Identification of subject-specific responses to footwear during running

Fabian Horst [1] ✉, **Fabian Hoitz** [2,3], **Nicolas Schons** [1], **Hendrik Beckmann** [1], **Benno M. Nigg** [3], **Wolfgang I. Schöllhorn** [1]

[1]Department of Training and Movement Science, Institute of Sport Science, Johannes Gutenberg-University Mainz, Germany; [2]Biomedical Engineering, Schulich School of Engineering, University of Calgary, Alberta, Canada; [3]Human Performance Laboratory, Faculty of Kinesiology, University of Calgary, Alberta, Canada

✉ **For correspondence:**
horst@uni-mainz.de (Fabian Horst)

Preprint: This manuscript is not peer reviewed.

Data availability: The data and code will be made public on GitHub upon publication.

Acknowledgment: This preprint was created using the LaPreprint template (https://github.com/roaldarbol/lapreprint) by Mikkel Roald-Arbøl.

Competing interests: The author declare no competing interests.

Citation: Horst, F., Hoitz, F., Schons, N., Beckmann, H., Nigg, B.M. & Schöllhorn, W.I. (2022). Identification of subject-specific responses to footwear during running [Preprint]. *SportRxiv*.

Abstract

Background Placing a stronger focus on subject-specific responses to footwear may lead to a better functional understanding of footwear effects on running and its influence on comfort perception, performance, and pathogenesis of injuries. Here, we investigate subject-specific responses to different footwear conditions within ground reaction force (GRF) data during running using a machine learning-based approach. We conduct our investigation in three steps, guided by the following hypotheses: (I) For each subject x footwear combination, unique GRF patterns can be identified. (II) For each subject, unique GRF characteristics can be identified across footwear conditions. (III) For each footwear condition, unique GRF characteristics can be identified across subjects.

Methods Thirty male subjects ran ten times at their preferred (self-selected) speed on a level and approximately 15 m long runway in five footwear conditions (barefoot, subject's own running shoe, and three standardised running shoes). We recorded three-dimensional GRFs for one right-foot stance phase per running trial and classified the vectorised GRFs using support vector machines.

Results The highest prediction accuracy was found for the subject x footwear classification (hypothesis I). The median prediction accuracy was 95.7 %. This is approximately 137 times higher than the zero-rule baseline (ZRB) of 0.7 %. Across footwear conditions, subjects could be discriminated with a median prediction accuracy of 89.7 % (approximately 27 times higher than the ZRB of 3.3 %). Across subjects, footwear conditions could be discriminated with a median prediction accuracy of 76.3 % (approximately 4 times higher than the ZRB of 20.0 %).

Conclusion Our results suggest that, during running, responses to footwear are unique to each individual subject and footwear design. As a result, considering subject-specific responses contribute to a more differentiated functional understanding of footwear effects. Incorporating holistic biomechanical data is auspicious for the (subject-specific) evaluation of the footwear effects, as unique interactions between subjects and footwear manifest in versatile ways. Machine learning methods have demonstrated their great potential to fathom subject-specific responses when evaluating and recommending footwear.

Introduction

Debates in sport biomechanics have discussed the effects of footwear on sports performance, comfort perception, and injury risks.[1,2] From these debates, footwear designs emerged that were aimed to reduce speculated risk factors of running-related injuries (e.g., excessive pronation or high impact forces). The effects of such footwear designs, however, remain elusive as contradictory findings regarding their influence on injury risks and biomechanics are frequently reported.[3,4] Some studies, for instance, showed that a shoe's midsole hardness affected ground reaction force (GRF) variables (e.g., loading rates, impact forces),[5,6] while other authors reported no (or even opposing) effects.[7–9]

One possible explanation for those contradictory findings could be methodological limitations. Current research strategies commonly focus on average responses to footwear (e.g., using estimates of central tendencies like mean values from groups of individuals). This approach, however, neglects footwear-related effects on individual subjects.[2,10] This is a crucial limitation as substantial differences between anatomy,[11,12] history of previous injuries,[13] and milage[13] were reported across individuals. Runners, therefore, have unique response to different footwear designs.[14] A notion that is supported by the concept of movement signatures: the finding of unique movement patterns for each individual in (barefoot) walking[15,16] and running with one's own shoe.[17,18] Consequently, group-based approaches may have led to an incomplete functional understanding of how footwear affects a subject's (unique) movement.[14,19,20]

Research strategies with a stronger focus on subject-specific responses to footwear designs have been discussed.[21] Those research strategies are categorised into *single-subject*[21] and *functional group*[2,22,23] based approaches. Either approach needs to consider holistic biomechanical data (i.e., several multi-dimensional and time-continuous variables) to map subject-specific responses to footwear because previous findings have shown that movement signatures[24,25] and responses to footwear[4] manifest in multiple interacting variables.

Machine learning models can process several multi-dimensional, time-continuous variables (e.g., three-dimensional lower-body joint angles).[10] Consequently, no reduction of the measured data and no pre-selection of single time-discrete variables is required. The potential of machine learning-based approaches has been demonstrated in previous studies investigating the uniqueness of movement patterns for each individual in (barefoot) walking[15,16] and running with one's own shoe.[17,18,26] Machine learning, therefore, is well suited to analyse the multi-faceted interactions in the responses of subjects to footwear interventions.[27–30]

Here, we continue this work and aim to explore the uniqueness of individual responses to different footwear conditions using a machine learning-based classification via support vector machines (SVMs). We conduct our investigation in three steps, guided by the following hypotheses:

(I) For each subject x footwear combination, unique GRF patterns can be identified (30 subjects x 5 footwear conditions = 150 unique patterns)
(II) For each subject, unique GRF characteristics can be identified across various footwear conditions (30 subjects = 30 unique patterns)
(III) For each footwear condition, unique GRF characteristics can be identified across various subjects (5 footwear conditions = 5 unique patterns)

Methods & Materials

Subjects and ethics statement

Thirty healthy, physically active male subjects (age: 20 - 28 yrs; height: 1.80 - 1.90 m; mass: 71.4 - 100.0 kg) that were free of lower extremity injuries participated in the study. Prior any testing, participants provided written informed consent. All experimental procedures were conducted in accordance with the Declaration of Helsinki and were approved by the ethical committee of the medical association Rhineland-Palatinate in Mainz (Germany).

