This is a repository copy of *Predicting longitudinal changes in joint contact forces in a juvenile population: Scaled generic versus subject-specific musculoskeletal models*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/162382/

Version: Published Version

Article:

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Predicting longitudinal changes in joint contact forces in a juvenile population: scaled generic versus subject-specific musculoskeletal models

Claude Fiifi Hayford, Erica Montefiori, Emma Pratt & Claudia Mazzà

Published online: 26 Jun 2020.

Submit your article to this journal ⬀

View related articles ⬀

View Crossmark data ⬀

Taylor & Francis
Taylor & Francis Group

Predicting longitudinal changes in joint contact forces in a juvenile population: scaled generic versus subject-specific musculoskeletal models

Claude Fiifi Hayford[a], Erica Montefiori[a], Emma Pratt[b] and Claudia Mazzà[a]

[a]Department of Mechanical Engineering and Insigneo Institute for in Silico Medicine, University of Sheffield, Sheffield, UK; [b]Gait Laboratory, Ryegate Children's Centre, Sheffield Children's Hospital NHS Foundation Trust, Sheffield, UK

ABSTRACT

Subject-specific musculoskeletal model use in clinical settings is limited due to development-associated time and effort burdens together with potential medical imaging unavailability. As an alternative, this study investigated consistency in estimating longitudinal changes in joint contact forces (JCF) between scaled generic and subject-specific models. For 11 children, joint kinematics and JCF were calculated using subject-specific and scaled generic models. JCF changes estimated by both models were strongly correlated for the hip and knee although JCF estimates varied between models. Findings suggest that within specified limits of accuracy, scaled generic models are sensitive enough to detect JCF changes consistent with subject-specific models.

Introduction

Three-dimensional gait analysis (3DGA), based on optoelectronic and force platform data, has become a mainstay in the study of human movement musculoskeletal disorders, providing useful information to guide treatment planning and rehabilitation (Baker et al. 2016; Wesseling et al. 2017). Most recent literature has shown that the utility of conventional 3DGA can be further augmented with musculoskeletal models (MSK)(Kainz et al. 2019; Montefiori et al. 2019b). These are mathematical representations of the body as a system of rigid bodies linked in a chain by joints and constraints and actuated by muscle forces. This formulation lends itself to rigid multibody dynamics and simulation that provides information such as estimates of changes in muscle length, muscle force and joint contact force that are not available using conventional 3DGA or would require the use of some instrumented prosthesis.

Most commonly used MSK models, typically referred to as scaled generic models, are based on data extrapolated from cadaveric specimens of healthy adults (Delp et al. 1990; Arnold et al. 2010; Modenese et al. 2011) which are scaled based on markers or anthropometry to match a subject. This poses a challenge when dealing with different populations, such as children and those with pathologic conditions (Duda et al. 1996; Bosmans et al. 2015, 2016). Imaging modalities such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) have been used to address this challenge by allowing for the increase in personalisation of these models through the inclusion of subject-specific details like bone geometry (Lenaerts et al. 2009), muscle paths and attachment (Scheys et al. 2008; Bosmans et al. 2015), as well as estimates of musculotendon parameters (Correa and Pandy 2011; Hainisch et al. 2012). This personalisation has been proven to increase the accuracy and reliability of these MSK models (Blemker et al. 2007; Lenaerts et al. 2009). Nonetheless, subject-specific models created in this way have cost and time burdens which limit their use in clinical settings (Valente et al. 2014). In addition, medical imaging may not be feasible or available, especially when conducting retrospective studies. In such instances, if sensitive enough to detect changes that are bigger than their expected limits of accuracy, generic models might represent a relatively easily implementable substitute.

A number of studies have compared the performance between generic models and image-based subject-specific models and concur that differences exist between the biomechanical measures estimated. When investigating a normal and a pathologic gait condition,

Scheys et al. (2008) found that differences existed between the moment arms and muscle-lengths estimated by the two types of models for 16 major muscles of the lower limb for each gait condition. For most of the muscles however, the changes in muscle length and moment arm estimates was found to be similar for the two approaches, for both normal and pathologic gait. Similarly, Correa et al. (2011) found significant differences in muscle moment arms when comparing generic and subject-specific models but also reported that both models were consistent in their predictions of muscle action. Muscle forces contribute to the magnitude of joint contact forces (JCFs) and differences are therefore expected to be observed in JCF estimates between the two models. However, it remains unclear to what extent these differences impact on longitudinal estimates of changes in the biomechanical variables predicted by generic models and how they differ from those predicted from subject-specific models. The aim of this study was to examine the suitability of using scaled generic models to predict longitudinal changes in biomechanical measures and how these predictions differ from those obtained from subject-specific models in a juvenile population. We hypothesised that despite differences in instantaneous estimates of JCFs, there would be no difference in the change in calculated JCF over time between models. If this hypothesis holds true, scaled generic models could be used to infer clinically meaningful information where interest is in change over time as opposed to absolute estimates such as in predicting or evaluating surgical outcomes, hence suggesting the feasibility of using them as alternative to the more accurate subject-specific models.

Materials and methods

Data collection

Data from 11 participants (age at initial observation: mean 11.5 (SD 3.2) years) were extracted from a dataset collected during the MD-PAEDIGREE project, which aimed at investigating disease progression in children with Juvenile Idiopathic Arthritis (JIA)(Montefiori et al. 2019b). Subject anthropometry was recorded at an initial observation (mass: 46.5 ± 18.0 kg, height: 1.4 ± 0.2 m) and at twelve months follow-up (mass: 51.4 ± 20.5 kg, height: 1.5 ± 0.2 m). Approval was obtained from the research ethics committees of the hospitals from which the data was collected. Gait data were collected over two observations (M0 and M12, 12 months apart) and across two laboratories, one using a 6-camera setup (BTS, SmartDX, 100 Hz) with two force plates (Kistler, 1 kHz) and the other, an 8-camera

system (Vicon, MX, 200 Hz) with two force plates (AMTI, OR6, 1 kHz). The Vicon PlugIn gait protocol (Vicon Motion Systems) augmented with the modified Oxford Foot Model (Stebbins et al. 2006) formed the set of forty-four markers used. Regional MRI of the foot and ankle was acquired for each participant at the two observations (M0 and M12) using a multi-slice multi-echo 3 D Gradient Echo (mFFE) with water only selection (0.5 mm in plane resolution and 1 mm slice thickness). MRI was also used to acquire entire lower limb images at an intermediate timepoint (six months from initial observation) using a 3 D T1-weighted fat-suppression sequence. In-plane resolution was 1 mm with a slice thickness of 1 mm. These images were used to clinically evaluate bone erosion and cartilage damage (Montefiori et al. 2019a).

