comfort being subjective and multifactorial, as described in Section 3.1 'Quantification of subjective comfort' (Menz & Bonanno, 2021; Miller et al., 2000; Nigg et al., 2015; Wegener et al., 2008). Though measurement equipment such as motion capture systems and force plates, regarded as gold standard for measuring movement, do however most often require the testing protocol to be executed inside laboratory conditions. This forces the test to be done within small spaces and the analysis to be of single steps. This limitation is due to transportation issues of the equipment, small calibration spaces and ambient light noise when using reflective marker motion capture while requiring larger economical investments. (Cole et al., 2014) These conditions being in a controlled artificial environment might not replicate natural running patterns which are normally performed during a running exercise (Aristizabal Pla et al., 2021; Svenningsen et al., 2020). Cheaper, more lightweight and easier to use technologies such as accelerometers and IMU's make it possible to implement biomechanical recordings over long distances and can be used both during inside and outside lab protocols. Accelerometers have also been concluded to reliably estimate locomotion variations on predictable and unpredictable irregular surfaces and movement dynamics in ecological environments hence possibly supporting analysis of a more natural running pattern than what can be achieved during inside lab protocols. (Aristizabal Pla et al., 2021; Svenningsen et al., 2020) The use of accelerometry and IMU's, due to price and easy use, might also be a more suitable option for running stores, smaller sized shoe companies or physiotherapists. This could be used as a complimentary measurement to get insight into how running shoes affects the running biomechanics, and possibly the perceived comfort, in order to assess shoes during development or when matching shoes to individual runners.

On the basis of the above, it is of interest to further investigate if one or more objective parameters extracted from accelerometer and IMU can be linked to perceived footwear comfort. The following experimental protocol will therefore be assessing perceived comfort using a VAS-questionnaire for two shoe conditions (see section 3.4.2 'Shoe conditions') during road running while using two IMUs to estimate peak resultant tibial acceleration (Peak TA-R) differences as a proxy measure for impact shocks as described by Sheerin et al. (2019) and kinematic variability differences as suggested by Meyer et al. (2018) respectively. This to investigate if these objective measures are affected by perceived footwear comfort.

3.3 Pilot test

Prior to the data collection, a pilot study was performed in the Qualisys laboratory in Gothenburg. This was performed with the objectives of collecting data to create a custom MATLAB data analysis script and to control and test the foot contact detection methods of Meyer et al. (2018) and Aubol & Milner (2020) of which the latter is also used in Garcia et al. (2021). Lastly it was controlled if the acceleration signals caused saturation of the IMU measurements.

Two pilot test participants, who only participated in the pilot test, ran at two different speeds (10 and 12 km/h) on a treadmill with equipped IMUs on the lower medial tibia and dorsum of the foot on the shoe upper. The instrumentation of the IMUs and running procedure followed the same method as described in Section 3.4.3 'Study design' though running at predetermined speeds. The participants were equipped with Shimmer3 IMUs on the right leg (Weight: 23.6 g; Dimensions: 51 x 34 x 14 mm) (Shimmer Research Ltd, Ireland) and Myon IMUs on the left leg (Weight: 5.3 g; Dimensions: 32.7 x 25.5 x 7.8 mm)(Myon AG, Switzerland). Both the Shimmer3 and Myon IMUs were configured to its maximum acceleration range of ± 16 g and gyroscope range of ± 2000 degrees per second (°/s). The Shimmer3 IMUs saved data on a local SD-card at a sampling frequency of 512 Hz while the Myon IMUs send accelerometer data wirelessly to an AKTOS EMG system receiver at a sampling frequency of 2000 Hz which was synchronized with a Qualisys motion capture system (Göteborg, Sweden).

The synchronized motion capture system recorded the test person in the sagittal plane with a high-speed camera at a sampling frequency of 256 Hz. The video material was used to inspect if the same acceleration signal characteristics occured when comparing the Myon IMUs acceleration signals to the aforementioned studies (Aubol & Milner, 2020; Meyer et al., 2018) at the time of initial foot contact. And further if the same acceleration signal characteristics (shape) were apparent in the Shimmer3 IMUs at the time of the detected foot contacts.

From the pilot test it was concluded that the initial foot contact detection methods by Meyer et al. (2018) and Aubol & Milner (2020) were adequate for the current study. Further it was tested if the detected foot contact time points of the IMU placed on distal medial tibia could adequately be used for the IMU placed on Dorsum after synchronization of IMU data. This

resulted in the Dorsum foot contact detection being slightly displaced compared to the data of Meyer et al. (2018). Therefore, it was chosen to use one step detection method for the IMU placed on tibia (Aubol & Milner, 2020) and one for the IMU placed on Dorsum (Meyer et al., 2018) for the subsequent experimental protocol. Further it could be concluded that no saturation of the IMU measurements was apparent for both the accelerometer and gyroscope why a \pm 16 g accelerometer range and a \pm 2000 °/s gyroscope range was deemed adequate. Lastly a sampling frequency of 512 Hz of the Shimmer IMUs was adequate in regard to the resolution of each gait cycle revealing the same signal characteristics as the Myon IMUs sampling at 2000 Hz. The 512 Hz was further in line with the recommendations of Sheerin et al. (2019).

3.4 Experimental Method

3.4.1 Participants

Fifteen healthy recreational runners participated in this study including seven males and eight females (age: 34.7 ± 6.6 years, weight: 70.8 ± 12.8 kg, height: 172.1 ± 9.0 cm) with a self-reported weekly running mileage of 16.9 ± 9.1 km, see Table 1. The participants were recruited through social media outlets and running clubs located in Sweden. In order to achieve a reasonably uniform group a number of exclusion criteria were determined. As the reliability of comfort assessment is population dependent (Matthias et al., 2021) it was decided to exclude runners with a weekly running mileage shorter than 5 km and longer than 25 km. This to achieve a group that resembles the recreational runner which runs a couple of times a week as this is the intended running group for the running shoe prototype. The participants had to wear shoe size EU 43 for males or EU 39 for women, as these were the given sizes of the running shoe prototypes. Further exclusion criteria were; outside the range of 18-50 years of age, being injured in lower extremity within 3 months prior to the study and pain in lower extremities or lower back when running. All participants gave informed written consent prior to the experimental protocol. The study was performed in accordance with the Helsinki Declaration.

Specification	All (n=15)	Males (n=7)	Females (n=8)
Body mass (kg)	70.8 ± 12.8	82.5 ± 6.2	60.6 ± 6.1
Height (cm)	172.1 ± 9.0	180.1 ± 4.6	165.1 ± 4.9
BMI	23.7 ± 2.6	22.2 ± 2.1	25.5 ± 1.9
Foot length (cm)	25.7 ± 1.5	27.2 ± 0.4	24.4 ± 0.7
Foot width (cm)	9.8 ± 0.6	10.3 ± 0.4	9.4 ± 0.4
Weekly milage (km)	16.9 ± 9.1	13.3 ± 5.8	20.1 ± 10.6

Table 1 - Average values and standard deviations are shown for mass, height, body mass index (BMI), foot length and width, and weekly mileage for all participants, and further divided into males and women.