Experimental protocol

Subjects performed ten running trials at their preferred (self-selected) speed along a level, 15 m long runway in four shod and one barefoot condition. The four running shoes were the New Balance Minimus, Adidas Adistar Boost, ON Cloudsurfer, and the subject's own running shoe. Footwear conditions were counterbalanced across subjects. Prior data collection, subjects performed twenty familiarisation runs in each new condition. Data acquisition took place on a single day in an indoor laboratory.

Data acquisition

Per running trial, three-dimensional GRFs were recorded at a frequency of 1,000 Hz for one right-foot stance phase using a floor embedded force plate (Kistler - Type 9287CA, Switzerland) located half-way along the runway. The recording was processed within the LabView 2010 (National Instruments, USA) framework. Subjects were instructed to focus on a gender-neutral face emoji (i.e., simple, open eyes and a flat, closed mouth) on the opposing wall of the laboratory to direct their visual attention away from targeting the force plate and ensure a natural run with an upright body position.

Data processing

The recorded three-dimensional GRFs - vertical, anterior-posterior, and medio-lateral - were filtered using a second order Butterworth bidirectional low-pass filter at a cut-off frequency of 50 Hz. Stance phase was determined based on the filtered vertical GRF using a threshold value of 10 N. Each GRF signal was time-normalised to 101 data points, corresponding to 100 % stance phase. GRF signals were normalised to the body weight, measured separately for each footwear condition. Data processing was performed exclusively within MATLAB 2021b (MathWorks, USA).

Data analysis

A total of 1,500 GRF recordings were classified using SVMs.[31] The L2-regularised L2-loss support vector classification of the Liblinear Toolbox 1.4.1[32] with a linear kernel function was applied. The regularisation parameter C was experimentally determined using a grid search within the range of C = 2^{-5}, $2^{-4.75}$, . . . , 2^{15} prior model training / testing. GRF signals were min-max normalised to range from 0 to 1 and concatenated when passed to the SVM models. The grid search and determination of normalisation min / max values were conducted exclusively based on recordings that were included in the training data. Three classification tasks were tested: (1) *subject x footwear*, where each subject-footwear combination represented one of 150 (30 subjects x 5 footwear conditions) possible classification outcomes (hypothesis (I)); (2) *subject*, where each subject represented one of 30 possible classification outcomes (hypothesis (II)); and (3) *footwear*, where each footwear represented one of 5 possible classification outcomes (hypothesis (III)).

Performance evaluation

For all classification tasks, a stratified five-fold cross-validation was used to evaluate the classification performance. Additionally, each individual recording was part of the test data once. For the subject x footwear classification, recordings from each combination of subject and footwear condition were distributed equally among the cross-validation folds. For the subject classification and the footwear classification, we ensured for each fold that the recordings of each footwear condition (in the subject classification) and each subject (in the footwear classification) were either part of the training or the test data.

For the footwear classification task, we also tested the effect of the number of subjects used to train the SVM model on the classification performance. For this purpose, a leave-subject out cross-validation was used, in which the number of subjects used for training was iteratively increased. The individual classification performances were compared to the zero-rule baseline, which refers to the theoretical accuracy obtained by assigning class labels according to the prior probabilities of the classes. Specifically, the target labels were set to the class with the largest cardinality in the training dataset, corresponding to 0.7 %, 3.3 %, and 20.0 % for subject x footwear, subject, and footwear, respectively.

Relevance score evaluation

Layer-wise relevance propagation[33] was used to decompose the predictions of the trained SVM models into relevance scores for each value i of the corresponding input vector. The relevance scores R_i were calculated based on the product of each value x_i of the input vector x and the weight w_i of the weight vector w of the trained SVM models:

$$R_i = x_i * w_i \tag{1}$$

Relevance scores indicate which information was used by the SVM model for its prediction. Positive scores represent variables supporting the classification, while negative scores represent variables speaking against a given classification. For this work, the true class labels were decomposed, and only positive input relevance scores were analysed, as negative scores highlight input values did not support the model's prediction of the ground truth class.[33] Subsequently, positive relevance scores were normalised to their respective maximum. All data analysis was performed within the MATLAB 2021b (MathWorks, USA) framework.

Results

Performance evaluation

Across classification tasks, the median prediction accuracy across the five-fold cross-validation was superior to the theoretical task-specific zero-rule baseline accuracy (Figure 1). The highest prediction accuracy was found for the subject x footwear classification task (hypothesis (I)), with a median value of 95.7 % (138 times higher than the respective zero-rule baseline of 0.7 %). For the subject classification task (hypothesis (II)), the median accuracy was 89.7 % (27 x zero-rule baseline of 3.3 %). For the footwear classification task (hypothesis (III)), the median accuracy was 76.3 % (4 x zero-rule baseline of 20.0 %).

In addition to the fold-wise performance evaluation (Figure 1), the prediction accuracy of the SVM models was also summarised according to individual subjects (Figure 2) and individual footwear conditions (Figure 3). The presented results (Figure 1 - 3) are all based on the same SVM models.