Modelling approaches

Subject-specific bone geometries for the two time-points were obtained by a single expert operator segmenting MRI images of the full lower-limb together with the regional foot and ankle images from each observation point, respectively. The full lower-limb geometries for each participant were subsequently coupled with the regional geometries to build subject-specific models (SubS) for each observation using NMSBuilder (Valente et al. 2017). For each SubS model, the hip was modelled as an ideal ball-and-socket joint, with ideal hinges for the knee, ankle and subtalar joints. The joint axes were defined by morphological fitting of articular surfaces isolated from the bone geometries, using a least square difference minimization approach. A supervised atlas registration procedure with a reference model (Delp et al. 1990) was used to estimate muscle attachments and via points, with manual adjustment against the MRI when needed. The maximum isometric force for each muscle in the SubS model were linearly scaled using the ratio of participant lower-limb mass, calculated as the product of the soft tissue volume and bone volume and their respective densities from the literature (White et al. 1987), and the lower limb mass of a generic model (Delp et al. 1990). Further details for generating the SubS are provided in Modenese et al. (2018) and Montefiori et al. (2019a).

The cadaver-based generic gait2392 model (Delp et al. 1990) formed the basis of the scaled generic models (Gen). The gait2392 model was scaled by each subject's mass and anthropometry based on experimental markers placed on anatomical landmarks and estimated joint centres using the Scale tool in

OpenSim 3.3 (Delp et al. 1990; 2007). Scaling was based on literature recommendations (Hicks et al. 2015; Kainz et al. 2017) using Harrington regression equation estimates of the hip joint centre (Harrington et al. 2007), midpoint of medial and lateral epicondyle and malleolus markers for the knee and ankle joint centres, respectively. Maximum isometric force of each muscle was scaled by the mass of the subject divided by the mass of the gait2392 model. Optimal fibre length was scaled to preserve the muscle-tendon length ratio in the gait2392 model. The Gen models consisted of a single lower limb model with 12 degrees of freedom (DoF) for consistency with the SubS which were unilateral.

Simulations were subsequently performed in OpenSim 3.3 using a minimum of three collected experimental gait trials for each participant. The OpenSim simulation pipeline included inverse kinematics, inverse dynamics, static optimisation and joint reaction analysis (Steele et al. 2012). For each model and observation, joint angles, joint moments, muscle forces and JCFs were obtained. In line with best practice, maximum root mean square tracking error and peak marker tracking error between experimental marker trajectories and virtual markers for each model were kept below the recommended 20 mm and 40 mm thresholds (Hicks et al. 2015), respectively for inverse kinematics. Static posture joint angles were considered as a zero reference in comparing kinematic outputs between the two models. Joint powers were calculated as the product of joint moment and angular velocity. The muscle force-length-velocity relationship was ignored for both models during the estimation of muscle activation and force during static optimisation. Simulated joint moments and JCFs were normalised by participant body weight (BW). The dynamic consistency of the simulations was assessed for both models as recommended (Steele et al. 2012; Hicks et al. 2015) with the values of residual moments applied at the pelvis which were less than 10% of the product of BW and height for all subjects. Residual forces were similarly less than 10% of BW with the exception of F_y. This was due to the lack of general actuators in the model to provide an equilibrating effect of force and moment on the pelvis in the absence of a torso and contralateral leg (Modenese and Phillips 2012).

Differences between Gen and SubS models

Group mean and standard deviation for each estimated variable were determined as the average of ensemble means of subject trials for all subjects for both models. Joint angles, joint moments, joint power, JCF and differences in JCF (ΔJCF) at the hip, knee and ankle were compared between models and observations using the nonparametric one sample paired t-tests from the *spm1d* statistical parametric mapping (SPM) package (Pataky 2012) in MATLAB (v9.5.0, R2018b, MathWorks, USA). Significance was evaluated at $\alpha < 0.05$.

For the JCFs, total waveform variability or goodness of fit was assessed using the Root Mean Square Deviation (RMSD) (Picerno et al. 2008) between Gen and SubS for each subject at each observation. The percentage difference (%*Diff*) in JCF estimates was calculated as the ratio of RMSD to range of JCF predicted by the SubS model for each participant.

The coefficient of determination (R^2) was calculated using the Linear Fit Method to assess waveform shape similarity (Iosa et al. 2014) of JCF estimates over the gait cycle between Gen and SubS for each participant. This method can be used as a robust measure of curve similarity in the analysis of gait data (Di Marco et al. 2018). Peak values of JCF during the loading response (P1, indicated as occurring within the first 20% of the gait cycle) and push off (P2, indicated as occurring between 40 to 60% of the gait cycle) for both models were extracted for each participant and analysed. Area under the JCF curves (AUC) were also calculated as measures of overall loading of the joint throughout the gait cycle and compared between the two models. A graphical representation of the different indices used is presented in the supplementary materials.

Consistency in longitudinal predictions

Inter-observation differences (Δ_t) between the values of JCFs, AUC and JCF P1 and P2 were calculated for each participant's Gen and SubS model to assess their agreement in estimating longitudinal changes. Inter-model differences (Δ_m) for these metrics at each observation were similarly calculated for each participant. Gen and SubS were judged as in agreement in predicting longitudinal changes in JCFs if the coefficient of determination R^2 calculated between $\Delta_t(JCF_{Gen})$ and $\Delta_t(JCF_{SubS})$ was greater than or equal to 0.6.

It has been previously reported that SubS output is affected by repeatability errors associated to operator input (Montefiori et al. 2019b). In order to account for this when assessing the differences between the two models, ad-hoc thresholds were calculated for hip (H), knee (K) and ankle (A) joints using publicly
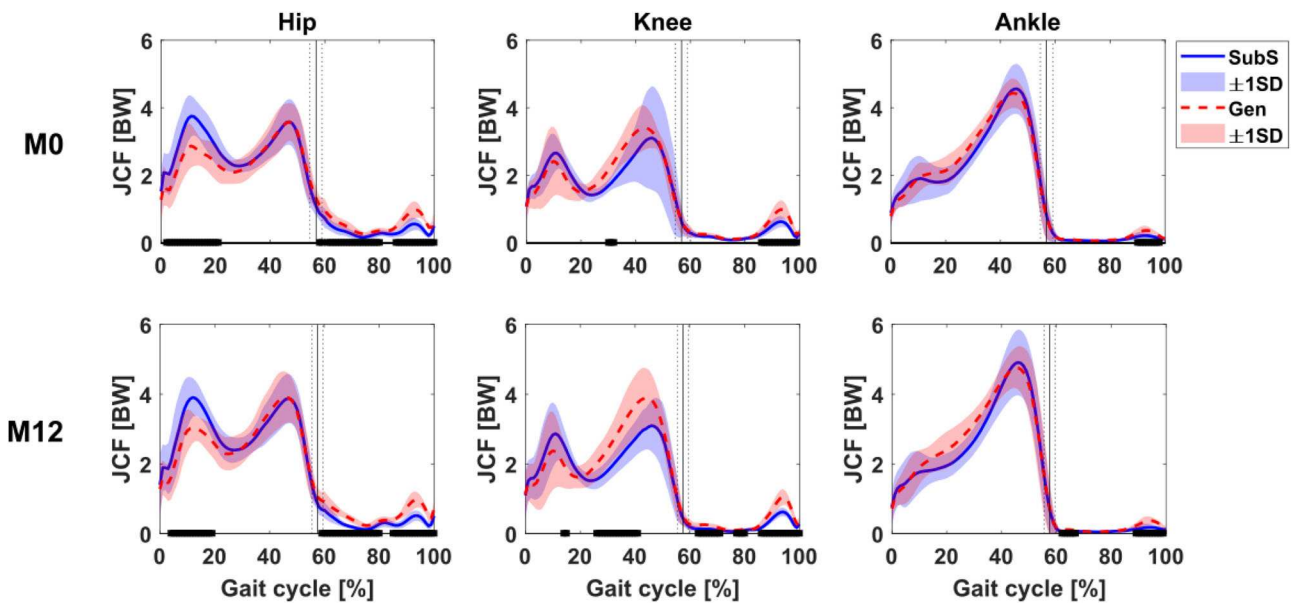
Figure 1. Comparison between Gen (red) and SubS (blue) model estimations of joint contact forces at observations M0 and M12. Black bars indicate region of gait cycle with significant statistical difference between the two models at P < 0.01.