3.4.2 Shoe conditions

Two different shoe conditions were tested in a balanced randomized order for each participant in this study. The two shoe models were a Nike Pegasus 38 and an Icebug Aura RB9X, see Table 2 and Figure 2 for comparison. No mechanical tests were performed on the shoes. Both shoes were intended as road running shoes with a neutral designed footstep

and had similar shoe constructions. Further the shoes had similar weight, midsole hardness (shore) and inside dimensions in length and width (see Table 2). Though the Aura RB9X (EU size 43) was 42 g (14.3 %) heavier than the Nike Pegasus 38 (EU size 43) and had a measured 3.8 shore (7.5 %) higher hardness than the Nike Pegasus 38 (EU size 43). The shoe weight was measured with a 1g precision weight scale (Soehnle, Germany), the midsole hardness with an expanded rubber hardness tester (GS-701N, Teclock, Japan), the internal shoe length with an internal shoe measurement stick (Springyard, Sweden) and the internal shoe width by measuring the widest point of the insole. The midsole hardness was further derived as an average of 10 measurements on the midsole of each shoe as each measurement varied slightly.

The shoes were chosen to see if differences in comfort and biomechanical IMU parameters, using the chosen method, can be found between similar shoes intended for the same segment opposite to a range of studies (Dinato et al., 2015; Jordan et al., 1997; W.-K. Lam et al., 2018; Luo et al., 2009; Meyer et al., 2018; Wegener et al., 2008; Yang et al., 2019) assessing shoes with different constructions and intended segments.

Table 2 - Comparison of shoe characteristics of Nike Pegasus 38 (shoe 1) and Icebug Aura RB9X (shoe 2) shown in size, weight, midsole hardness, inside length and inside width.

Specification	Shoe 1		Sho	be 2
Model and size (EU)	Pegasus 39	Pegasus 43	Aura 39	Aura 43
Weight (g)	252.0	294.0	258.0	336.0
Midsole hardness (shore)	51.0	49.9	50.5	54.3
Inside length (cm)	25.1	28.1	25.1	28.1
Inside width (cm)	8.8	9.8	9.0	9.8





Figure 2 - Visual comparison of the shoe models Nike Pegasus 38 (Shoe 1) and Icebug Aura RB9X (Shoe 2) shown from the right side, bottom, top, back and front view.

3.4.3 Study design

The experimental protocol of this study was carried out within 1.5 hours, as visualized in Figure 3 (See Appendix 1 'Test protocol' for experimental protocol checklist). Prior to the rest of the protocol each participant signed an informed written consent and filled out a questionnaire regarding their running history and the exclusion criteria of this study after receiving both written and verbal information of the purpose of the study, see section 3.4.1 'Participants'.



Figure 3 - An overview of the experimental protocol divided into the first part of the test inside the laboratory setting (20 min) followed by the second part in-field on an asphalt paved road (70 min). The timeline and sequence of each step is indicated by arrows.

The body height, body mass and foot width and length (Brannock device, The Brannock Device Company, USA) were measured inside a laboratory setting before the following field test. Then a 7-minute warmup procedure was completed on the outside asphalt paved road, in order to achieve familiarization of the running surface as per recommendations (Huang et al., 2022; Mohr et al., 2021). The participant was instructed to find a self-selected steady and comfortable running pace during the warm up. The warm up procedure was divided into two as the participants first ran 3.5 minutes to achieve a steady and comfortable pace. Then the

last 3.5 minutes of the warm up were recorded (Garmin Forerunner 735XT) and the average pace (min/km) was saved for the following running procedure.

Thereafter the participants went through a practice of the calibration and running procedure prior to data collection. This included familiarization of the Garmin Virtual Partner (Garmin Forerunner 735XT) for helping the participant keep the recorded average pace. Also the practice included an instruction in how to answer the VAS-questionnaire and familiarization of its use as the participant assessed their own running shoes during the practice. This practice of the procedure included 200 m of running and the data was not used for further analysis.

The participant was then instrumented with a firmly secured IMU (Shimmer3) on the distal medial tibia proximal to the distal curve caused by the medial malleolus of the right leg (Tibia IMU), as per recommendations (Sheerin et al., 2019) using self-adhesive gauze (Fixomull®), see Figure 4. The Tibia IMU was not removed until the protocol was completed for both shoe conditions.



Figure 4 - Fixation of IMU on the distal medial tibia of the right leg using Fixomull self-adhesive gauze, as per recommendations (Sheerin et al., 2019).

In a balanced randomized order, the participants put on the first of the two shoe conditions while being blinded to the shoe characteristics by wearing blindfolds. For this the test leader assisted with putting on the shoes while the participant did the lace tightening themselves as preferred. Further the participants were instructed, and continuously reminded, to not look at the shoes during the entirety of the test protocol, as blinding is considered important to reduce the risk of bias of comfort assessments (Matthias et al., 2021). Thereafter the second IMU (Shimmer3) was secured on the dorsum of the right foot on the upper of the shoe (Dorsum IMU). First by fixating the IMU to the shoelaces with double-sided adhesive tape (Sanoj Tape, Sweden) which was standardized using the same laces for both shoes for each participant. Then by further fixating the IMU firmly to the shoe with self-adhesive gauze (Fixomull®) on the shoe upper, see Figure 5. After instrumentation of the Dorsum IMU the blindfold was removed from the participant.



Figure 5 - Fixation of IMU on the dorsum of the right foot on the upper of the shoe. 1. A piece of double-sided adhesive tape is placed on standardized laces. 2. The IMU is firmly attached to the tape. 3a. Further fixation of the IMU using Fixomull self-adhesive gauze shown from top. 3b. Further fixation shown from side view with IMU orientation indication.

Prior to the test of each shoe condition the participant was informed with instructions on how to complete the test. The test of each shoe condition began at the starting point when the test leader started the data collection of the IMUs manually. Then the participant would complete an IMU calibration procedure, for subsequent Dorsum IMU data frame rotation. The calibration procedure consisted of a 10-second guiet standing phase followed by three stamps with their right foot and a second 10-second quiet standing phase. Afterwards the participants ran 400 meters at the same velocity as the measured self-selected comfortable pace during the warm up. The speed was controlled by the participant by using the Garmin Virtual Partner set to target the previously saved average pace. The Virtual Partner gives constant visual information as to if you run too fast or too slow. Finally, the participant completed the same IMU calibration procedure as in the beginning, though ending with a 60-second standing phase to ensure no data was lost on the IMUs. The data collection ended when the test leader manually stopped the data collection of the IMUs. After the running procedure the participant assessed the comfort of the shoe condition by filling out the 100 mm VAS-questionnaire consisting of 5 items: 'Overall shoe comfort', 'Heel cushioning', 'forefoot cushioning', 'shoe stability' and 'forefoot flexibility', see section 3.1 'Quantification of subjective comfort' and Appendix 2. The above running procedure was repeated for the next shoe condition after a break of 5-7 minutes, depending on the duration of equipment preparation prior to the next running procedure. The participants ran on the same asphalt paved walking and cycling road during both the warm up and the running procedure for the two shoe conditions. The asphalt road had a minor curvature in the end and was predominantly flat with an elevation gain of 4 meters (Garmin connect, Garmin Ltd.), over the length of the 400 m completed during the running procedure.

3.4.4 Measurement equipment

The Shimmer3 IMUs fixated on the tibia and dorsum of the foot were identical (Weight: 23.6 g; Dimensions: 51 x 34 x 14 mm) (Shimmer Research Ltd, Ireland). The IMUs were configured to record at 512 Hz in agreement with recommendations stating that tibial accelerations can be recorded at a sampling frequency ranging from 300 to 600 Hz (Sheerin et al., 2019). The frame rate of 512 Hz was chosen as it was the power of two with a binary base of 2 within the range of 300-600 Hz which was preferred by the configuration software (ConsensysBASIC v1.6.0). Both the Dorsum IMU and the Tibia IMU were configured to

collect the highest wide-range accelerometer data with a range of ± 16 g (1g ≈ 9.82 m/s²). This was chosen to ensure that the signal range did not exceed the IMU capture range when running as rapid changes in the signal occur during initial ground contact (Sheerin et al., 2019). The Dorsum IMU was further configured to record gyroscope data at a range of ± 2000 degrees per second (°/s) which was the highest selectable range, for the same reasons as the argumentation for the accelerometer range. Both accelerometer and gyroscope data were controlled to be within the capture range during pilot tests. Prior to the experimental protocol of this study the IMUs were calibrated using the 9DOF Calibration Application software (v2.10) as according to the User Manual (9DoF Calibration Application, Rev 2.10a). This was done to ensure that the accelerometer axis would measure 9,82 m/s² when affected by the gravitational pull and 0 m/s² if not, when stationary. And further that the gyroscope axis would measure 0 °/s when stationary (9DoF Calibration Application, Rev 2.10a).

