The Journal of Arthroplasty xxx (2019) 1–9

Contents lists available at ScienceDirect

The Journal of Arthroplasty

journal homepage: www.arthroplastyjournal.org

Current Preclinical Testing of New Hip Arthroplasty Technologies Does Not Reflect Real-World Loadings: Capturing Patient-Specific and Activity-Related Variation in Hip Contact Forces

David E. Lunn, PhD [a, b], Enrico De Pieri, PhD [c], Graham J. Chapman, PhD [a, b, d], Morten E. Lund, PhD [e], Anthony C. Redmond, PhD [a, b, *], Stephen J. Ferguson, PhD [c]

[a] Institute for Rheumatic and Musculoskeletal Medicine, University of Leeds, Leeds, United Kingdom
[b] NIHR Leeds Biomedical Research Centre, Leeds, United Kingdom
[c] Institute for Biomechanics, ETH Zurich, Zurich, Switzerland
[d] School of Health Sciences, University of Central Lancashire, Preston, United Kingdom
[e] AnyBody Technology A/S, Aalborg, Denmark

ARTICLE INFO

Article history:
Received 2 August 2019
Received in revised form
29 August 2019
Accepted 3 October 2019
Available online xxx

Keywords:
total hip arthroplasty
hip contact force
functional outcomes
activities of daily living
biomechanics

ABSTRACT

Background: Total hip arthroplasty (THA) implants are routinely tested for their tribological performance through regulatory preclinical wear testing (eg, ISO-14242). The standardized loading conditions defined in these tests consist of simplified waveforms, which do not specifically represent in vivo loads in different groups of patients. The aim of this study is to investigate, through musculoskeletal modeling, patient-specific and activity-related variation in hip contact forces (HCFs) in a large cohort of THA patients during common activities of daily living (ADLs).
Methods: A total of 132 THA patients participated in a motion-capture analysis while performing different ADLs, including walk, fast walk, stair ascent, and descent (locomotor); sit to stand, stand to sit, squat, and lunge (nonlocomotor). HCFs were then calculated using the AnyBody Modeling System and qualitatively compared across all activities. The influence of gender on HCFs was analyzed through statistical parametric mapping analysis.
Results: Systematic differences were found in HCF magnitudes and individual components in both locomotor and nonlocomotor ADLs. The qualitative analysis of the ADLs revealed a large range and a large variability in forces experienced at the hip during different activities. Significant differences in the 3-dimensional loading patterns were observed between males and females across most activities.
Conclusion: THA patients present a large variability in the forces experienced at the hip joint during their daily life. The interpatient variation might partially explain the heterogeneity observed in implant survival rates. A more extensive preclinical implant testing standard under clinically relevant loading conditions has been advocated to better predict and avoid clinical wear problems.

Total hip arthroplasty (THA) is the most effective form of treatment for severe hip osteoarthritis [1–3], reducing pain and restoring mobility in arthritic patients [4–6]. Monitoring of implant survivorship revealed survival rates greater than 95% at 10 years, but with this number falling to 58% after 25 years [7]. However, the overall demand for THA is expected to increase in the future as a consequence of a demographic shift toward an aging population [8,9].

David E. Lunn and Enrico De Pieri equally contributed to this work.

One or more of the authors of this paper have disclosed potential or pertinent conflicts of interest, which may include receipt of payment, either direct or indirect, institutional support, or association with an entity in the biomedical field which may be perceived to have potential conflict of interest with this work. For full disclosure statements refer to https://doi.org/10.1016/j.arth.2019.10.006.

Funding Source: The funding source had no role in the study design, collection, analysis, and interpretation of the data; in the writing of the manuscript; or in the decision to submit the manuscript for publication.