The majority of individuals were correctly identified regardless of the footwear condition. A few individuals (e.g., subject 10, 26), however, could not be identified across all footwear conditions. Generally, matching movement patterns to individuals (across all footwear conditions) was most precise when individuals ran in their own shoes (accuracy: 96.3 %) and least accurate when individuals ran barefoot (accuracy: 68.7 %).

Regardless of the subject, running barefoot was identified most accurately with 86.7 %. The non-standardised footwear condition (i.e., subject's own shoes) was recognised with the lowest accuracy (58.0 %). Interestingly, the accuracy at which footwear conditions were correctly identified varied greatly across individuals (38.0 - 100.0 %).

As Figure 4 shows, the prediction accuracy of the SVM models reached a saturation when using data from at least 10 subjects (median accuracy: 76.0 %), which could be slightly increased by adding the data from further subjects (median accuracy for 22 subjects: 84.0 %).

Relevance score evaluation

Across all classification tasks, aggregated relevance scores of the vertical ground reaction force (GRF_V) trajectory were lowest (Figure 5). Aggregated relevance scores of the medio-lateral (GRF_{ML}) and anterior-posterior (GRF_{AP}) ground reaction force trajectories were comparable in all but one classification tasks: within the footwear classification task, aggregated relevance scores in GRF_{AP} were substantially higher than aggregated scores in GRF_{ML} (Figure 5C).

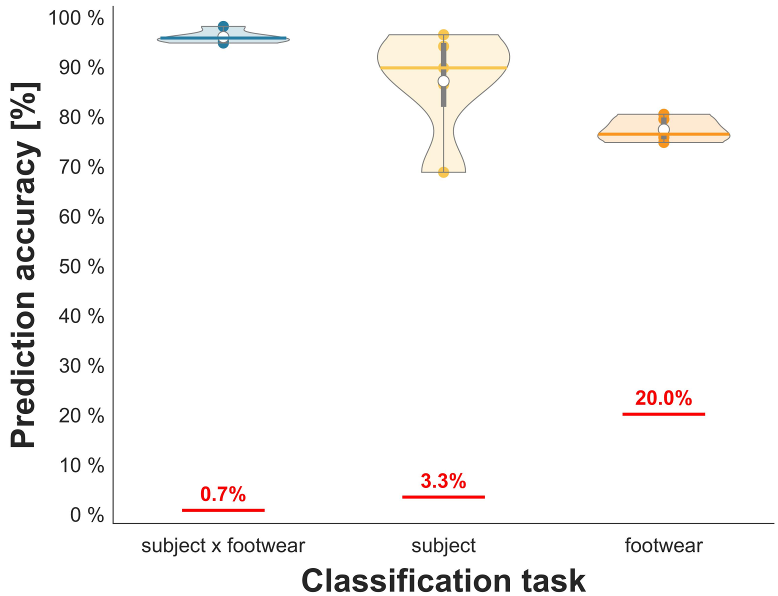

Figure 1. Performance evaluation across the five-fold cross-validation shown as violin plots with median (solid line), mean (white dot), and prediction accuracy of the individual folds (coloured dots). Support vector machine models were trained for the three employed classification tasks: subject x footwear (in blue on the left), subject (in yellow in the middle), and footwear (in orange on the right). The task-specific zero-rule baseline values (i.e., 0.7 % (subject x footwear classification), 3.3 % (subject classification), and 20.0 % (footwear classification)) are shown as a red line. Created using the MATLAB code provided by Bechtold et al. (2022).[34]

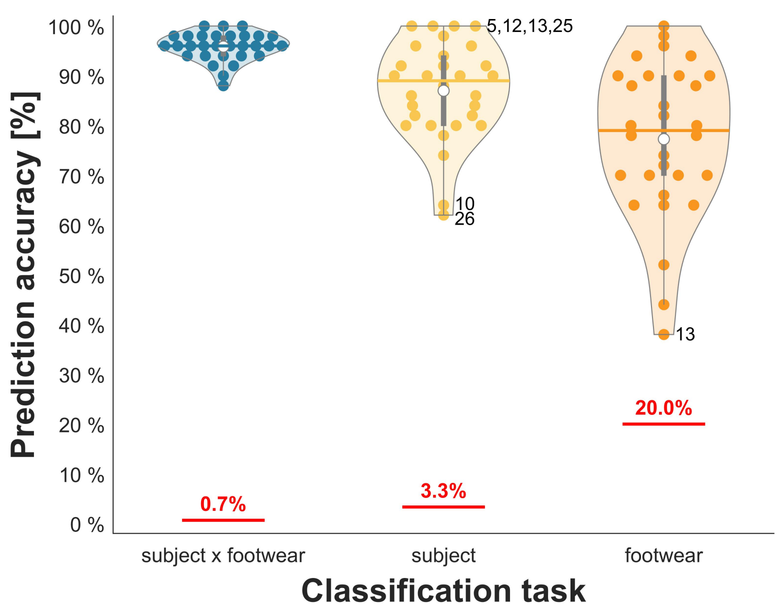

Figure 2. Performance evaluation across the subjects shown as violin plots with median (solid line), mean (white dot), and prediction accuracy of the individual subjects (coloured dots). The numbers in the violin plots represent selected subjects that are discussed in sections Performance evaluation, Subject classification (across footwear conditions), and Footwear classification (across subjects). Support vector machine models were trained for the three employed classification tasks: subject x footwear (in blue on the left), subject (in yellow in the middle), and footwear (in orange on the right). The task-specific zero-rule baseline values (i.e., 0.7 % (subject x footwear classification), 3.3 % (subject classification), and 20.0 % (footwear classification)) are shown as a red line. Created using the MATLAB code provided by Bechtold et al. (2022).[34]