3.4.5 Data processing and analysis

The analysis was conducted using MATLAB (version R2021b, Mathworks, Massachusetts, U.S). A custom script was developed for the data processing, analysis and visualization of the IMU accelerometer and gyroscope data in this study. Prior to the following data processing of the IMUs, the minor deviations from 0 m/s² and 0 °/s in the accelerometers and gyroscope after calibration, when not affected by the gravitational pull and when stationary, were subtracted for each of the three IMU axes.

3.4.5.1 Dorsum IMU - Relative variability

The pre and post calibration procedure was first identified from the Dorsum IMU data of each running procedure. This was done by differentiating the most vertical axis (z-axis) of the IMU relative to the sagittal axis, see Figure 5. Then the time index of the standing phases was identified as periods of at least 7 seconds where the differentiated z-axis signal did not surpass a threshold of 3 m/s³, using the regionprops function. This was found to be adequate during pilot testing. Then the running phase was identified as being present in between the last time index of the second stand phase and the first time-index of the third stand phase, see Figure 6.





Figure 6 - Shows the Dorsum IMU acceleration signal (az) and differentiated acceleration (az_diff) for a participant in the IMU z-axis for the entire calibration and running procedure. The procedure phases are automatically identified as first stand phase (1st stand), calibration stamps (cal. stamps), second stand phase (2nd stand) and the running phase between 'run start' and 'run end' as indicated on the figure. The blue box indicates the stand phase data during quiet standing used for subsequent IMU data coordinate system rotation. The shoe symbol in the lower left corner indicates the IMU z-axis orientation.

Within the time of the identified running phase every initial foot contact was identified following the detection method of Meyer et al. (2018) validated by the authors against a force plate with a 10N threshold sampling at 2400 Hz. The identification of initial foot contacts allowed the IMU signal to be divided into gait cycles. The detection method was performed using the z-axis acceleration. As the z-axis direction of the IMUs used in this study was opposite of the one used by Meyer et al. (2018), the z-axis acceleration signal was inverted. According to the authors the foot strike detection then followed the steps, see Figure 7:

- 1. The z-axis acceleration signal was low-pass filtered with a cut-off frequency of 20 Hz for detection of peaks in the filtered signal present right before initial foot contact.
- 2. The raw z-axis acceleration was differentiated.
- 3. After each peak in the filtered signal the time index of the point where the filtered signal was first lower than the differentiated signal, indicated the foot contact.

To find these peaks of the filtered signal and ignore other present peaks not of interest, the average step duration was first found. From pilot testing it was found that a minimum distance of 75% of the average step duration between positive peaks in the filtered signal resulted in the peaks of interest using the *findpeaks* function. The average step duration was found by identifying the maximum positive peaks of the raw z-axis acceleration representing each stance phase and thereafter dividing the duration of the identified running phase with the number of stance phases.



Figure 7 - Visualization of the method of the foot contact detection for the Dorsum IMU. 1. The peak of the filtered IMU z-axis acceleration (black line) was identified (blue circle) prior to rapid changes in the z-axis acceleration (gray line) identified by the maximum peak (red circle). 2. Following the peak in the filtered acceleration the time

point when the filtered acceleration becomes lower than the differentiated z-axis acceleration (dotted gray line) indicates the initial foot contact (red cross). 3. After IMU data coordinate system rotation, the IMU-data of each axis corresponding to the anatomical coordinate system of the foot was cut into 50 gait cycles in the middle of the defined running phase. In the bottom left it is indicated which coordinate system is referred to and its orientation.

After the initial foot contacts were identified, the coordinate system of the IMU data of each axis, was rotated and aligned with the anatomical coordinate system of the foot, according to the standards of the International Society of Biomechanics referring to the paper of Wu & Cavanagh (1995). This was done for both accelerometer and gyroscope data. See Figure 8 and Figure 9 which shows the difference of the coordinate system between the IMU and the foot and visualizes the following description of the coordinate system rotation of the IMU data. The IMU data was rotated about its x-axis and y-axis as it was assumed that these were the axes most prone to tilt, as the IMU was aligned with the foot length. The rotation of the coordinate system was done by calculating the Dorsum IMU tilt around its x and y-axis using the average x, y and z-axis acceleration measured during the first stand phase of the calibration procedure, indicated on Figure 6. For calculating the tilt around the x-axis and y-axis the following equations were used:

$$\theta_x = \tan^{-1}\left(\frac{a_y}{a_z}\right) \;\;;\;\; \theta_y = \tan^{-1}\left(\frac{a_x}{a_z}\right)$$

Of which θ_x and θ_y are the tilt angles in radians around the x and y-axis, and a_x, a_y and a_z are the average measured acceleration during quiet standing in each axis of the IMU coordinate system.

 θ_x and θ_y were then used to perform a rotation of the IMU coordinate system first around the x-axis and thereafter around the y-axis also described as a roll and pitch respectively, using the following transformation matrices:

 $Roll \ matrix = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_x) & -\sin(\theta_x) \\ 0 & \sin(\theta_x) & \cos(\theta_x) \end{bmatrix} \quad ; \quad Pitch \ matrix = \begin{bmatrix} \cos(\theta_y) & 0 & \sin(\theta_y) \\ 0 & 1 & 0 \\ -\sin(\theta_y) & 0 & \cos(\theta_y) \end{bmatrix}$

The roll and pitch matrices were then multiplied with the data vector for each of the IMU axes (x, y and z):

$$vector_{rotated} = (vector \cdot Roll \ matrix) \cdot Pitch \ matrix$$

$$\downarrow$$

$$vector_{rotated} = \left(\begin{bmatrix} v_1 \\ v_2 \\ v_{n...} \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_x) & -\sin(\theta_x) \\ 0 & \sin(\theta_x) & \cos(\theta_x) \end{bmatrix} \right) \cdot \begin{bmatrix} \cos(\theta_y) & 0 & \sin(\theta_y) \\ 0 & 1 & 0 \\ -\sin(\theta_y) & 0 & \cos(\theta_y) \end{bmatrix}$$

Of which vector_{rotated} is the rotated data vector aligned with the anatomical coordinate system, v_1, v_2 and v_n are elements of the data vector, and θ_x and θ_y are the IMU tilt angles in radians around the x and y-axis.

Finally, after the coordinate system rotation, to match the IMU data with the anatomical coordinate system, the Dorsum IMU z-axis was defined as the anatomical x-axis, the Dorsum IMU y-axis was defined as the anatomical y-axis and the Dorsum IMU x-axis was inverted and defined as the the anatomical z-axis (see, Figure 8):

$$A_x = IMU_z$$
; $A_y = IMU_y$; $A_z = -IMU_x$

Of which A_x , A_y and A_z are the anatomical axes and IMU_x , IMU_y and IMU_z are the IMU axes. In the following all IMU data will be referred to in regards to the specified anatomical coordinate system of the foot.