* Reprint requests: Anthony C. Redmond, PhD, Leeds Institute of Rheumatic and Musculoskeletal Medicine, NIHR Leeds Biomedical Research Centre, 2nd Floor, Chapel Allerton Hospital, Harehills Lane, Leeds LS7 4SA, United Kingdom.

https://doi.org/10.1016/j.arth.2019.10.006

The ever-improving survivorship of joint arthroplasties is evidence of the important continued innovation and improvement in implant design and surgical techniques, which has led to better implant fixation options, improved wear performance, and reduction in perioperative and postoperative complication rates [2,3]. The outcomes of novel implant design are not however always better than for existing implants [10,11]. Occasionally, innovation leads to unforeseen problems such as early implant failure, as exemplified by the DePuy articular surface replacement (ASR) hip implant, which failed because of higher than anticipated wear between the 2 metal bearing surfaces [12–14]. One lesson to be learned from the ASR was the differential in failure rates for population-level factors such as gender, which had not been identified during preclinical testing but which resulted in significantly higher failure rates in women for the ASR implant [15]. These demographic-dependent design shortcomings are not only associated with frank and widespread early failures but can be observed, albeit more subtly, in broader registry data, with younger patients and, in contrast to the ASR hip, male patients typically having an increased lifetime risk of revision [16,17]. It has also become apparent that different implant combinations perform better in different patient groups [18]. While such information is useful when gathered retrospectively and is well suited to monitoring performance of tried and tested combinations such as metal-on-polyethylene bearings, it would be better to be able to predict the likely outcomes of novel designs during any preclinical testing.

The majority of hip arthroplasty failures are caused by wear [19] which is a consequence of load and motion as determined by the amount and type of physical activity undertaken by the patient [20]. Additionally, different patient factors such as age, gender, weight, activity level, and patient-specific kinematic patterns have shown a correlation with wear [20–22]. Therefore, the differential failures due to wear are potentially predictable and testable.

Joint arthroplasties are routinely tested for their tribological performance before being introduced to the market by means of standardized tests [23]. Current regulatory preclinical testing standards, such as the ISO 14242-1, define standardized loading conditions consisting of simplified and stylized waveforms, which do not directly represent in vivo loads and motion in different groups of patients. The loading profile defined in the ISO 14242-1 preclinical testing standard is given in Figure 1. A more extensive implant testing under clinically relevant loading conditions [23] has been suggested to be warranted to predict and avoid clinical wear problems, which could have been better anticipated in the case of the ASR implant system [14].

These deficiencies in preclinical testing have been highlighted when comparing the ISO 14242-1 testing waveforms to real-world hip joint contact forces (HCFs) measured through instrumented implants [24], particularly when comparing the ISO model to the larger and varied loading pattern observed when performing real-world activities of daily living (ADLs) [25]. Due to the inherently invasive nature of in vivo HCF measurement via instrumented implants, data are only available for a small number of patients and thus has not captured the variation which exists in larger populations. Advances in computational techniques such as musculoskeletal modeling have shown potential for estimating accurate HCFs noninvasively [26], and these techniques are much better suited to describing the load variability observed in larger populations [27].

The aim of the present study is to explore differences in HCFs between patient groups in a relatively large sample of hip arthroplasty cases and to further investigate these differences during a selection of the real-world ADLs to which a hip implant is typically exposed in vivo.

Methods

A total of 132 THA patients were recruited into the study through a clinical database of surgical cases as part of the LifeLongJoints (LLJ) patients' cohort. Inclusion criteria for the hip arthroplasty group were as follows: between 1 and 5 years THA postsurgery, older than 18 years of age, no lower limb joint replaced other than hip joint(s), fully pain free, and not having any other orthopedic or neurologic problem which may compromise gait. Ethical approval was obtained via the UK national NHS ethics (IRAS) system and all participants provided informed, written consent.

Motion-Capture Data Acquisition

Patients undertook a series of ADLs during which lower-limb kinematics and kinetics were acquired using a 10 camera Vicon system (Vicon MX; Oxford Metrics, UK) sampling at 100 Hz, integrated with 2 force plates (AMTI, Watertown, MA) capturing at 1000 Hz. The CAST marker set was used to track lower limb segment kinematics in 6 degrees of freedom. A more detailed description can be found in [28]. For the THA group, the operated limb (or in bilateral cases, the most recently operated limb) was used for analysis.