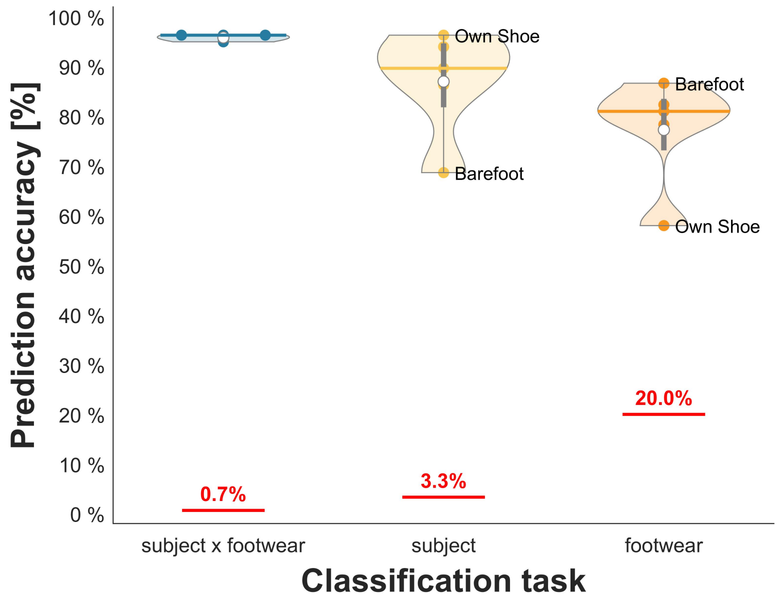

Figure 3. Performance evaluation across the footwear conditions shown as violin plots with median (solid line), mean (white dot), and prediction accuracy of the individual footwear conditions (coloured dots). Support vector machine models were trained for the three employed classification tasks: subject x footwear (in blue on the left), subject (in yellow in the middle), and footwear (in orange on the right). The task-specific zero-rule baseline values (i.e., 0.7 % (subject x footwear classification), 3.3 % (subject classification), and 20.0 % (footwear classification)) are shown as a red line. Created using the MATLAB code provided by Bechtold et al. (2022).[34]

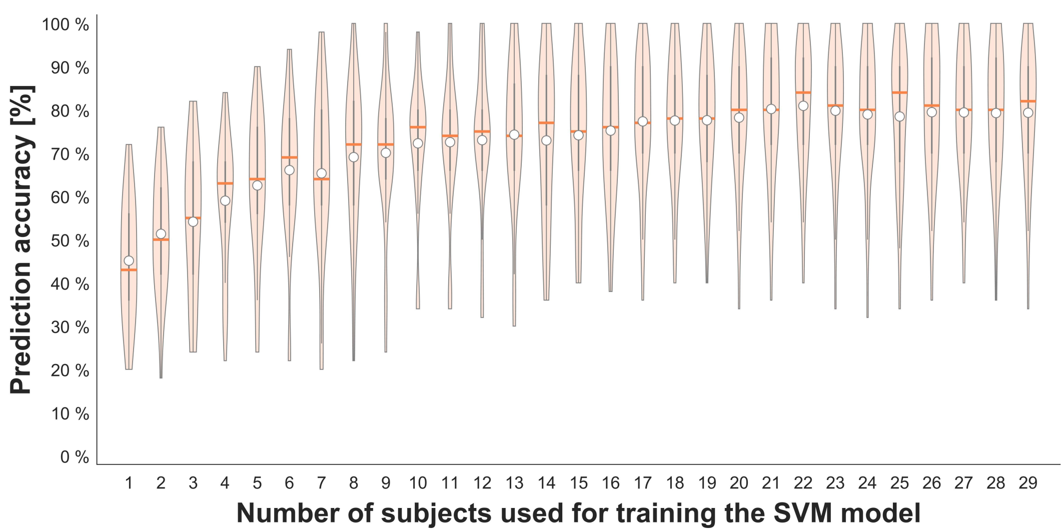

Figure 4. Performance evaluation of the machine learning models trained for footwear classification with different number of subjects used for training the support vector machine (SVM) models. The prediction accuracy was obtained using a leave-subject-out cross-validation configuration shown as violin plots with median (solid line) and mean (white dot). Created using the MATLAB code provided by Bechtold et al. (2022).[34]