available data (Montefiori et al. 2019b): $\Delta_m(AUC_H) = 29.17$, $\Delta_m(AUC_K) = 7.88$, $\Delta_m(AUC_A) = 4.42$, $\Delta_m(P1_H) = 0.45$ BW, $\Delta_m(P2_H) = 1.27$ BW, $\Delta_m(P1_K) = 0.36$ BW, $\Delta_m(P2_K) = 0.64$ BW and $\Delta_m(P1_A) = 0.94$ BW. If corresponding $\Delta_m(JCF$ peaks) and $\Delta_m(AUC)$ were lower than these thresholds at both observations, then the longitudinal output from the two models were considered to be in agreement for that participant. If the differences between the two models were bigger than these thresholds for at least one of the observations, then the differences (Δ_t) obtained for the two models were considered as being in agreement if consistent in signs.

Finally, significance of inter-model and inter-observation differences in estimates of RMSD, JCF peak values and area under JCF curve were assessed with the Wilcoxon signed-rank test in MATLAB. Effect size statistics for these estimates was also calculated using Cohen's d estimate with a pooled standard deviation from SubS and Gen. All statistical tests were conducted at $\alpha < 0.05$.

Results

Assessment of MRI images did not highlight any clinically meaningful changes in bone erosion and cartilage damage between the two time points.

Estimates of joint kinematics, moments and power in the sagittal plane at all observations and for all participants are presented as supplementary figures. Profile shapes of these estimates over the whole gait cycle were overall similar between the models,

although there existed significant ($P < 0.05$) differences at instances in the gait cycle. The JCFs estimated by the two models showed similar waveform profile (Figure 1).

The Gen tended to have lower estimates (average difference of 0.8 BW) of hip JCF during the loading response. This difference was found to be significant at both observed timepoints ($P < 0.001$ at both M0 and M12). JCF estimates were generally similar at the ankle with some significant differences reported during the stance phase of the gait cycle. The Gen also estimated higher JCF at the knee during push off (P2); this was however found to be not significant. The higher P2 prediction at the knee by the Gen was coincident with the prediction of a higher gastrocnemius medialis muscle force (see supplementary material) by the Gen in the same phase of the gait cycle.

Differences in JCF between the Gen and SubS were then analysed for each of the three joints (Table 1). Group median RMSD results ranged from 0.30 to 0.60. These values were similar for M0 and M12 with no statistical difference between RMSD values recorded at the two timepoints. The largest mean RMSD was recorded at the knee with a value of 0.59. The knee also had the largest variability between participants' RMSD as indicated by its standard deviation. Similar trends were observed when looking at %Diff, where the knee values doubled those at the ankle.

The coefficient of determination values (Table 1) were greater than 0.7 for all participants, with the highest correlations observed at the ankle (higher

Table 1. Inter-model analysis of gait waveform profile at M0 and M12.

	M0			M12		
	RMSD (IQR)	%Diff (SD)	R^2(range)	RMSD (IQR)	%Diff (SD)	R^2(range)
Hip	0.38 (0.17)	11 (3)	0.87 - 0.97	0.40 (0.18)	10 (2)	0.77 - 0.98
Knee	0.43 (0.27)	17 (6)	0.74 - 0.98	0.63 (0.41)	17 (7)	0.71 - 0.98
Ankle	0.32 (0.15)	8 (3)	0.93 - 0.99	0.29 (0.22)	8 (5)	0.95 - 0.99

Median RMSD, interquartile range (IQR) and range of JCF curve similarity (correlation) between Gen and SubS model estimates for the hip, knee and ankle joints of 11 participants. %Diff is the mean RMSD expressed as a percentage of the range of normalised JCF estimated by SubS.
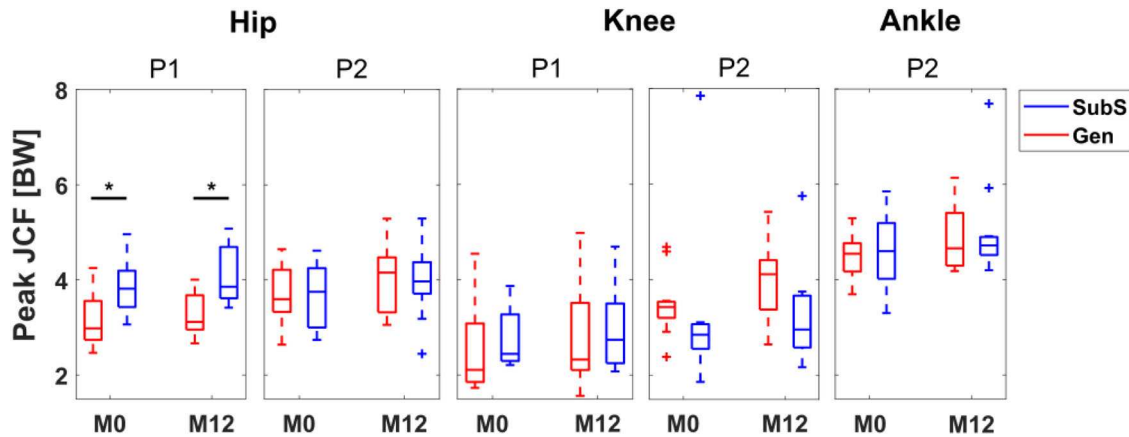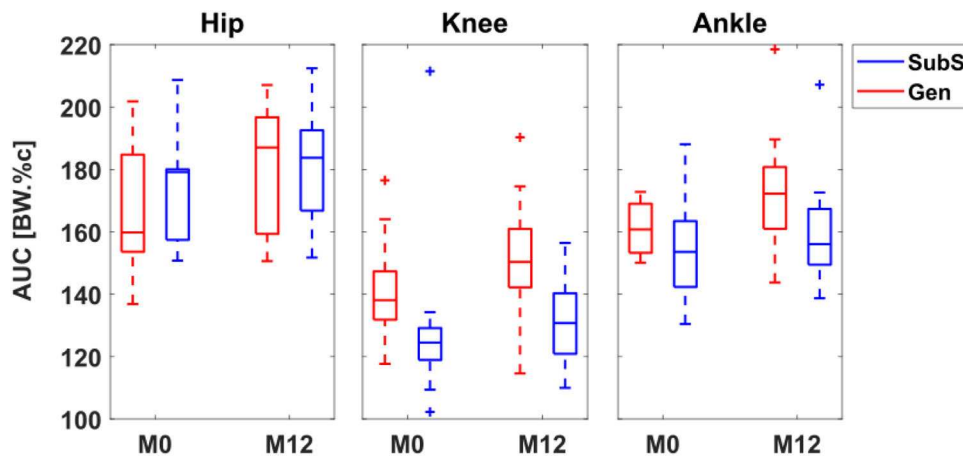

Figure 2. Boxplot distribution of P1 and P2 JCF estimates for scaled generic and subject-specific models at two observations, M0 and M12. * indicates significant difference at $P < 0.05$.