Patient Characteristics

This patient cohort has been previously shown to demonstrate a large variability in hip loadings during gait, which were shown to be dependent on patient characteristics, particularly on body mass index and the patients' functional ability determined by their self-selected walking speed [27]. To further investigate the load

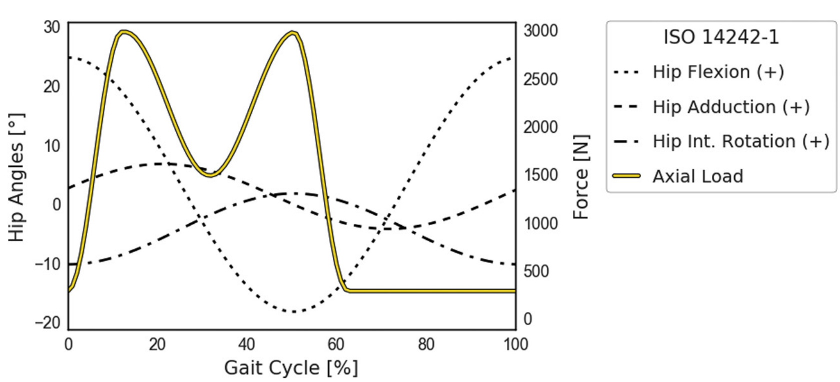

Fig. 1. Current preclinical testing standard ISO 14242-1. Axial load reported in the solid yellow line.

variability in this patients' cohort, we stratified the patients by gender, which represents an important differentiator of implant survivorship. Patients were allowed to individually exclude activities that they were not able to perform relatively comfortably. Patient demographics for each activity are reported in Table 1.

Activities of Daily Living

The ADLs are grouped into 2 categories: locomotor tasks (walk, fast walk, stair ascent, and stair descent) and nonlocomotor tasks (sit to stand, stand to sit, squat, and lunge). Information regarding the protocol of each task can be found at (https://doi.org/10.5518/319), while a brief description is provided below.

Walking Tasks

Patients undertook 2 walking conditions (1) at a self-selected walking speed (hereafter referred to as a normal walk) and (2) a fast walk, where patients were instructed to walk "as fast as possible without running" along a 10-m walkway. All trials were time-normalized from heel strike (0%), to heel strike (100%) and interpolated to 1% steps (101 points).

Stair Negotiation

Patients were asked to ascend and descend 3 steps at self-selected comfortable speed, without the use of a handrail. The stair case was mounted and bolted to the force plates [29] to collect ground reaction force data. All trials were time-normalized from foot strike (0%), to foot strike (100%) and interpolated to 1% steps (101 points).

Standing and Sitting

During the sitting and standing trials, patients sat on a platform with the feet shoulder-width apart, each foot positioned on a separate force plate in a fixed position. The seat height was matched to the level the patient's tibial plateau. Patients were then asked to stand and return to a seated position without use of the arms which were held out straight ahead, to avoid any occlusion of the markers.

Lunge

Lunge was chosen to replicate relevant sports activities such as lawn green bowls and tennis. Patients were asked to stand with both feet on one force plate and lunge forward, leading with the study limb, onto the adjacent force plate return to standing.

Squat

Squatting or a variation of a squat is performed on a daily basis [30] and therefore is important to assess. Patients were positioned with one foot on each force plate shoulder width apart and were asked to squat as low as comfortably possible with arms out in front of them to avoid marker occlusion.

Data Processing

All markers were labeled and gap-filled using the spline fill function in Vicon Nexus 2.5 (Vicon MX, Oxford Metrics, UK), before the labeled marker coordinates and kinetic data were exported to Visual 3D modeling software (C-motion) for further analysis. Kinematic data were filtered using a low-pass (6 Hz) Butterworth filter. Ground reaction force (GRF) data were filtered using a low-pass Butterworth filter (25 Hz).

Musculoskeletal Modeling

Musculoskeletal simulations were performed using a commercially available software (AnyBody Modeling System, version 7.1, Aalborg, Denmark). A detailed musculoskeletal model of the lower limb [26] based on a cadaveric dataset [31] was scaled to match the anthropometrics of each patient based on marker data collected during a static trial [32]. Marker trajectories and GRF data from each gait trial served as input to an inverse dynamics analysis, based on a third-order-polynomial muscle recruitment criterion, to calculate muscle forces and HCFs. A total of 2148 trials were processed and analyzed through the toolkit AnyPyTools [33]. The HCF components were defined in a common femur-based reference frame [24] and averages for each patient during each individual ADL were computed.