Figure 8 - Shows the IMU dorsum orientation and placement on the shoe upper (x, y, z), the anatomic standard coordinate system of the foot (x_1 , y_1 , z_1) adapted from Wu & Cavanagh (1995), and the order of coordinate system rotation equal to the calculated IMU tilt (θ) around the IMU x and y-axis respectively.



Figure 9 - Shows the IMU z-axis orientation (black circle) in relation to the anatomic standard coordinate system of the foot (x, y, z) (Wu & Cavanagh, 1995), derived from the IMU acceleration data from the quiet standing phase during the calibration procedure. First the IMU acceleration data is rotated around its x-axis (black dot) whereafter the data is rotated around its y-axis (red dot), aligning the IMU with the anatomic coordinate system.

Thereafter 51 foot contacts in the middle of the IMU signal were identified and used to divide the middle part of the signal into 50 gait cycles, see Figure 7. This omits the first and last gait cycles to ensure data analysis of more constant velocity running and not analysis of acceleration/deceleration phases. Then the identified gait cycles were time normalized from 0% to 100% consisting of the amount of data points equal to the median step duration in measured frames, using the *interp1* function. The median was chosen so that the steps with the longest and shortest duration were approximately equally affected by the normalization procedure. Then, the average and standard deviation (STD) was calculated for the late swing phase for each time point between 75%-95% of the 50 normalized gait cycles, see

Figure 10. Finally the Root Mean Square (RMS) of the mean values and standard deviation values were found and used to calculate the Relative variability (RV) resulting in one RV value for each of the IMU data vectors per participant for each shoe (Meyer et al., 2018):

Relative variablity (RV) =
$$\frac{RMS_{std}}{RMS_{avg}}$$

Where RMS_{std} is the RMS of the standard deviation values and RMS_{avg} is the RMS of the average values.



Means and standard deviations of each time point of time-normalized gait cycles

Figure 10 - The mean and standard deviations (STD) for each time point are shown for the late swing phase (75 - 95% of gait cycle) for both acceleration and angular rotation data of the x, y and z-axis of the anatomic coordinate system, indicated by the foot symbol in the lower left corner.

3.4.5.2 Tibia IMU - peak resultant tibial acceleration

The pre and post calibration procedure was first identified for the Tibia IMU using the same method as for the Dorsum IMU, see Section 3.4.5.1 'Dorsum IMU', with the only difference being the use of the y-axis as the most vertical axes of the IMU coordinate system, see Figure 11.



Figure 11 - Shows the Tibia IMU acceleration signal (ay) and differentiated acceleration (ay_diff) for one testperson in the IMU y-axis for the entire calibration and running procedure. The procedure phases are automatically identified as first stand phase (1st stand), calibration stamps (cal. stamps), second stand phase (2nd stand) and the running phase between 'run start' and 'run end' as indicated on the figure.

Thereafter the acceleration data of the Tibia IMU was band pass filtered using a 2. order Butterworth filter (10 - 60 Hz) as per recommendations (Sheerin et al., 2019). The method of Aubol & Milner (2020) was used to identify initial foot contact for the Tibia IMU, recommended for field test and if some acceleration peak values can be expected to be low which may occur with varying self-selected running velocities, as some participants had an average comfortable running velocity as low as 2.55 m/s (9.2 km/h). Following the method of the authors, the resultant acceleration (modulus) was first calculated as the vector sum of the measured accelerations of all three axes of the IMU, using the following equation:

$$a_{res} = \sqrt{a_x^2 + a_y^2 + a_z^2}$$

Where a_{res} is the resultant acceleration and a_x , a_y and a_z are the acceleration vectors for each of the IMU coordinate system axes. This square of each acceleration vector and square root of the sum is used to accommodate both negative and positive accelerations that occur during running. As a rapid increase is present in the resultant acceleration signal right after foot contact, these rapid changes were first identified as maximum peaks for the resultant acceleration signal of the running phase, using the *findpeaks* function. Thereafter the local minima occurring within 75 ms before each of the identified maximum peaks was identified and considered initial foot contact, see Figure 12. (Aubol & Milner, 2020; Garcia et al., 2021). With the identified foot contact times, the 51 foot contacts in the middle of the resultant acceleration signal were identified and used to divide the resultant acceleration signal into 50 gait cycles. The peak resultant acceleration was then identified within the first 40% of each of the 50 gait cycles and the mean peak resultant acceleration was calculated, see Figure 12. This resulted in one mean peak resultant tibial acceleration value for each running test per participant for each shoe.



Figure 12 - Visualization of the foot contact detection method for the Tibia IMU. 1. The resultant acceleration is first calculated from the x, y and z-axis accelerations. 2. The maximum peak (red circle) of the resultant acceleration signal is found and the minimum acceleration (blue circle) is found within 75 ms prior to the maximum peak indicating foot contact. 3. The resultant acceleration signal is cut into 50 gait cycles within the middle of the signal from the identified foot contacts.

3.4.5.3 VAS questionnaire

For each participant one 100 mm VAS questionnaire was answered after the running protocol for each of the two shoe conditions, see Appendix 2. The participant would draw a line on the 100 mm scale indicating their perceived comfort for each of the five comfort items; 'overall shoe comfort' and 'Heel cushioning', 'forefoot cushioning', 'shoe stability' and 'forefoot flexibility'. The distance for each comfort item was then measured from the left word anchor 'not comfortable at all' (0mm) to the drawn line on the scale towards the right word anchor 'most comfortable imaginable', see Figure 13. This resulted in five comfort scores for each shoe condition per participant.

Heel cushioning

Please rate your perception of the cushioning underneath your heel when running in the shoe The further to the right, you perceived the cushioning in the heel to be more comfortable when running. The further to the left, you perceived the cushioning in the heel to be less comfortable. (Too much or too little cushioning)



Figure 13 - Comfort assessment of the comfort item 'Heel cushioning', assessed by drawing a vertical line on the 100 mm Visual Analogue Scale (VAS) adapted from Bishop et al. (2020) (indicated by a vertical black line). The length is measured from 'not comfortable at all' (0 mm) to the vertical black line. The Heel cushioning comfort score of this example would correspond to 67 mm.

3.4.5.4 Statistical analysis

Mean ± standard deviation (STD), median and Interquartile Range (25-75 IQR) are presented. All statistical analysis was carried out in SPSS Statistics (v. 28.0, 64-bit edition, IBM corp., United States). Each participant was represented by one average or median value for each of the six RV variables and one for the peak resultant tibial acceleration from the IMU results and the five VAS questionnaire items for each shoe condition. The variables were both divided into Shoe 1 and Shoe 2, Nike Pegasus and Icebug Aura respectively, and additionally divided into running procedure 1 and running procedure 2 (running the first and second time, respectively). This was done to both see if any significant differences between the variables occurred between the two shoe conditions and to see if any significant differences between the first and second running procedure and therefore check for order effect (carry-over effect). For the data divided into the first and second running procedure half the participants started with shoe 1 and the other half started with shoe 2. Further half the males and half the females started with shoe 1 and the other half started with shoe 2.

To compare the means of the variables between Shoe 1 and Shoe 2 and between Test 1 and Test 2 and to test for significant differences, a parametric paired-samples t-test was chosen, if the data did not violate the statistical assumptions. If the assumption was violated a non-parametric Wilcoxon signed-ranks test was carried out to compare the variable medians. If the assumptions of the Wilcoxon signed-ranks test were violated a Sign test was carried out. See Figure 14 that shows the flow chart of the statistical approach.



Figure 14 - The flowchart shows the order of steps of the statistical analysis indicated by arrows and which choices are made depending on if violations of the statistical assumptions are present. First the assumptions of no significant outliers and a normal distributed population of variable mean differences for the paired-samples t-test was examined. If no violations were present a Paired-samples t -test was performed. If violations were present a check for the Wilcoxon Signed-rank test was performed. If no violations were present a Sign test was instead performed.