Figure 3. Boxplot distribution of overall joint loading estimates calculated as area under BW-normalised JCF curve. AUC expressed as BW*%Gait Cycle (BW.%c).

than 0.9) indicating a very strong linear relationship between JCF predictions between the Gen and SubS. Again, estimates at the knee showed the widest range of values. The comparison of selected peaks highlighted a large inter-subject variability in differences between the two models, however the hip loading response peak, P1, was consistently and significantly lower in the Gen than in the SubS (P(M0) = 0.006, P(M12) = 0.002, Figure 2). Further, Cohen's d values

(-1.31 and -1.48 for P1 and P2, respectively) suggested a high relevance of this difference.

Comparison of longitudinal differences in JCF, Δ_t(JCF) estimates for each model revealed that the Gen reported higher mean differences than the SubS for the hip and ankle joint although this did not reach significance. Peak Δ_t(JCF) was lower at the hip for the Gen compared to the SubS. Both models showed an overall increase going from M0 to M12 in

Table 2. Inter-model differences in participant estimates at M0.

| | $\|\Delta_m(AUC)\|[BW.\%c]$ | | | $\|\Delta_m(Peak)\|[BW]$ | | | | | $R^2[\Delta_m(JCF)]$ | | |
	H	K	A	$P1_H$	$P2_H$	$P1_K$	$P2_K$	$P2_A$	H	K	A
S1	9.1	**14.1**	**11.0**	0.4	0.3	**0.8**	0.6	0.5	**0.94**	0.75	**0.96**
S2	0.1	**13.6**	0.8	0.3	0.2	0.2	0.6	**0.9**	**0.93**	0.90	**0.93**
S3	3.2	**15.7**	**10.5**	**0.9**	0.1	**0.5**	0.7	0.6	**0.89**	0.85	**0.99**
S4	6.9	**39.5**	0.7	**1.2**	0.1	**0.5**	**1.6**	0.6	**0.92**	0.94	**0.98**
S5	7.7	**42.2**	**16.4**	0.4	0.6	**0.7**	**1.7**	0.3	**0.94**	0.94	**0.99**
S6	9.0	**22.4**	1.0	**0.8**	0.7	0.3	**1.2**	**0.9**	**0.87**	0.74	**0.93**
S7	21.5	**12.1**	**17.9**	**1.2**	0.5	0.4	**0.7**	0.4	**0.96**	0.88	**0.96**
S8	12.1	**20.1**	**13.3**	0.5	0.2	**0.8**	0.4	0.6	**0.92**	0.98	**0.98**
S9	6.2	**14.4**	**30.3**	0.1	0.5	0.3	0.3	0.6	**0.97**	0.92	**0.98**
S10	17.5	**15.4**	**29.9**	**1.0**	0.2	0.2	0.5	0.6	**0.95**	0.89	**0.99**
S11	**37.3**	**91.5**	0.3	**1.5**	**1.0**	**1.1**	**4.9**	0.0	**0.91**	0.86	**0.99**
n	10	0	4	4	11	5	5	10	11	11	11

Absolute values of inter-model differences (Δ_m) at M0. Values in bold indicate greater than applied thresholds. [n] is the number of participants for which Gen and SubS were considered to be in agreement based on Δ_m. AUC expressed as BW*%Gait Cycle (BW.%c).

Table 3. Inter-model differences in participant estimates at M12.

| | $\|\Delta_m(AUC)\|[BW.\%c]$ | | | $\|\Delta_m(Peak)\|[BW]$ | | | | | $R^2[\Delta_m(JCF)]$ | | |
	H	K	A	$P1_H$	$P2_H$	$P1_K$	$P2_K$	$P2_A$	H	K	A
S1	19.5	**41.6**	**75.1**	0.4	0.4	0.2	**1.8**	**1.3**	**0.91**	0.78	**0.95**
S2	9.5	**29.6**	**27.0**	0.0	0.4	0.3	**1.3**	0.1	**0.95**	0.91	**0.96**
S3	11.3	**25.2**	7.3	**0.6**	0.9	**0.6**	**1.6**	**0.9**	**0.90**	0.81	**0.98**
S4	4.8	**38.7**	**26.0**	**1.4**	0.5	0.1	**1.7**	**1.9**	**0.90**	0.95	**0.98**
S5	1.9	**34.3**	3.1	**1.1**	0.4	0.3	**1.4**	0.1	**0.92**	0.95	**0.99**
S6	1.1	**26.3**	9.9	**1.3**	0.6	**0.9**	**1.6**	0.6	**0.77**	0.71	**0.95**
S7	0.1	**16.2**	2.0	**0.9**	0.2	**0.7**	**1.1**	0.2	**0.95**	0.82	**0.97**
S8	12.4	**16.9**	**24.5**	0.3	0.5	0.2	0.4	0.7	**0.98**	0.98	**0.99**
S9	5.4	**10.0**	**18.7**	**0.6**	0.0	0.1	0.6	0.1	**0.98**	0.92	**0.98**
S10	9.8	**10.4**	5.4	**0.6**	0.5	0.4	0.5	0.5	**0.98**	0.88	**0.97**
S11	20.3	**41.9**	5.1	**1.5**	0.6	**1.0**	**3.1**	0.1	**0.95**	0.85	**0.99**
n	11	0	2	3	11	7	3	9	11	11	11

Absolute values of inter-model differences (Δ_m) at M12. Values in bold indicate greater than applied thresholds. [n] is the number of participants for which Gen and SubS were considered to be in agreement based on Δm. AUC expressed as BW*%Gait Cycle (BW.%c).

maximum values of $\Delta_t(JCF)$ for all joints (hip: 0.9 (0.3) BW, 1.0 (0.3) BW; knee: 0.8 (0.3) BW, 0.7 (0.3) BW; ankle: 1.0 (0.8) BW, 0.9 (0.6) BW for Gen and SubS, respectively).

A large within group variability (maximum SD: 28.5 BW.%c) was observed for the overall joint loading (AUC, Figure 3) measures, particularly at the hip and at the ankle, for both Gen and SubS. No group differences were found for these values, even if a tendency was observed at the knee, where SubS predictions of JCF at both time points tended to be on average, 10% lower than that of Gen, with higher estimates (>40%) from SubS observed only for one subject (still true at both time points).

Inter-model differences (Δ_m) at the knee were observed for the majority of participants at both time points when considering the AUC and JCF peaks whereas the opposite was true for all joints, looking at the waveform correlation between model predictions of JCF (Tables 2 and 3).

The matrix in Table 4 shows an overall good agreement between predictions from the two models

for most of the subjects, except for the R^2 that at the ankle showed a disagreement for about half of the participants.

Discussion

The aim of this study was to evaluate the consistency in measures of JCF and JCF changes over time, using subject-specific (SubS) and scaled generic (Gen) MSK models applied to 3 D gait data from a group of children with Juvenile Idiopathic Arthritis.