HCF Analysis

The mean resultant HCFs, with relative ranges of variation, predicted across this cohort were qualitatively compared to measurements from instrumented implants reported in the Orthoload database [34] for matching ADLs.

Mean resultant HCFs and their individual components, with associated 95% confidence intervals, are also qualitatively compared across different ADLs and the peak values are reported.

Additionally, individual patients' loading profiles across activities were investigated and the data from 1 representative low-functioning and 1 high-functioning patient were reported in relation to the cohort as a whole. Functional level was defined by the self-selected gait speed as reported previously [27,28].

Statistical Parametric Mapping Analysis

The mean computed HCFs for each patient and activity were then normalized to each patient's body mass. The normalized HCFs were analyzed using Statistical Parametric Mapping (SPM; www.spm1D.org, v0.4). The 3 individual force components were regarded as a vector field, describing the 3-dimensional variation over time of the HCF vector trajectory. A 2-sample Hotelling's T^2 test, the vectorial analog of a scalar *t*-test [35], was carried out to evaluate the influence of gender on the contact forces. The use of vector field analysis takes into consideration covariance between force

Table 1
Patient Demographics for Each Activity of Daily Living.

Activity	No. of Patients	Body Mass (kg)	Height (cm)	BMI (kg/m^2)	Male/Female	Age (y)	Years Since THA
Walk	132	78.10 (12.79)	166.28 (8.40)	28.20 (3.85)	66/66	71.62 (7.61)	2.80 (1.42)
Fast	117	78.59 (12.81)	167.36 (8.08)	27.99 (3.71)	62/55	70.56 (7.31)	2.84 (1.43)
Ascent	49	80.13 (13.81)	167.55 (9.37)	28.50 (4.03)	28/21	69.90 (7.70)	3.00 (1.47)
Descent	47	79.87 (14.12)	168.01 (9.34)	28.22 (3.92)	28/19	70.00 (7.87)	3.09 (1.46)
Sit	131	78.08 (12.83)	166.25 (8.42)	28.20 (3.86)	65/66	71.57 (7.61)	2.82 (1.42)
Stand	131	78.08 (12.83)	166.25 (8.42)	28.20 (3.86)	65/66	71.57 (7.61)	2.82 (1.42)
Squat	34	78.45 (11.80)	169.74 (6.23)	27.20 (3.60)	23/11	67.24 (6.28)	3.18 (1.59)
Lunge	35	75.89 (11.64)	167.23 (6.41)	27.09 (3.53)	22/13	70.29 (6.85)	2.57 (1.58)

Values are reported as mean (SD) unless otherwise stated.
SD, standard deviation; BMI, body mass index; THA, total hip arthroplasty.

components, thus reducing errors due to covariation bias. Technical details and practical examples are provided elsewhere [35]. The output test statistic SPM $\{T^2\}$ was evaluated at each point in the time series of each activity. Significance level was set at $\alpha = 0.05$, and the corresponding T^{2*} critical threshold was calculated based on the temporal smoothness of the input data through random field theory. Finally, the probability that similar suprathreshold regions would have occurred from equally smooth random waveforms was calculated. Post hoc scalar field *t*-tests were also conducted using SPM on each force component separately, with Bonferroni-corrected significance threshold levels set at $\alpha = 0.05/3 = 0.017$. Only differences which were statistically significant for more than 2% of the gait cycle are discussed.

Results

HCFs During Activities of Daily Living

The predicted resultant contact forces for the new LLJ patients' cohort showed comparable trends and mean absolute values with previous HCF data derived from the small-sample instrumented prosthesis Orthoload studies for all the compared activities (Fig. 2). Standing up from a chair presented a lower peak HCF value compared to the patients fitted with instrumented prostheses, although one of the instrumented implant patients was reported to have confounding contralateral hip pain. Stair ascent and descent showed similar trends and peak values, although with a shift in the temporal frame. The ranges of variation in the predicted HCF were generally wider, particularly for the locomotive activities, as might be expected from a larger cohort of patients.