All variable mean differences were first tested for significant outliers by visually assessing boxplots. If outliers were present being more than 1.5 times the box-length they were kept in the analysis as they were not considered extreme enough to necessarily influence the data. To control for a potential influence an additional paired-samples t-test was done on the same data excluding the identified outliers to see if the same conclusion was apparent for both the t-tests with and without the outliers. If the t-test conclusions were inconsistent the outliers were instead considered as a violation of the assumption of no significant outliers. Further if outliers were present being more than 3 times the box-length it was considered a violation of the assumption of no significant outliers. (Zar, 2010)

Then it was tested if the population of variable mean differences was normally distributed using a Shapiro-Wilk test recommended for sample sizes below n=50. The Shapiro-Wilk test has the null hypothesis that the variable in question is a normally distributed population, which is rejected if p < 0.05. The mean differences were further visually analyzed through histograms and Q-Q plots to verify the results from the Shapiro-Wilk test. (Zar, 2010)

A paired samples t-test was performed for normally distributed populations of mean differences with no significant outliers. A Wilcoxon signed-ranks test was performed if one of the populations of the mean differences violated any of the two assumptions, if the assumption of symmetrically distributed differences was not violated. If a violation of the assumption of a symmetric distribution was seen a Sign test was performed instead.

The paired samples t-test tests for differences between the population mean whereas the Wilcoxon signed-ranks test and the Sign test tests for difference between the median of paired observations. The statistical tests have the null hypothesis that the paired population means (parametric test) or medians (non-parametric tests) are equal. If p < 0.05 the null hypothesis is rejected which concludes that the means (parametric test) or medians (non-parametric tests) are significantly different. (Zar, 2010)

A Bonferroni correction was used to compensate for multiple comparisons of the six RV variables from the Dorsum IMU and the five comfort items, as these are somewhat dependent and therefore increases risk of committing a type I error. The Bonferroni correction reduces the α -value by dividing the original α -value with the number of dependent variables analyzed:

$$\alpha_{corrected} = \frac{0.05}{k}$$

Where $\alpha_{corrected}$ is the bonferroni corrected α -value and k is the number of dependent variables. This corresponded to a $\alpha_{corrected}$ of $\alpha = 0.008$ for the six RV variables and $\alpha = 0.01$ for the five comfort variables. Tendencies were reported for p-values below 0.1. Though the Bonferroni correction reduces the risk of committing a type I error one should be aware that it also increases the risk of committing a type II error, and therefore not rejecting the null hypothesis when it should be rejected. (Zar, 2010)

Furthermore, the effect size was calculated (Cohen's d) for each mean comparison, as an indication of how big a difference is present between two samples, as a larger effect size indicates a bigger difference. Here a small effect size is equal to 0.2, a medium effect size equal to 0.5 and a large effect size equal to 0.8. The following equation is used to calculate Cohen's d. (Cohen, 1998):

$$d = \frac{M}{SD}$$

Where d is Cohen's d, M is the mean difference between the two samples and SD is the standard deviation of the mean difference.

Thereafter, an explorative analysis was conducted to see if any correlations between the comfort items and the biomechanical variables were present, if any significant differences were found between both comfort items and between biomechanical variables. For this a correlation coefficient was found (Pearson correlation). The result of the correlation coefficient would be in the range of -1 to 1, with a result close to 0 would indicate an absence of correlation and a result close to 1 or -1 would indicate a high degree of correlation. (Zar, 2010)

4.0 Results

This section will compare the results of the 100 mm VAS comfort items, the RV variables and the peak TA-R first between the two shoe conditions (Shoe 1 and Shoe 2) and thereafter between the two running procedures (Test 1 and Test 2). The visual representation of the results in Figure 15 and 16 are shown as mean \pm STD (bar plots and error bars) for variables used in parametric statistical analysis while median and 25-75 IQR (boxplots) are shown for variables used in non-parametric statistical analysis. No further correlation analysis was performed as no significant differences were found between Shoe 1 and Shoe 2 or Test 1 and Test 2.

4.1 Shoe 1 versus shoe 2

The assumptions test showed that outliers larger than 3 box plot-lengths were present for mean differences of three variables (Shoe stability, RV ax and RV gy) and violation of normal distributed mean differences was present for two variables (RV ax and RV gz) (Shapiro-Wilk, p < 0.05), see Table 3. As a result non-parametric testing (Wilcoxon signed-rank test) was performed for 'Shoe stability', 'RV ax', 'RV gy' and 'RV gz' while parametric testing (Paired-samples t-test) was performed for the remaining variables. The Wilcoxon signed-rank tests were performed after visually inspecting the shape of the distribution of the median differences, as these were considered symmetric and therefore no violation of the assumptions were present. Further outliers larger than 1.5 box-lengths were found for mean differences of six of the remaining variables. These were kept in the further parametric analysis though control paired-samples t-tests without the identified outliers were performed. These showed the same conclusions as the t-tests performed including the outliers, described in the following, indicating that the outliers did not influence the statistical conclusion.

Table 3 - Io	dentified quan	tity of outliers	being >1.5 box	-lengths and >3	box-lengths follow	ved by the computed
p-value of	the Shapiro-V	Vilk test of no	rmality, for the	mean difference	e of each of the	analyzed variables. *
indicates if	a significant	difference is	found indicating	y violation of a r	normally distributed	d population of mean
differences.						

Variable	Outliers >1.5 box-lengths	Outliers > 3 box-lengths	Shapiro-Wilk (sig.)
Shoe comfort	0	0	0.872
Heel cushioning	1	0	0.483
Forefoot cushioning	0	0	0.772
Shoe stability	1	1	0.085
Forefoot flexibility	2	0	0.062
RV ax	1	1	0.005*
RV ay	1	0	0.689
RV az	1	0	0.256
RV gx	2	0	0.091
RV gy	1	1	0.065
RV gz	1	0	0.025*
Peak tibial acceleration	0	0	0.715
Run duration	2	0	0.916

Assumptions	test of difference	s - Shoe 1 vs Shoe 2
Assumptions	cest of anticitatioe	5 0110C 1 45 0110C 2

* indicates a significant difference (p<0.05) rejecting the null-hypothesis

For the comfort items, a general trend towards higher mean VAS scores was seen for Shoe 1 compared to Shoe 2 for four comfort items, see Figure 15. Of these trends the highest was

seen for 'Overall comfort' (Shoe 1: $75.2 \pm 17.1 \text{ mm}$, Shoe 2: $60.0 \pm 24.0 \text{ mm}$). Further a visually high variance of the individual data points is apparent for the comfort items with a trend towards higher standard deviations (STD) for Shoe 2 compared to Shoe 1. No systematic difference was seen between data points for males and females except a visual difference between males and females for the Shoe stability score of Shoe 1.

The results of the RV variables of the Dorsum IMU showed no apparent differences between the variable population mean of Shoe 1 and Shoe 2. The highest RV was seen for the angular velocity around the x-axis (Gyro x) and the acceleration in the z-axis (Acc z) of up to a mean of 46.6 % and 37.8 % respectively followed by the angular velocity around the y-axis (Gyro y) of up to a median of 21.1 %. The remaining RV variables were all below a RV of 16 %. Further a low visual variance in the individual data points was seen for the RV variables except acceleration in the z-axis (Acc z) and angular velocity around the x and y-axis (Gyro x and Gyro y). No systematic difference between data points for males and females was seen. For the peak resultant tibial acceleration from the Tibia IMU no apparent visual differences were seen between Shoe 1 and Shoe 2.