Several indices have been used by different authors to analyse the temporal curves usually obtained from gait data, with some looking into distinct parameters such as the mean value at a specified event, while others look at how a parameter of interest changes over the whole cycle (Kadaba et al. 1989; Chau 2001; Picerno et al. 2008; Di Marco et al. 2018). These different approaches yield complementary information associated to changes in peak values, amplitudes and phases of the curves which can all be of interest, depending on the outcome (e.g., maximum force vs

Table 4. Agreement in longitudinal changes (M12 - M0) in selected metrics between Gen and SubS predictions of JCF for each of 11 participants.

	Δ_t(AUC)			Δ_t(Peak)[BW]					$R^2[\Delta_t$(JCF)]			
	H	K	A	$P1_H$	$P2_H$	$P1_K$	$P2_K$	$P2_A$	H	K	A	Tot_s
S1	✓	✓	×	✓	✓	×	✓	×	✓	✓	×	7
S2	✓	✓	✓	✓	✓	✓	✓	×	✓	✓	×	9
S3	✓	×	×	×	✓	✓	×	✓	✓	×	✓	6
S4	✓	✓	✓	×	✓	✓	✓	✓	✓	✓	✓	10
S5	✓	✓	×	×	✓	✓	✓	✓	✓	✓	×	8
S6	✓	✓	✓	✓	✓	✓	✓	✓	×	×	✓	9
S7	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	×	10
S8	✓	×	×	✓	✓	×	✓	✓	✓	✓	✓	9
S9	✓	✓	✓	✓	✓	✓	✓	✓	✓	×	✓	10
S10	✓	✓	×	×	✓	✓	✓	✓	×	✓	×	7
S11	×	✓	✓	✓	✓	✓	✓	✓	×	×	×	7
Tot_g	10	9	7	7	11	9	10	9	8	7	5	

Change (Δ_t) in area under JCF curve (AUC), peak 1 (P1) and peak 2 (P2) and coefficient of determination for Gen against SubS ΔJCF. Subscripts represent hip (H), knee (K) and ankle (A). (✓) agreement, (×) disagreement. Tot_g and Tot_s refer to the number of participants with agreement for a particular metric and number of metrics in agreement per participant, respectively.

impulse). For this reason, in this study indices from both groups were used to capture the salient features of the subject gait as well as any time-dependent patterns in the data.

Within sessions, the outputs of the two models were similar in terms of range of motion and waveform profiles of joint angles, net joint moments and powers, despite the significant differences observed when looking at individual time instances between models over the gait cycle. Kainz et al. (2018) also found similarities when comparing joint kinematics and kinetics between scaled generic and MRI-based models of typically developing children. As expected in this study, differences were likewise observed in the JCF profiles and estimates between the two models, in line with what has generally been reported in the literature (Lenaerts et al. 2008; Scheys et al. 2008; Valente et al. 2014; Song et al. 2019). In particular, the range of JCF estimates observed for both models were comparable to values reported by other studies for the hip (Carriero et al. 2014), knee (Valente et al. 2014) and ankle (Prinold et al. 2016). The mean JCF peaks were also comparable to previous independent work conducted on a subset of children from the same cohort at a different time point observation (Modenese et al. 2018). Differences observed between the models likely originate from the personalisation of muscle origin and insertion points as well as joint centre and axis locations, both in children (Kainz et al. 2018) and adults (Scheys et al. 2008). An assessment of some selected muscle forces estimated by both models showed a general concurrence in timing of activity and magnitude during the gait cycle for most muscles (see supplementary data). As per the JCFs, the differences in calculated muscle forces between the two models remained similar across observations for all muscles, which was unsurprising, since muscle forces are known to be the main contributors to JCF (Correa et al. 2010) predictions.

When looking at specific points on the JCF loading profiles (Figure 1), lower hip loading response peaks (P1) were predicted by the Gen compared to the SubS. This is in contrast with what was reported by Wesseling et al. (2016), in a sample of adult subjects. An explanation for this disagreement could be the different methods used to calculate the maximal isometric force for Gen and SubS models in this study, this was kept the same for both models by Wesseling et al. However, it has been previously suggested that this should not have a significant influence on output muscle force and JCF estimates (Valente et al. 2014; Wesseling et al. 2016; Modenese et al. 2018). Further study would be needed to further explore this specific aspect.

A large variability between subjects was observed for the JCFs at all the joints, particularly at the knee when using SubS models. This high between-subject variability was in line with what was previously found when looking at joint kinematics, moments and knee JCF in a larger group from the same cohort (Montefiori et al. 2019a, 2019b). This may be attributable to participants adopting a variety of loading strategies to attenuate pain or discomfort resulting from swelling or inflammation of their joints (Montefiori et al. 2019b). This variability was partially masked by the scaled model with a reduced between subject variability indicating its less sensitive intersubject nature. We similarly observed larger between model differences in AUC at the knee (Tables 2 and 3), but these were expected since the SubS was implemented to have a simplified knee joint (extension/flexion) compared to the Gen which had an

additional prescription of tibial translation as a function of knee joint angle. Nonetheless, the latter was able to capture longitudinal changes at individual level consistently with the subject specific model for the hip and knee especially, as suggested by the results of the agreement matrix (Table 4). Taking the knee for example, although predictions of AUC were different between models at each observation, a majority of participants had $\Delta_t(AUC)$ in agreement in terms of whether there was an increase or decrease.

Comparing the differences between the JCF profiles estimated at the two time points for each of the models, it was observed that Gen and SubS provided consistent information in terms of increased or decreased JCFs between different phases of the gait cycle going from M0 to M12, even if these changes were of different magnitude. The difference in magnitude was particularly prominent at the ankle, which also explains why the number of participants with between model agreement (Tot_g) for $R^2[\Delta_t(JCF)]$ (Table 4) was the lowest for this joint. This was despite the observation that predictions of JCFs at the ankle had the most highly correlated waveforms between models at each observation. At group level, no statistically significant difference in longitudinal change in JCFs between Gen and SubS was observed. Overall, these results indicate that both models were able to account for changes in the JCFs likely attributable to changes in the patient's pathological condition.

The sample size involved in this study is small, although larger than other studies comparing generic and image-based subject-specific models (Correa et al. 2011; Bosmans et al. 2015; Wesseling et al. 2016; Kainz et al. 2018). Moreover, the investigated group is a good representation of a very heterogeneous patient population (Montefiori et al. 2019b), as also indicated by the reported between subject variability in calculated JCF, which may be considered an advantage in terms of applicability of the reported results. It must be acknowledged, however, that even if Juvenile Idiopathic Arthritis might cause bone deformities, these were not evident in the investigated group of children and no clinically meaningful longitudinal changes emerged from the analysis of their MRIs. This does of course limit the generalisability of the reported results to populations with large bony deformities. The inability of scaled generic models to account for significant anatomical alterations, such as increased femoral anteversion or tibial torsion, have been reported to impact significantly on predictions of moment arm lengths (Scheys et al. 2008).

The SubS was considered as the gold standard in this study as it is assumed to be more representative of the subject's anatomy than the generic models. This assumption of course has its limitations such as errors in operator input but was the only one possible due to the unavailability of longitudinal data from instrumented prosthesis, especially for children.