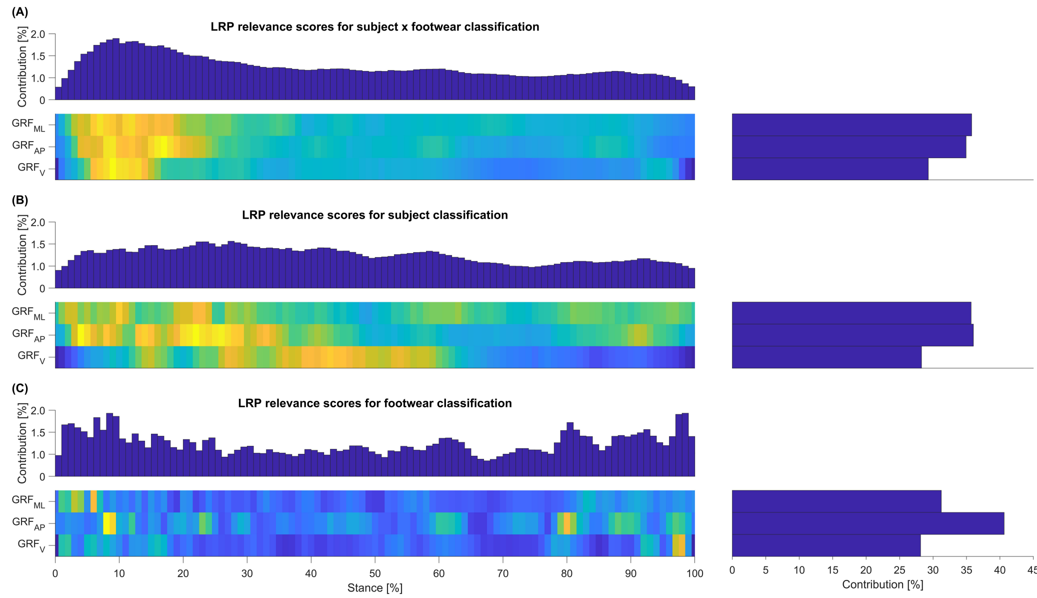

Figure 5. Input relevance evaluation of the machine learning models, created using the MATLAB code provided by Hoitz et al. (2021).[17] Input relevance scores obtained by Layer-wise Relevance Propagation (LRP) for the employed classification tasks: **(A)** subject x footwear, **(B)** subject, **(C)** footwear. For each subfigure **(A-C)**: The top part on the left shows the summed contribution of the relevance scores for each of the 101 time points of the stance phase. In the bottom part on the left, lighter colours indicate variables of high relevance, while darker colours indicate variables of low relevance. The bottom right part highlights the summed contribution of relevance scores of each of the ground reaction forces (GRFs), namely medio-lateral (GRF_{ML}), anterior-posterior (GRF_{AP}), and vertical (GRF_{V}).

From a temporal perspective, GRFs during the first 5 - 20 % of the stance phase were particularly relevant to the predictions of SVM models in subject x footwear classification task (Figure 5A). For the subject classification task (Figure 5B), regions with high summed relevance scores were more spread throughout the stance phase. Regions of high relevance were observed at 0 - 40 % (GRF_{AP}) and at 25 - 60 % (GRF_{V}) of stance. In the footwear classification task, input values with high relevance scores were spread across the entirety of stance with peak scores in the range of 0 - 10 % and 80 - 100 % (Figure 5C).

Discussion

Movement patterns of thirty subjects that ran overground in four shod conditions and one barefoot condition were analysed using support vector machines and layer-wise relevance propagation. Specifically, three-dimensional ground reaction forces were classified across three tasks: subject x footwear, subject, and footwear. The results for the different classification tasks will be discussed separately as they provide unique and novel insights to the field of running biomechanics.

Subject x footwear classification

When movement patterns were classified according to their respective subject *and* footwear condition, the accuracy of the support vector machine models were almost perfect (median accuracy: 95.7 %). Hypothesis (I), therefore, was supported by the findings of this work. This outcome suggests that each combination of runner and footwear results in a unique movement response that is identified by machine learning-based analysis. In other words, for each possible combination of runner and footwear, there is a high probability that unique and recognisable movement pattern exist. A notion supported by previous studies that reported variation in responses to footwear between subjects.[7,14] In combination with the works cited above, the presented results suggest that the effect of footwear on biomechanical variables should be studied by considering subject-specific responses more strongly, as was suggested previously by Bates et al. (1983).[14]

Subject classification (across footwear conditions)

When movement patterns were classified according to their respective subject, the median accuracy of the SVM models was approximately 90 %, supporting the second hypothesis of this work. This outcome suggests that subjects expressed individual movement characteristics, *regardless* of the tested footwear condition. An interpretation that is corroborated by previous findings that highlighted unique movement patterns for each tested individual in a barefoot[15,16,24] or a single non-standardised shod condition.[17,18] These distinct differences in movement patterns are likely the result of unique anatomical characteristics,[11,12] muscle activation strategies,[25,35] and prevalence of previous injuries.[13] Because participants were recognised *regardless* of the footwear condition, one may speculate that subject-specific movement characteristics changed only minimally across the different running conditions. This is relevant to the paradigm of the preferred movement path that suggests that runners maintain a consistent movement pattern when changing between similar footwear conditions.[2,19,22] In fact, runners were recognised with an accuracy of 96.3 % when running in their own shoes, supporting the idea of a consistent movement pattern (Figure 3). Interestingly, this accuracy dropped to 68.7 % when runners switched to the barefoot condition, suggesting that the subject's barefoot movement patterns were not as easily identified based on the shod movement patterns. This suggests that shod running is more fundamentally different from barefoot running for some individuals. A finding supported by previous studies.[19] The degree to which a movement pattern can be maintained (or cannot be maintained) is runner specific (Figure 2). Some subjects appeared to have an easier time maintaining a similar movement pattern across the footwear conditions than others. This is evident by the high classification accuracies of these runners (e.g., subjects 5, 12, 13, 25) compared to runners with low classification accuracies (e.g., subjects 10, 26). The ability of a runner to maintain a movement signature in different running conditions may influence the runner's ability to perform well (i.e., reduced energy consumption) with a given footwear and / or reduce risk of injury associated with changes in footwear. In any case, a reliable characterisation of an individual's movement signature (i.e., unique movement characteristics) must encompass a broad range of footwear / running conditions to account for the changes in movement patterns induced by distinctly different running conditions.

Footwear classification (across subjects)

When movement patterns were classified according to their respective footwear condition, the support vector machine model's median accuracy was 76.3 %. This outcome suggests that certain footwear-induced changes in movement patterns are consistent *across* subjects. However, the model's accuracy varied greatly across subjects (Figure 2), suggesting that some individuals *did not* exhibit consistent footwear-induced changes in movement patterns. In other words, when individuals react in a consistent manner to a footwear intervention and movement patterns across individuals are similar, the support vector machine model can predict the correct footwear condition with a high degree of accuracy. When individuals react differently to footwear interventions and the movement patterns of individuals are not similar, however, the model's accuracy drops drastically. Hypothesis (III), therefore, was partly supported by the findings of this work. Consequently, for some subjects, responses to a given footwear condition are comparable. These subjects may have similar functional needs towards footwear designs.[2,22,23] Given the finding of unique subject x footwear responses (in section Subject x footwear classification), future research needs to address whether there are subjects whose responses to "all" footwear conditions are comparable.