Figure 15 - Mean values ± STD (indicated by bar plots and error bars) are shown for variables included in the parametric statistical tests while Median values and 25-75 IQR (indicated by boxplots) are shown for variables included in the non-parametric tests. Further the individual data points are shown for Shoe 1 and Shoe 2 for each of the five comfort items, six relative variability (RV) variables of the Dorsum IMU and peak resultant tibial acceleration of the Tibia IMU. The individual data points are further divided into 'Males' and 'Females' indicated by blue and red dots respectively. The x-axis shows the different variables while the y-axis shows the measured units; VAS score, relative variability and acceleration for the comfort items, Dorsum IMU and Tibia IMU respectively.

The paired-samples t-test showed that no statistical differences were present between the means of the variables as none of the p-values < α -values, see Table 4. However 'Shoe comfort' showed a tendency of difference with a medium effect size between Shoe 1 and Shoe 2 as this result would have been a significant difference if the Bonferroni correction was not included. The remaining variables had a low effect size (< Cohen's d = 0.5). Further the results of the Wilcoxon signed-rank test showed that no statistical differences were present between the medians of the remaining variables as none of the p-values < α -values, see Table 5.

Table 4 - The mean difference and its standard deviation (STD) are shown for each variable followed by the t-value, p-value and α -value as a result of the paired-samples t-test between Shoe 1 and Shoe 2. Finally the computed Cohen's d is shown.

Paired-samples t-test - Shoe 1 vs Shoe 2

• • • • • • • • • • • • • • • • • • • •						
Variable	Mean difference	STD	t-value	p-value	α-value	Cohen's d
Shoe comfort	15.2	27.0	2.181	0.047	0.01	0.563
Heel cushioning	10.2	34.7	1.138	0.274	0.01	0.294
Forefoot cushioning	9.0	32.9	1.061	0.307	0.01	0.274
Forefoot flexibility	5.7	20.6	1.065	0.305	0.01	0.275
RV ay	0.1	1.3	0.334	0.744	0.008	0.086
RV az	2.4	15.5	0.608	0.553	0.008	0.157
RV gx	-2.0	15.3	-0.512	0.617	0.008	-0.132
Peak tibial acceleratio	6.5	18.5	1.353	0.197	0.008	0.349
Run duration	-0.3	2.3	-0.440	0.666	0.05	-0.114

* indicates a significant difference rejecting the null-hypothesis

Table 5 - The median difference, statistical z-score, p-value and α -value are shown for each variable as a result of the Wilcoxon signed-rank test between Shoe 1 and Shoe 2.

Wilxocon signed-rank test - Shoe 1 vs Shoe 2						
Variable	Median difference	z-score	p-value	α-value		
Shoe stability	7.00	-1.22	0.221	0.01		
RV ax	-0.31	0.11	0.910	0.008		
RV gy	1.24	-1.36	0.173	0.008		
RV gz	0.06	-0.28	0.776	0.008		

* indicates a significant difference rejecting the null-hypothesis

4.2 Test 1 versus Test 2

The assumption test results showed that outliers larger than 3 box-lengths were present for mean differences of two variables (RV gy and RV gz) and violation of normal distributed mean differences was present for four variables (Shoe stability, RV gx, RV gy and RV gz) (Shapiro-Wilk, p < 0.05), see Table 6. As a result non-parametric testing (Wilcoxon signed-rank test) was performed for 'Shoe stability', 'RV gx', 'RV gy' and 'RV gz' while parametric testing (Paired-samples t-test) was performed for the remaining variables. The Wilcoxon signed-rank tests were performed after visually inspecting the shape of the distribution of the median differences, as these were considered symmetric and therefore no violation of the assumptions were present. Further outliers larger than 1.5 box-lengths were present for mean differences of four of the remaining variables. These were kept in the further analysis though control paired-samples t-tests without the identified outliers were performed. These showed the same conclusions as the t-tests performed including the outliers, described in the following, indicating that the outliers did not influence the statistical conclusion.

Table 6 - Identified quantity of outliers being >1.5 box-lengths and >3 box-lengths followed by the computed p-value of the Shapiro-Wilk test of normality, for the mean difference of each of the analyzed variables. * indicates if a significant difference is found indicating violation of a normally distributed population of mean differences.

Assumptions test of uniferences - rest 1 vs rest 2							
Variable	Outliers >1.5 box-lengths	Outliers > 3 box-lengths	Shapiro-Wilk (sig.)				
Shoe comfort	0	0	0.989				
Heel cushioning	1	0	0.801				
Forefoot cushioning	0	0	0.976				
Shoe stability	2	0	0.038*				
Forefoot flexibility	2	0	0.086				
RV ax	2	0	0.054				
RV ay	0	0	0.873				
RV az	0	0	0.146				
RV gx	2	0	0.006*				
RV gy	2	1	0.016*				
RV gz	0	1	0.042*				
Peak tibial acceleration	0	0	0.988				
Run duration	2	0	0.851				

Assumptions test of differences - Test 1 vs Test 2

* indicates a significant difference (p<0.05) rejecting the null-hypothesis

For the comfort items a general trend of higher mean VAS scores was seen for Test 2 compared to Test 1 for four comfort items, of which the highest trend was seen for 'Heel cushioning' (Test 1: 55.5 mm, Test 2: 70.7 mm), see Figure 16. A visually high variance of the individual data points was apparent for all five comfort items. Further, no systematic difference was seen between data points for males and women, except a visual difference between males and females for the Shoe stability score for Test 2.

No apparent visual differences were seen between the mean or median of the six RV variables between Test 1 and Test 2. The highest RV was seen for the angular velocity around the x-axis (Gyro x) and the acceleration in the z-axis (Acc z) of up to a median of 41.4 % and a mean of 38.7% respectively followed by the angular velocity around the y-axis (Gyro y) of up to a median of 21.7 %. The remaining RV variables were all below a RV of 16%. Further no systematic difference between data points for males and females was seen

For the peak resultant tibial acceleration from the Tibia IMU no apparent visual differences were seen between Test 1 and Test 2.



Test 1 versus Test 2

Figure 16 - Mean values ± STD (indicated by bar plots and error bars) are shown for variables included in the parametric statistical tests while Median values and 25-75 IQR (indicated by boxplots) are shown for variables included in the non-parametric tests. Further the individual data points are shown for Shoe 1 and Shoe 2 for each of the five comfort items, six relative variability (RV) variables of the Dorsum IMU and peak resultant tibial acceleration of the Tibia IMU. The individual data points are further divided into 'Males' and 'Females' indicated by blue and red dots respectively. The x-axis shows the different variables while the y-axis shows the measured units; VAS score, relative variability and acceleration for the comfort items, Dorsum IMU and Tibia IMU respectively.

The paired-samples t-test showed that no statistical differences were present between the means of the variables as none of the p-values < α -values, see Table 7. Though 'Heel cushioning' showed a tendency of difference with a small effect size between Test 1 and Test 2. The remaining variables showed a small effect size (< Cohen's d = 0.5). Lastly the results of the Wilcoxon signed-rank test showed that no statistical differences were present between the medians of the variables as none of the p-values < α -values, see Table 8.

Table 7 - The mean difference and its standard deviation (STD) are shown for each variable followed by the t-value, p-value and α -value as a result of the paired-samples t-test between Test 1 and Test 2. Finally the computed Cohen's d is shown.