Conclusion

In conclusion, this study evaluated the consistency between scaled generic and subject-specific model estimates of longitudinal changes in JCF for a population of children with JIA. By using different metrics for reporting JCF, it was shown that even if the estimates of JCF can be highly different at a single time-point, the two models showed agreement when calculating the longitudinal difference in joint contact forces, particularly at the hip. It is hence suggested, albeit with caution, that scaled generic models can be used as an initial and easily implementable modelling approach when interest is in trends rather than exact estimates.

Disclosure statement

Acknowledgement

Funding

References

Arnold EM, Ward SR, Lieber RL, Delp SL. 2010. A model of the lower limb for analysis of human movement. Ann Biomed Eng. 38(2):269–279.

Baker R, Esquenazi A, Benedetti MG, Desloovere K. 2016. Gait analysis: clinical facts. Eur J Phys Rehabil Med. 52(4):560–574.

Blemker SS, Asakawa DS, Gold GE, Delp SL. 2007. Image-based musculoskeletal modeling: applications, advances, and future opportunities. J Magn Reson Imaging. 25(2): 441–451.

Bosmans L, Jansen K, Wesseling M, Molenaers G, Scheys L, Jonkers I. 2016. The role of altered proximal femoral geometry in impaired pelvis stability and hip control during CP gait: a simulation study. Gait Posture. 44: 61–67.

Bosmans L, Valente G, Wesseling M, Van Campen A, De Groote F, De Schutter J, Jonkers I. 2015. Sensitivity of predicted muscle forces during gait to anatomical variability in musculotendon geometry. J Biomech. 48(10): 2116–2123.

Carriero A, Zavatsky A, Stebbins J, Theologis T, Lenaerts G, Jonkers I, Shefelbine SJ. 2014. Influence of altered gait patterns on the hip joint contact forces. Comput Methods Biomech Biomed Engin. 17(4):352–359.

Chau T. 2001. A review of analytical techniques for gait data. Part 1: Fuzzy, statistical and fractal methods. Gait Posture. 13(1):49–66.

Correa TA, Baker R, Graham HK, Pandy MG. 2011. Accuracy of generic musculoskeletal models in predicting the functional roles of muscles in human gait. J Biomech. 44(11):2096–2105.

Correa TA, Crossley KM, Kim HJ, Pandy MG. 2010. Contributions of individual muscles to hip joint contact force in normal walking. J Biomech. 43(8):1618–1622.

Correa TA, Pandy MG. 2011. A mass-length scaling law for modeling muscle strength in the lower limb. J Biomech. 44(16):2782–2789.

Delp SL, Anderson FC, Arnold AS, Loan P, Habib A, John CT, Guendelman E, Thelen DG. 2007. OpenSim: open-source software to create and analyze dynamic simulations of movement. IEEE Trans Biomed Eng. 54(11): 1940–1950.

Delp SL, Loan JP, Hoy MG, Zajac FE, Topp EL, Rosen JM. 1990. An interactive graphics-based model of the lower extremity to study orthopaedic surgical procedures. IEEE Trans Biomed Eng. 37(8):757–767.

Di Marco R, Scalona E, Pacilli A, Cappa P, Mazzà C, Rossi S. 2018. How to choose and interpret similarity indices to quantify the variability in gait joint kinematics. Int Biomech. 5(1):1–8.

Duda GN, Brand D, Freitag S, Lierse W, Schneider E. 1996. Variability of femoral muscle attachments. J Biomech. 29(9):1185–1190.

Hainisch R, Gfoehler M, Zubayer-Ul-Karim M, Pandy MG. 2012. Method for determining musculotendon parameters in subject-specific musculoskeletal models of children developed from MRI data. Multibody Syst Dyn. 28(1–2):143–156.

Harrington ME, Zavatsky AB, Lawson SE, Yuan Z, Theologis TN. 2007. Prediction of the hip joint centre in adults, children, and patients with cerebral palsy based on magnetic resonance imaging. J Biomech. 40(3): 595–602.

Hicks JL, Uchida TK, Seth A, Rajagopal A, Delp SL. 2015. Is my model good enough? Best practices for verification and validation of musculoskeletal models and simulations of movement. J Biomech Eng. 137(2):020905.

Iosa M, Cereatti A, Merlo A, Campanini I, Paolucci S, Cappozzo A. 2014. Assessment of waveform similarity in clinical gait data: the linear fit method. Biomed Res Int. 2014:214156.

Kadaba MP, Ramakrishnan HK, Wootten ME, Gainey J, Gorton G, Cochran GV. 1989. Repeatability of kinematic, kinetic, and electromyographic data in normal adult gait. J Orthop Res. 7(6):849–860.

Kainz H, Hoang H, Pitto L, Wesseling M, Van Rossom S, Van Campenhout A, Molenaers G, De Groote F, Desloovere K, Jonkers I. 2019. Selective dorsal rhizotomy improves muscle forces during walking in children with spastic cerebral palsy. Clin Biomech (Bristol, Avon)). 65:26–33.

Kainz H, Hoang HX, Stockton C, Boyd RR, Lloyd DG, Carty CP. 2017. Accuracy and reliability of marker-based approaches to scale the pelvis, thigh, and shank segments in musculoskeletal models. J Appl Biomech. 33(5):354–360.

Kainz H, Wesseling M, Pitto L, Falisse A, Van Rossom S, Van Campenhout A, De Groote F, Desloovere K, Carty C, Jonkers I. 2018. O 107 – Impact of subject-specific musculoskeletal geometry on estimated joint kinematics, joint kinetics and muscle forces in typically developing children. Gait Posture. 65:223–225.

Lenaerts G, Bartels W, Gelaude F, Mulier M, Spaepen A, Van der Perre G, Jonkers I. 2009. Subject-specific hip geometry and hip joint centre location affects calculated contact forces at the hip during gait. J Biomech. 42(9):1246–1251.

Lenaerts G, De Groote F, Demeulenaere B, Mulier M, Van der Perre G, Spaepen A, Jonkers I. 2008. Subject-specific hip geometry affects predicted hip joint contact forces during gait. J Biomech. 41(6):1243–1252.

Modenese L, Montefiori E, Wang A, Wesarg S, Viceconti M, Mazza C. 2018. Investigation of the dependence of joint contact forces on musculotendon parameters using a codified workflow for image-based modelling. J Biomech. 73:108–118.

Modenese L, Phillips A. 2012. Prediction of hip contact forces and muscle activations during walking at different speeds. Multibody Syst Dyn. 28(1–2):157–168.

Modenese L, Phillips AT, Bull AM. 2011. An open source lower limb model: Hip joint validation. J Biomech. 44(12):2185–2193.

Montefiori E, Modenese L, Di Marco R, Magni-Manzoni S, Malattia C, Petrarca M, Ronchetti A, de Horatio LT, van Dijkhuizen P, Wang A, et al. 2019a. An image-based kinematic model of the tibiotalar and subtalar joints and its application to gait analysis in children with Juvenile Idiopathic Arthritis. J Biomech. 85:27–36.

Montefiori E, Modenese L, Di Marco R, Magni-Manzoni S, Malattia C, Petrarca M, Ronchetti A, de Horatio LT, van Dijkhuizen P, Wang A, et al. 2019b. Linking joint impairment and gait biomechanics in patients with juvenile idiopathic arthritis. Ann Biomed Eng. 47(11):2155–2167.