Limitations and implications for footwear research

One major limiting factor is the fact that the GRFs are integral variables that summarise accelerations of the centre of mass of all body segments and does not distinguish properly between specific influences. Future research involving a combination of bilateral kinematic, kinetic, and electromyographic data is needed to relate the presented results to a functional perspective. We consider the use of GRF data as an example to demonstrate how machine learning models (i.e., Support Vector Machines), together with Explainable Artificial Intelligence methods (i.e., Layer-wise Relevance Propagation), can enrich the evaluation of footwear effects.

The identification of unique movement patterns for each combination of runner and footwear provides evidence for subject-specific responses to footwear (in section Subject x footwear classification). Current research strategies often focus on estimates of central tendencies from groups, such as mean values, which inevitably blur subject-specific responses to footwear. Just because subject-specific responses to footwear occur in opposing directions (i.e., increase in one runner, decrease or no change in another), measures of central tendencies may capture neither response adequately.[20] Furthermore, it may be one of the reasons why endeavours frequently report contradictory results regarding footwear design features and their effects on biomechanics and / or injury risk.[4] Time-discrete ground reaction force variables (e.g., impact peaks, loading rates), which are frequently used to report injury risks,[36–38] appear to be particularly vulnerable to subject-specific responses to footwear (Figure 5A). It is essential, then, to consider subject-specific responses to gain an improved functional understanding on footwear effects during running and the potential relationship to running injury risks, as has been suggested previously.[14]

The finding that the movement patterns of most subjects have unique features (in section Subject classification (across footwear conditions)) regardless of footwear (including barefoot running) supports research strategies that assess the effects of footwear at the individual level, i.e., single-subject approaches.[21] It seems promising to integrate machine learning methods into these approaches, as the features that are unique to an individual (regardless of footwear) appear to vary widely between individuals (Figure 5B). Subject-specific machine learning models could therefore be used to predict in which footwear design the movement patterns are most similar to a desired reference movement pattern (e.g., preferred movement path,[3,19,39] or habitual movement path[40,41]). Machine learning models have the advantage that they can be trained on any number of multi-dimensional and time-continuous (biomechanical) variables and that no thresholds for low / high deviations need to be set for their predictions.

Another approach to better take into account the fact that individuals respond differently to footwear is to look for individuals who respond similarly to footwear (i.e., functional group approaches). On the one hand, machine learning approaches can be used to identify groups of individuals with similar responses to footwear,[23] and the effect of footwear can be evaluated by considering these groups (as is often done when distinguishing between forefoot and rearfoot runners in studies[42,43]). On the other hand, machine learning approaches can be used to learn different strategies for classifying footwear effects from individual subjects without explicit prior knowledge.[44] That is, machine learning models can consider various responses to footwear at the same time in their predictions. This allows to map the effect of footwear on running patterns in a more versatile and differentiated way than group-based statistical comparisons based on single time-discrete variables. However, the data from a crucial minimum number of subjects seem to be required to train machine learning models that can represent a wide range of individuals (Figure 4). When the number of subjects is relatively small, as often the case in biomechanical studies to date, there is a risk that the running patterns of individual subjects cannot be mapped well (e.g., subject 13 in our study). A representative database containing a large number of subjects could, however, not only allow to consider subgroups, but even allow the consideration on an individual level using the most similar subject in the database in the sense of a “digital twin”. This could have a great potential especially for the prediction of long-term effects of footwear for individual subjects. Given that predictions of machine learning models for footwear classification are characterised by a plethora of versatile features (Figure 5C), our findings imply that considering multi-dimensional and time-continuous biomechanical data (e.g., full-body kinematics and kinetics) appear to be promising for (subject-specific) evaluations of footwear effects.

Conclusion

The present findings suggest that unique movement signatures (across footwear) and unique responses to each footwear design can be modelled for each individual subject. Our results support the idea that considering subject-specific responses is advantageous for a better understanding of the functional effects of footwear during running. The incorporation of different multi-dimensional and time-continuous biomechanical data (e.g., whole body kinematics and kinetics) seems to be similar auspicious for a more differentiated (subject-specific) evaluation of the effects of footwear. Machine learning methods seems to be a promising and valuable extension to previous (subject-specific) approaches for footwear evaluation and recommendation.

References

[1] Nigg, B. M., Stefanyshyn, D., Cole, G., & Boyer, K. Footwear research - past, present and future. In: *Proceedings: 7th Symposium on Footwear Biomechanics*. Ed. by Hamill, J., Hardin, E., & Williams, K. Cleveland, Ohio, USA: Case Western Reserve University Printing, 2005: pp. 14–17.

[2] Nigg, B. M. *Biomechanics of Sport Shoes*. Calgary, Alberta, Canada: Topline Printing, 2010.

[3] Nigg, B. M., Baltich, J., Hoerzer, S., & Enders, H. 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) (2015): 1290–1294. DOI: 10.1136/bjsports-2015-095054.

[4] Hoitz, F., Mohr, M., Asmussen, M., Lam, W.-K., Nigg, S., & Nigg, B. The effects of systematically altered footwear features on biomechanics, injury, performance, and preference in runners of different skill level: a systematic review. *Footwear Science* **12**(3) (2020): 193–215. DOI: 10.1080/19424280.2020.1773936.