Paired-samples t-test - Test 1 vs Test 2							
Variable	Mean difference	STD	t-value	p-value	α-value	Cohen's d	
Shoe comfort	-2.7	31.1	-0.332	0.745	0.01	-0.086	
Heel cushioning	-15.1	32.7	-1.791	0.095	0.01	-0.462	
Forefoot cushioning	-6.1	33.6	-0.700	0.495	0.01	-0.181	
Forefoot flexibility	-6.7	20.3	-1.287	0.219	0.01	-0.332	
RV ax	-0.2	4.3	-0.172	0.866	0.008	-0.044	
RV ay	-0.5	1.2	-1.577	0.137	0.008	-0.407	
RV az	4.1	15.1	1.045	0.314	0.008	0.270	
Peak tibial acceleratio	-0.1	19.6	-0.014	0.989	0.008	-0.004	
Run duration	0.5	2.3	0.900	0.384	0.05	0.232	

* indicates a significant difference rejecting the null-hypothesis

Table 8 - The median difference, statistical z-score, p-value and α -value are shown for each variable as a result of the Wilcoxon signed-rank test between Test 1 and Test 2.

Wilxocon signed-rank test - Test 1 vs Test 2

White consigned fails test i vs fest z								
Variable	Median difference	z-score	p-value	α-value				
Shoe stability	3.00	0.028	0.977	0.01				
RV gx	-0.65	-0.06	0.955	0.008				
RV gy	-1.03	0.51	0.609	0.008				
RV gz	-0.34	1.31	0.191	0.008				

* indicates a significant difference rejecting the null-hypothesis

5.0 References

- Agresta, C. E., Peacock, J., Guadagnino, S., Carmichael, A., & Gonzalez, R. (2021). The test-retest reliability of ride perception during treadmill and outdoor running. *Footwear Science*, 1–11. https://doi.org/10.1080/19424280.2021.1994021
- Aristizábal Pla, G., Hollville, E., Schütte, K., & Vanwanseele, B. (2021). The Use of a Single Trunk-Mounted Accelerometer to Detect Changes in Center of Mass Motion Linked to Lower-Leg Overuse Injuries: A Prospective Study. *Sensors*, *21*(21), 7385. https://doi.org/10.3390/s21217385
- Arun Faisal, I., Waluyo Purboyo, T., & Siswo Raharjo Ansori, A. (2019). A Review of Accelerometer Sensor and Gyroscope Sensor in IMU Sensors on Motion Capture. *Journal of Engineering and Applied Sciences*, *15*(3), 826–829. https://doi.org/10.36478/jeasci.2020.826.829
- Aubol, K. G., & Milner, C. E. (2020). Foot contact identification using a single triaxial accelerometer during running. *Journal of Biomechanics*, *105*, 109768. https://doi.org/10.1016/j.jbiomech.2020.109768
- Bishop, C., Buckley, J. D., Esterman, A. E., & Arnold, J. B. (2020). The running shoe comfort assessment tool (RUN-CAT): Development and evaluation of a new multi-item assessment tool for evaluating the comfort of running footwear. *Journal of Sports Sciences*, 38(18), 2100–2107. https://doi.org/10.1080/02640414.2020.1773613
- Brown, T. (2010). Construct Validity: A Unitary Concept for Occupational Therapy
 Assessment and Measurement. *Hong Kong Journal of Occupational Therapy*, 20(1),
 30–42. https://doi.org/10.1016/S1569-1861(10)70056-5
- Chen, H., Nigg, B. M., Hulliger, M., & de Koning, J. (1995). Influence of sensory input on plantar pressure distribution. *Clinical Biomechanics*, *10*(5), 271–274. https://doi.org/10.1016/0268-0033(95)99806-D
- Cole, M. H., van den Hoorn, W., Kavanagh, J. K., Morrison, S., Hodges, P. W., Smeathers, J. E., & Kerr, G. K. (2014). Concurrent Validity of Accelerations Measured Using a

Tri-Axial Inertial Measurement Unit while Walking on Firm, Compliant and Uneven Surfaces. *PLoS ONE*, *9*(5), e98395. https://doi.org/10.1371/journal.pone.0098395

- Crowley, P., Vuillerme, N., Samani, A., & Madeleine, P. (2021). The effects of walking speed and mobile phone use on the walking dynamics of young adults. *Scientific Reports*, *11*(1), 1237. https://doi.org/10.1038/s41598-020-79584-5
- Dadafshar, M. (2015). Accelerometer and gyroscopes sensors: Operation, sensing and applications (5830). Maxim Integrated.
- De Looze, M. P., Kuijt-Evers, L. F. M., & Van Dieën, J. (2003). Sitting comfort and discomfort and the relationships with objective measures. *Ergonomics*, *46*(10), 985–997. https://doi.org/10.1080/0014013031000121977
- Dinato, R. C., Ribeiro, A. P., Butugan, M. K., Pereira, I. L. R., Onodera, A. N., & Sacco, I. C. N. (2015). Biomechanical variables and perception of comfort in running shoes with different cushioning technologies. *Journal of Science and Medicine in Sport*, *18*(1), 93–97. https://doi.org/10.1016/j.jsams.2013.12.003
- Garcia, M. C., Gust, G., & Bazett-Jones, D. M. (2021). Tibial acceleration and shock attenuation while running over different surfaces in a trail environment. *Journal of Science and Medicine in Sport*, *24*(11), 1161–1165.

https://doi.org/10.1016/j.jsams.2021.03.006

Giandolini, M., Romain, J.-P., Horvais, N., & Nigg, B. M. (2020). Midsole Properties Affect the Amplitude of Soft Tissue Vibrations in Heel-Toe Runners. *Medicine and Science in Sports and Exercise*, *52*(4), 884–891.

https://doi.org/10.1249/MSS.00000000002194

- Helander, M. G., & Zhang, L. (1997). Field studies of comfort and discomfort in sitting. *Ergonomics*, *40*(9), 895–915. https://doi.org/10.1080/001401397187739
- Hennig, E. M. (2011). Eighteen years of running shoe testing in Germany—A series of biomechanical studies. *Footwear Science*, *3*(2), 71–81. https://doi.org/10.1080/19424280.2011.616536

Huang, M., Mo, S., Pak-Kwan Chan, P., Chan, Z. Y. S., Zhang-Lea, J. H., & Cheung, R. T. H.

(2022). The influence of running shoes on familiarization time for treadmill running biomechanics evaluation. *Sports Biomechanics*, 1–14. https://doi.org/10.1080/14763141.2022.2046144

- Jordan, C., Payton, C. J., & Bartlett, R. M. (1997). Perceived comfort and pressure distribution in casual footwear. *Clinical Biomechanics*, *12*(3), S5. https://doi.org/10.1016/S0268-0033(97)88312-X
- Koplan, J. P., Rothenberg, R. B., & Jones, E. L. (1995). The natural history of exercise: A
 10-yr follow-up of a cohort of runners. *Medicine and Science in Sports and Exercise*,
 27(8), 1180–1184.
- Lam, C. K.-Y., Mohr, M., Nigg, S., & Nigg, B. (2018). Definition and quantification of 'ride' during running. *Footwear Science*, *10*(2), 77–82. https://doi.org/10.1080/19424280.2018.1453872
- Lam, W.-K., Liebenberg, J., Woo, J., Park, S.-K., Yoon, S.-H., Cheung, R. T.-H., & Ryu, J.
 (2018). Do running speed and shoe cushioning influence impact loading and tibial shock in basketball players? *PeerJ*, *6*, e4753. https://doi.org/10.7717/peerj.4753
- Lambrecht, S., & del-Ama, A. J. (2014). Human Movement Analysis with Inertial Sensors. In
 J. L. Pons & D. Torricelli (Eds.), *Emerging Therapies in Neurorehabilitation* (Vol. 4, pp. 305–328). Springer Berlin Heidelberg.
 https://doi.org/10.1007/978-3-642-38556-8_16
- Lindorfer, J., Kröll, J., & Schwameder, H. (2019). Comfort assessment of running footwear: Does assessment type affect inter-session reliability? *European Journal of Sport Science*, *19*(2), 177–185. https://doi.org/10.1080/17461391.2018.1502358
- Liu, Z.-L., Lam, W.-K., Zhang, X., Vanwanseele, B., & Liu, H. (2021). Influence of heel design on lower extremity biomechanics and comfort perception in overground running. *Journal of Sports Sciences*, 39(2), 232–238. https://doi.org/10.1080/02640414.2020.1813410
- Luo, G., Stergiou, P., Worobets, J., Nigg, B., & Stefanyshyn, D. (2009). Improved footwear comfort reduces oxygen consumption during running. *Footwear Science*, *1*(1),