Pataky TC. 2012. One-dimensional statistical parametric mapping in Python. Comput Methods Biomech Biomed Eng. 15(3):295–301.

Picerno P, Cereatti A, Cappozzo A. 2008. Joint kinematics estimate using wearable inertial and magnetic sensing modules. Gait Posture. 28(4):588–595.

Prinold JA, Mazza C, Di Marco R, Hannah I, Malattia C, Magni-Manzoni S, Petrarca M, Ronchetti AB, Tanturri de Horatio L, van Dijkhuizen EH, et al. 2016. A patient-specific foot model for the estimate of ankle joint forces in patients with juvenile idiopathic arthritis. Ann Biomed Eng. 44(1):247–257.

Scheys L, Spaepen A, Suetens P, Jonkers I. 2008. Calculated moment-arm and muscle-tendon lengths during gait differ

substantially using MR based versus rescaled generic lower-limb musculoskeletal models. Gait Posture. 28(4):640–648.

Scheys L, Van Campenhout A, Spaepen A, Suetens P, Jonkers I. 2008. Personalized MR-based musculoskeletal models compared to rescaled generic models in the presence of increased femoral anteversion: effect on hip moment arm lengths. Gait Posture. 28(3):358–365.

Song K, Anderson AE, Weiss JA, Harris MD. 2019. Musculoskeletal models with generic and subject-specific geometry estimate different joint biomechanics in dysplastic hips. Comput Methods Biomech Biomed Eng. 22(3):259–270.

Stebbins J, Harrington M, Thompson N, Zavatsky A, Theologis T. 2006. Repeatability of a model for measuring multi-segment foot kinematics in children. Gait Posture. 23(4):401–410.

Steele KM, Demers MS, Schwartz MH, Delp SL. 2012. Compressive tibiofemoral force during crouch gait. Gait Posture. 35(4):556–560.

Steele KM, van der Krogt MM, Schwartz MH, Delp SL. 2012. How much muscle strength is required to walk in a crouch gait? J Biomech. 45(15):2564–2569.

Valente G, Crimi G, Vanella N, Schileo E, Taddei F. 2017. nmsBuilder: Freeware to create subject-specific musculoskeletal models for OpenSim. Comput Methods Programs Biomed. 152:85–92.

Valente G, Pitto L, Testi D, Seth A, Delp SL, Stagni R, Viceconti M, Taddei F. 2014. Are subject-specific musculoskeletal models robust to the uncertainties in parameter identification? Plos One. 9(11):e112625.

Wesseling M, De Groote F, Bosmans L, Bartels W, Meyer C, Desloovere K, Jonkers I. 2016. Subject-specific geometrical detail rather than cost function formulation affects hip loading calculation. Comput Methods Biomech Biomed Eng. 19(14):1475–1488.

Wesseling M, De Groote F, Meyer C, Corten K, Simon JP, Desloovere K, Jonkers I. 2016. Subject-specific musculoskeletal modelling in patients before and after total hip arthroplasty. Comput Methods Biomech Biomed Eng. 19(15):1683–1691.

Wesseling M, Ranz EC, Jonkers I, et al. 2017. Objectifying treatment outcomes using musculoskeletal modelling-based simulations of motion. In: Müller B, Wolf SI, Brueggemann G-P, editors. Handbook of human motion. Cham: Springer International Publishing; p. 1–25.

White DR, Woodard HQ, Hammond SM. 1987. Average soft-tissue and bone models for use in radiation dosimetry. Br J Radiol. 60(717):907–913.