[5] Sterzing, T., Schweiger, V., Ding, R., Cheung, J. T.-M., & Brauner, T. Influence of rearfoot and forefoot midsole hardness on biomechanical and perception variables during heel-toe running. *Footwear Science* **5**(2) (2013): 71–79. DOI: 10.1080/19424280.2012.757810.

[6] Sterzing, T., Custoza, G., Ding, R., & Cheung, J. T.-M. Segmented midsole hardness in the midfoot to forefoot region of running shoes alters subjective perception and biomechanics during heel-toe running revealing potential to enhance footwear. *Footwear Science* **7**(2) (2015): 63–79. DOI: 10.1080/19424280.2015.1008589.

[7] Cole, G. K., Nigg, B. M., Fick, G. H., & Morlock, M. M. Internal loading of the foot and ankle during impact in running. *Journal of Applied Biomechanics* **11**(1) (1995): 25–46. DOI: 10.1123/jab.11.1.25.

[8] Oriwol, D., Sterzing, T., & Milani, T. L. The position of medial dual density midsole elements in running shoes does not influence biomechanical variables. *Footwear Science* **3**(2) (2011): 107–116. DOI: 10.1080/19424280.2011.613857.

[9] Baltich, J., Maurer, C., & Nigg, B. M. Increased vertical impact forces and altered running mechanics with softer midsole shoes. *PloS one* **10**(4) (2015): e0125196. DOI: 10.1371/journal.pone.0125196.

[10] Schöllhorn, W. I., Nigg, B. M., Stefanyshyn, D. J., & Liu, W. Identification of individual walking patterns using time discrete and time continuous data sets. *Gait & Posture* **15**(2) (2002): 180–186. DOI: 10.1016/s0966-6362(01)00193-x.

[11] Uhl, A. & Wild, P. Personal identification using eigenfeet, ballprint and foot geometry biometrics. In: *2007 First IEEE International Conference on Biometrics: Theory, Applications, and Systems*. IEEE, 2007: pp. 1–6. ISBN: 978-1-4244-1596-0. DOI: 10.1109/BTAS.2007.4401924.

[12] Kanai, R. & Rees, G. The structural basis of inter-individual differences in human behaviour and cognition. *Nature Reviews Neuroscience* **12**(4) (2011): 231–242. DOI: 10.1038/nrn3000.

[13] van Mechelen, W. Running injuries. A review of the epidemiological literature. *Sports Medicine* **14**(5) (1992): 320–335. DOI: 10.2165/00007256-199214050-00004.

[14] Bates, B. T., Osternig, L. R., Sawhill, J. A., & James, S. L. An assessment of subject variability, subject-shoe interaction, and the evaluation of running shoes using ground reaction force data. *Journal of Biomechanics* **16**(3) (1983): 181–191. DOI: 10.1016/0021-9290(83)90125-2.

[15] Pataky, T. C., Mu, T., Bosch, K., Rosenbaum, D., & Goulermas, J. Y. Gait recognition: highly unique dynamic plantar pressure patterns among 104 individuals. *Journal of the Royal Society, Interface* **9**(69) (2012): 790–800. DOI: 10.1098/rsif.2011.0430.

[16] Horst, F., Mildner, M., & Schöllhorn, W. I. One-year persistence of individual gait patterns identified in a follow-up study - a call for individualised diagnose and therapy. *Gait & Posture* **58** (2017): 476–480. DOI: 10.1016/j.gaitpost.2017.09.003.

[17] Hoitz, F., Tscharner, V. von, Baltich, J., & Nigg, B. M. Individuality decoded by running patterns: movement characteristics that determine the uniqueness of human running. *PloS one* **16**(4) (2021): e0249657. DOI: 10.1371/journal.pone.0249657.

[18] Hoitz, F., Fraeulin, L., Tscharner, V. von, Ohlendorf, D., Nigg, B. M., & Maurer-Grubinger, C. Isolating the unique and generic movement characteristics of highly trained runners. *Sensors* **21**(21) (2021). DOI: 10.3390/s21217145.

[19] Nigg, B. M., Vienneau, J., Smith, A. C., Trudeau, M. B., Mohr, M., & Nigg, S. R. The preferred movement path paradigm: influence of running shoes on joint movement. *Medicine and Science in Sports and Exercise* **49**(8) (2017): 1641–1648. DOI: 10.1249/MSS.0000000000001260.

[20] Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences of the United States of America* **115**(27) (2018). DOI: 10.1073/pnas.1711978115.

[21] Bates, B. T. Single-subject methodology: an alternative approach. *Medicine and Science in Sports and Exercise* **28**(5) (1996): 631–638. DOI: 10.1097/00005768-199605000-00016.

[22] Nigg, B. M. The role of impact forces and foot pronation: a new paradigm. *Clinical Journal of Sport Medicine* **11**(1) (2001): 2–9. DOI: 10.1097/00042752-200101000-00002.

[23] Hoerzer, S., Tscharner, V. von, Jacob, C., & Nigg, B. M. Defining functional groups based on running kinematics using self-organizing maps and support vector machines. *Journal of Biomechanics* **48**(10) (2015): 2072–2079. DOI: 10.1016/j.jbiomech.2015.03.017.

[24] Horst, F., Lapuschkin, S., Samek, W., Müller, K.-R., & Schöllhorn, W. I. Explaining the unique nature of individual gait patterns with deep learning. *Scientific Reports* **9**(1) (2019): 2391. DOI: 10.1038/s41598-019-38748-8.

[25] Aeles, J., Horst, F., Lapuschkin, S., Lacourpaille, L., & Hug, F. Revealing the unique features of each individual's muscle activation signatures. *Journal of the Royal Society, Interface* **18**(174) (2021): 20200770. DOI: 10.1098/rsif.2020.0770.