25-29. https://doi.org/10.1080/19424280902993001

- Matthias, E. C., Banwell, H. A., & Arnold, J. B. (2021). Methods for assessing footwear comfort: A systematic review. *Footwear Science*, *13*(3), 255–274. https://doi.org/10.1080/19424280.2021.1961879
- Melvin, J. M. A., Preece, S., Nester, C. J., & Howard, D. (2014). An investigation into plantar pressure measurement protocols for footwear research. *Gait & Posture*, 40(4), 682–687. https://doi.org/10.1016/j.gaitpost.2014.07.026
- Menz, H. B., & Bonanno, D. R. (2021). Footwear comfort: A systematic search and narrative synthesis of the literature. *Journal of Foot and Ankle Research*, *14*(1), 63. https://doi.org/10.1186/s13047-021-00500-9
- Meyer, C., Mohr, M., Falbriard, M., Nigg, S. R., & Nigg, B. M. (2018). Influence of footwear comfort on the variability of running kinematics[†]. *Footwear Science*, *10*(1), 29–38. https://doi.org/10.1080/19424280.2017.1388296
- Miller, J. E., Nigg, B. M., Liu, W., Stefanyshyn, D. J., & Nurse, M. A. (2000). Influence of foot, leg and shoe characteristics on subjective comfort. *Foot and Ankle International*, 21(9), 759–767. https://doi.org/10.1177/107110070002100908
- Mills, K., Blanch, P., & Vicenzino, B. (2010). Identifying Clinically Meaningful Tools for Measuring Comfort Perception of Footwear. *Medicine & Science in Sports & Exercise*, *42*(10), 1966–1971. https://doi.org/10.1249/MSS.0b013e3181dbacc8
- Mohr, M., von Tscharner, V., Nigg, S., & Nigg, B. M. (2021). Systematic reduction of leg muscle activity throughout a standard assessment of running footwear. *Journal of Sport and Health Science*, S209525462100003X.

https://doi.org/10.1016/j.jshs.2021.01.003

- Molloy, J. M. (2016). Factors Influencing Running-Related Musculoskeletal Injury Risk Among U.S. Military Recruits. *Military Medicine*, *181*(6), 512–523. https://doi.org/10.7205/MILMED-D-15-00143
- Mündermann, A., Nigg, B. M., Humble, R. N., & Stefanyshyn, D. J. (2003). Orthotic comfort is related to kinematics, kinetics, and EMG in recreational runners. *Medicine and*

Science in Sports and Exercise, 35(10), 1710–1719.

https://doi.org/10.1249/01.MSS.0000089352.47259.CA

- Mündermann, A., Nigg, B. M., Stefanyshyn, D. J., & Humble, R. N. (2002). Development of a reliable method to assess footwear comfort during running. *Gait and Posture*, *16*(1), 38–45. https://doi.org/10.1016/S0966-6362(01)00197-7
- Mündermann, A., Stefanyshyn, D. J., & Nigg, B. M. (2001). Relationship between footwear comfort of shoe inserts and anthropometric and sensory factors: *Medicine and Science in Sports and Exercise*, 33(11), 1939–1945.
 https://doi.org/10.1097/00005768-200111000-00021
- Nigg, B. M., Baltich, J., Hoerzer, S., & Enders, H. (2015). Running shoes and running injuries: Mythbusting and a proposal for two new paradigms: "Preferred movement path" and "comfort filter." *British Journal of Sports Medicine*, *49*(20), 1290–1294. https://doi.org/10.1136/bjsports-2015-095054
- Novacheck, T. F. (1998). The biomechanics of running. *Gait and Posture*, 7(1), 77–95. https://doi.org/10.1016/S0966-6362(97)00038-6
- Samani, A. (2020). *An introduction to signal processing for non-engineers* (First edition). CRC Press/Taylor & Francis Group.
- Sheerin, K. R., Reid, D., & Besier, T. F. (2019). The measurement of tibial acceleration in runners—A review of the factors that can affect tibial acceleration during running and evidence-based guidelines for its use. *Gait & Posture*, 67, 12–24. https://doi.org/10.1016/j.gaitpost.2018.09.017
- Shields, N., Lim, P., Wollersheim, D., Nikolopoulos, N., Barrett, J., Evans, A., Taylor, N. F., & Munteanu, S. (2017). Do foot posture, deformity, and footwear fit influence physical activity levels in children with Down syndrome? A prospective cohort study. *Journal of Intellectual & Developmental Disability*, *42*(4), 332–338.
 https://doi.org/10.3109/13668250.2016.1237623
- Sun, X., Wing-Kai, L., Zhang, X., Wang, J., & Fu, W. (2020). Systematic Review of the Role of Footwear Constructions in Running Biomechanics: Implications for

Running-Related Injury and Performance. ©*Journal of Sports Science and Medicine*, *19*(©Journal of Sports Science and Medicine), 20–37.

- Svenningsen, F. P., Pavailler, S., Giandolini, M., Horvais, N., & Madeleine, P. (2020). A narrative review of potential measures of dynamic stability to be used during outdoor locomotion on different surfaces. *Sports Biomechanics*, *19*(1), 120–140. https://doi.org/10.1080/14763141.2019.1642953
- Taunton, J. E. (2002). A retrospective case-control analysis of 2002 running injuries. *British Journal of Sports Medicine*, *36*(2), 95–101. https://doi.org/10.1136/bjsm.36.2.95
- Van Alsenoy, K., van der Linden, M. L., Girard, O., & Santos, D. (2021). Increased footwear comfort is associated with improved running economy – a systematic review and meta-analysis. *European Journal of Sport Science*, 1–13. https://doi.org/10.1080/17461391.2021.1998642
- Wegener, C., Burns, J., Penkala, S., & Sc, G. D. E. S. (2008). Effect of Neutral-Cushioned Running Shoes on Plantar Pressure Loading and Comfort in Athletes with Cavus
 Feet: A Crossover Randomized Controlled Trial. *The American Journal of Sports Medicine*, *36*(11), 2139–2146. https://doi.org/10.1177/0363546508318191
- Winter, S. C., Gordon, S., Brice, S. M., Lindsay, D., & Barrs, S. (2020). A Multifactorial Approach to Overuse Running Injuries: A 1-Year Prospective Study. *Sports Health: A Multidisciplinary Approach*, *12*(3), 296–303. https://doi.org/10.1177/1941738119888504
- Wu, G., & Cavanagh, P. R. (1995). ISB recommendations for standardization in the reporting of kinematic data. *Journal of Biomechanics*, *28*(10), 1257–1261.
 https://doi.org/10.1016/0021-9290(95)00017-C
- Yang, F., Liu, Z.-L., Ma, R.-Y., & Lam, W.-K. (2019). Using plantar pressure to predict insole comfort in running. *Footwear Science*, *11*(sup1), S15–S17. https://doi.org/10.1080/19424280.2019.1606054
- Zar, J. H. (2010). *Biostatistical analysis* (5th ed). Pearson education international.