Appendix
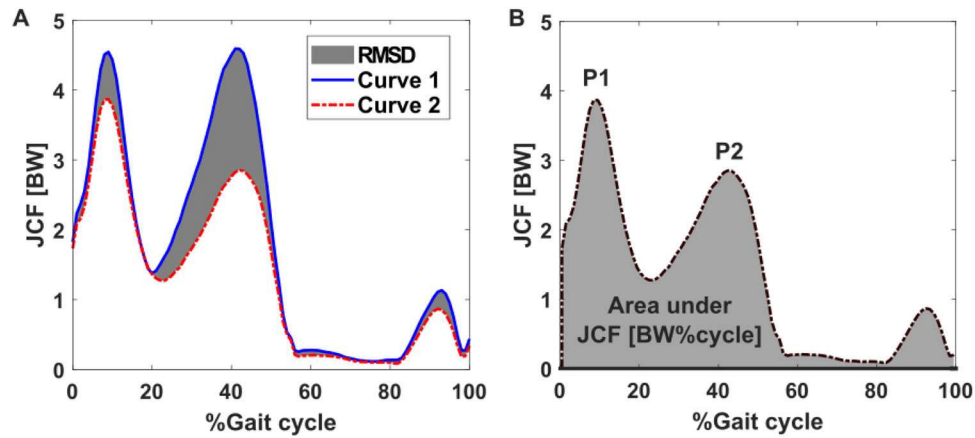
Figure A1 Different indices for the description and analysis of the simulation results. A: Goodness of fit was assessed with the RMSD. B: Joint loading during the loading response (P1) and push off (P2) phases were analysed using the peak values and overall joint loading by the area under JCF/BW curve.
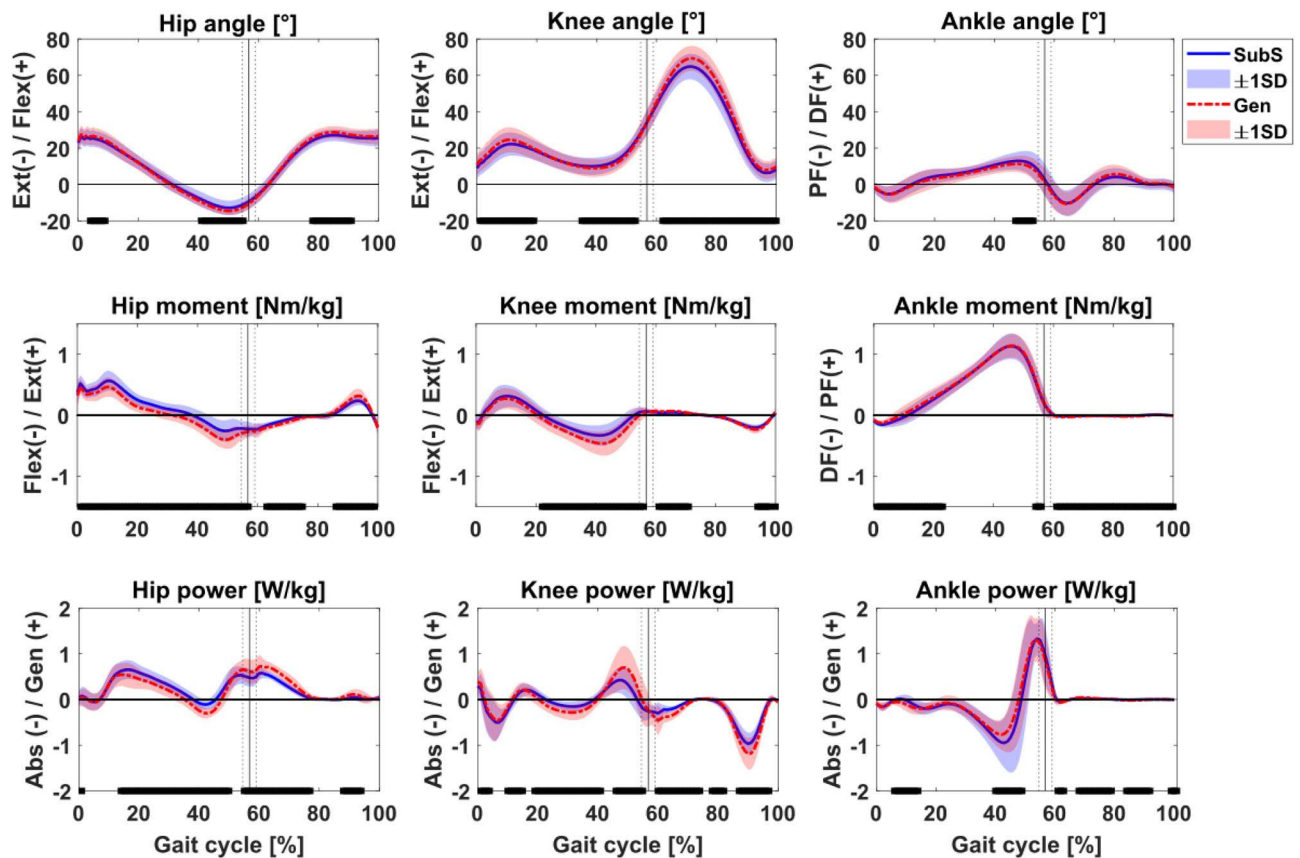
Figure A2 Comparison between Gen (red) and SubS (blue) model estimations of sagittal plane joint angles, moments and powers at all observations for 11 juvenile participants. Joint moments and powers normalised by body mass. Black bars indicate significance at $P < 0.05$ according to the non-parametric one-sample paired t-test. Extension/Flexion (Ext/Flex), Plantarflexion/Dorsiflexion (PF/DF) and Absorption/Generation (Abs/Gen).
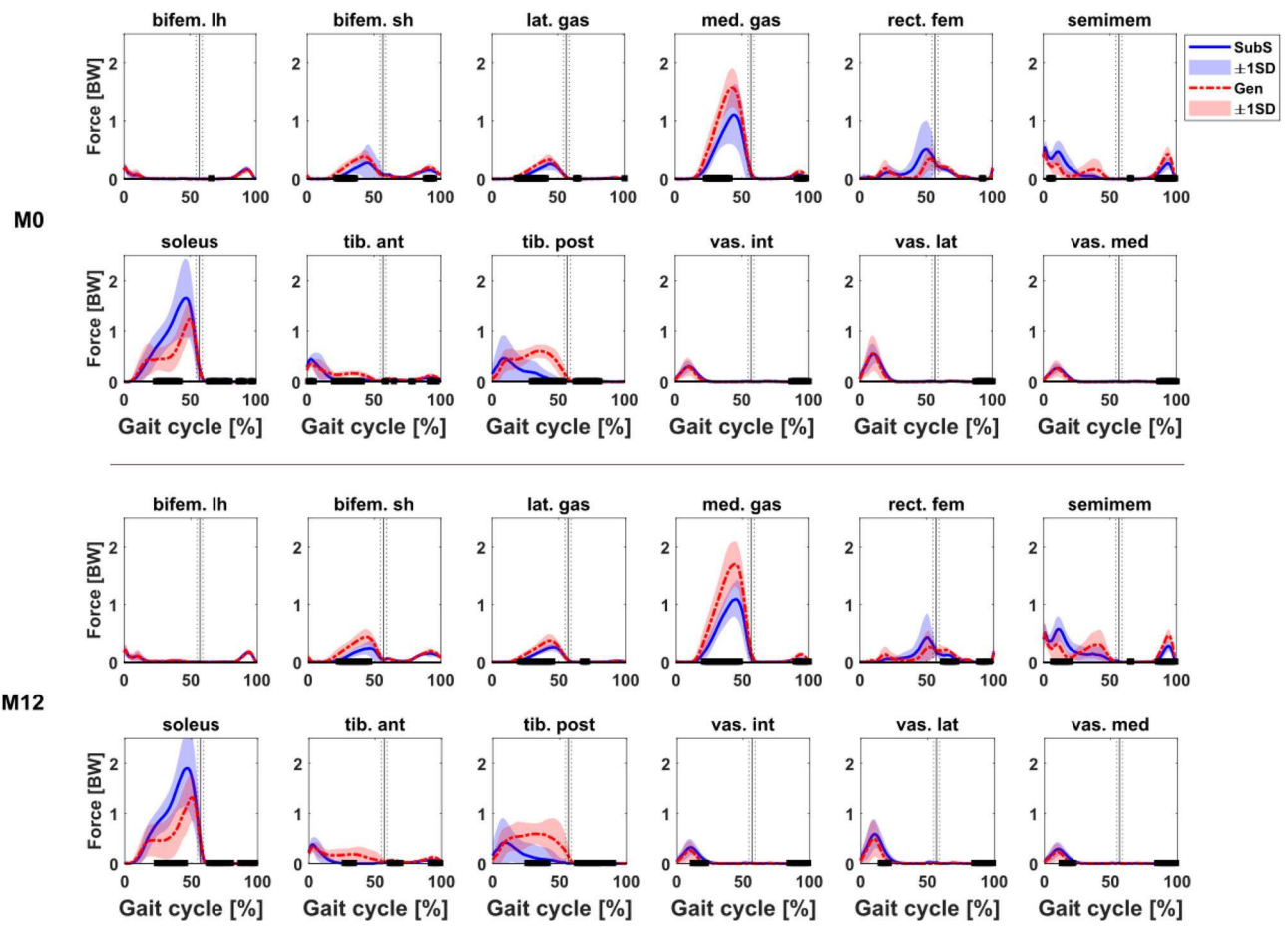
Figure A3. Muscle forces estimated by the Gen (red) and SubS (blue) models for 12 selected muscles at M0 and M12. Black bars indicate statistical significance at P < 0.05 between models. Selected muscles are biceps femoris long head, biceps femoris short head, gastrocnemius lateralis, gastrocnemius medialis, rectus femoris, semimembranosus, soleus, tibialis anterior, tibialis posterior, vastus intermedius, vastus lateralis and vastus medialis.

Table A1. p-values for Wilcoxon sign ranked test between M0 and M12 for Δ_m(RMSD).

	Hip	Knee	Ankle
p-value	0.929	0.285	0.959
Effect size (Cohen's d)	0.069	−0.037	−0.302

Table A2. P-values for Wilcoxon sign ranked test between Gen and SubS for Δ_t(RMSD).

	Hip	Knee	Ankle
p-value	0.965	0.779	0.859
Effect size (Cohen's d)	−0.043	−0.156	0.092

Table A3. p-values for Wilcoxon sign ranked test between Gen and SubS estimates of AUC(JCF).

	M0			M12		
	Hip	Knee	Ankle	Hip	Knee	Ankle
p-value	0.286	0.050	0.182	0.424	0.050	0.091
Effect size (Cohen's d)	−0.320	0.454	0.541	−0.072	1.009	0.648