[26] Schöllhorn, W. I. & Bauer, H. U. Erkennung von individuellen Laufmustern mit Hilfe von neuronalen Netzen. In: *Informatik im Sport*. Ed. by Mester, J. & Perl, J. Cologne, Germany: Sport Buch Strauss, 1998: pp. 169–176.

[27] Schöllhorn, W. I. Applications of artificial neural nets in clinical biomechanics. *Clinical Biomechanics* **19**(9) (2004): 876–898. DOI: 10.1016/j.clinbiomech.2004.04.005.

[28] Prakash, C., Kumar, R., & Mittal, N. Recent developments in human gait research: parameters, approaches, applications, machine learning techniques, datasets and challenges. *Artificial Intelligence Review* **49**(1) (2018): 1–40. DOI: 10.1007/s10462-016-9514-6.

[29] Phinyomark, A., Petri, G., Ibáñez-Marcelo, E., Osis, S. T., & Ferber, R. Analysis of big data in gait biomechanics: current trends and future directions. *Journal of Medical and Biological Engineering* **38**(2) (2018): 244–260. DOI: 10.1007/s40846-017-0297-2.

[30] Halilaj, E., Rajagopal, A., Fiterau, M., Hicks, J. L., Hastie, T. J., & Delp, S. L. Machine learning in human movement biomechanics: best practices, common pitfalls, and new opportunities. *Journal of Biomechanics* **81** (2018): 1–11. DOI: 10.1016/j.jbiomech.2018.09.009.

[31] Cortes, C. & Vapnik, V. Support-vector networks. *Machine Learning* **20**(3) (1995): 273–297. DOI: 10.1007/BF00994018.

[32] Fan, R. E., Chang, K.-W., Hsieh, C.-J., Wang, X.-R., & Lin, C.-J. Liblinear: a library for large linear classification. *Journal of Machine Learning Research* **9** (2008): 1871–1874.

[33] Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K.-R., & Samek, W. On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *PloS one* **10**(7) (2015): e0130140. DOI: 10.1371/journal.pone.0130140.

[34] Bechtold, B., Fletcher, P., Holden, S., & Gorur-Shandilya, S. *Violin plots for Matlab*. 2022. URL: https://zenodo.org/record/4559847 (visited on Oct. 31, 2022).

[35] Hug, F., Vogel, C., Tucker, K., Dorel, S., Deschamps, T., Le Carpentier, É., & Lacourpaille, L. Individuals have unique muscle activation signatures as revealed during gait and pedaling. *Journal of Applied Physiology* **127**(4) (2019): 1165–1174. DOI: 10.1152/japplphysiol.01101.2018.

[36] van der Worp, H., Vrielink, J. W., & Bredeweg, S. W. Do runners who suffer injuries have higher vertical ground reaction forces than those who remain injury-free? A systematic review and meta-analysis. *British Journal of Sports Medicine* **50**(8) (2016): 450–457. DOI: 10.1136/bjsports-2015-094924.

[37] Johnson, C. D., Tenforde, A. S., Outerleys, J., Reilly, J., & Davis, I. S. Impact-related ground reaction forces are more strongly associated with some running injuries than others. *The American Journal of Sports Medicine* **48**(12) (2020): 3072–3080. DOI: 10.1177/0363546520950731.

[38] Schmida, E. A., Wille, C. M., Stiffler-Joachim, M. R., Kliethermes, S. A., & Heiderscheit, B. C. Vertical loading rate is not associated with running injury, regardless of calculation method. *Medicine and Science in Sports and Exercise* **54**(8) (2022): 1382–1388. DOI: 10.1249/MSS.0000000000002917.

[39] Nigg, B. M., Mohr, M., & Nigg, S. R. Muscle tuning and preferred movement path - a paradigm shift. *Current Issues in Sport Science* (2) (2017): 007. DOI: 10.36950/CISS{\textunderscore}2017.007.

[40] Trudeau, M. B., Willwacher, S., Weir, G., Rohr, E., Ertel, C., Brüggemann, G.-P., & Hamill, J. A novel method for estimating an individual's deviation from their habitual motion path when running. *Footwear Science* **11**(3) (2019): 135–145. DOI: 10.1080/19424280.2019.1615004.

[41] Willwacher, S., Mählich, D., Trudeau, M. B., Hamill, J., Weir, G., Brüggemann, G.-P., & Bratke, G. The habitual motion path theory: evidence from cartilage volume reductions in the knee joint after 75 minutes of running. *Scientific Reports* **10**(1) (2020): 1363. DOI: 10.1038/s41598-020-58352-5.

[42] Ahn, A. N., Brayton, C., Bhatia, T., & Martin, P. Muscle activity and kinematics of forefoot and rearfoot strike runners. *Journal of Sport and Health Science* **3**(2) (2014): 102–112. DOI: 10.1016/j.jshs.2014.03.007.

[43] Gruber, A. H., Boyer, K. A., Derrick, T. R., & Hamill, J. Impact shock frequency components and attenuation in rearfoot and forefoot running. *Journal of Sport and Health Science* **3**(2) (2014): 113–121. DOI: 10.1016/j.jshs.2014.03.004.

[44] Slijepcevic, D., Horst, F., Lapuschkin, S., Horsak, B., Raberger, A.-M., Kranzl, A., Samek, W., Breiteneder, C., Schöllhorn, W. I., & Zeppelzauer, M. Explaining machine learning models for clinical gait analysis. *ACM Transactions on Computing for Healthcare* **3**(2) (2022): 1–27. DOI: 10.1145/3474121.