The comparison of individual force components across ADLs (Fig. 3) reveals qualitative differences between the waveform profiles. The different locomotive tasks (Fig. 3A) show higher resultant mean peak values for fast walk (3086.1N), stair ascent (2822.7N), and stair descent (2897.5N) compared to level walking (2449.1N). Additionally, stair ascent and descent demonstrated an extended and higher HCF from heel strike to toe off compared to level walking, while fast walking in our cohort is characterized by a more pronounced excursion in HCF magnitude, with higher peak values and a lower force during midstance. Similar trends emerged for the proximodistal component (Fig. 3B). Fast walk and stair ascent present mean peak medial forces approximately 25% higher compared to level walking and stair descent (Fig. 3D). Similarly, fast walk and stair ascent HCF are also characterized by a concurrent higher peak posterior force compared to level walking, while stair descent present an extended posterior load throughout the loaded phase (Fig. 3C).

The larger kinematic variability in the nonlocomotive tasks translated in more evident waveform differences in the contact forces. Lunge, as the only activity that creates an intentional asymmetry in the load distribution between the 2 limbs, yielded a higher resultant HCF, with a mean peak value of 2506.1N, compared to squat (1694.4N), stand up (1280.4N), and sit down (1247.2N; Fig. 3E). The same trend could be observed for the proximodistal and mediolateral force components. Lunges also result in a peak posterior force that is approximately 3 times higher than the other activities (Fig. 3G).

HCFs Stratified by Gender

The vector-field analysis of HCF revealed significant differences between male and female patients during all locomotive activities, as well as sit down and stand up from a chair (Fig. 4). During walking, significant differences of up to 0.49*body weight (BW) higher in males were observed between 5% and 14%, 28% and 44%, 57% and 72%, and 91% and 96% of the walking cycle. For fast walking, significant differences of up to 0.56*BW greater in males were observed between 6% and 16%, 58% and 69%, and 90% and 96% of the walking cycle, while stair descent presented significant differences (up to 0.28*BW higher in females) between 25% and 31% of the activity. Despite males and females presenting similar HCF magnitudes during stair ascent, the vector-field analysis also revealed significant differences between 43% and 57% of the stair ascent cycle in the order of 0.46*BW, indicating that differences between male and females exist in the 3-dimensional trajectory of the force vector (Fig. 4C). The test statistics continuum SPM $\{T^2\}$ obtained from the vector-field analysis, as well as the full results of the post hoc *t*-tests for the individual force components, are reported for each activity as supplementary material (Figs. A1-A8).

Discussion

This study has highlighted the general variability in the magnitudes and patterns of hip loading that might be expected in larger cohort and has identified statistically significant and clinically meaningful differences between males and females following THA, across a range of ADLs. The large interpatient variability might, in vivo, be expected to lead to differing amounts of wear and differing failure rates in subgroups of patients undergoing hip arthroplasty. Full datasets for 1 representative high-functioning patient and 1 lower-functioning patient demonstrating this variability are available at https://doi.org/10.5518/319. Previous studies have demonstrated that applications of musculoskeletal models can be used to reliably predict contact forces for a large cohort of patients during gait [26,27]. It was previously shown that different patient characteristics influence both kinematics [28] and loads experienced at hip [27], with patient's overall functionality being a highly influential factor in determining variability in kinematics and kinetics during gait. The present study has also further illustrated the comparability of the computational modeling approach to the Orthoload dataset across 5 additional ADLs.

It is worth noting however that our methods do have number of limitations which are inherent when using computational modeling. The HCFs predicted in this study were obtained from scaled generic models and a certain level of error in the prediction of forces might persist, due to uncertainty in marker positioning [32] and lack of subject-specific anthropometric imaging data [36]. Additionally, scaled generic models do not account fully for anatomic differences between genders [37] or patient-specific implant measures, which could have improved the models' predictions [38].

We found significant differences between males and females in HCFs normalized by BW across all locomotor activities as well as sit down and stand up from a chair. Differences in the HCF vectorial trajectories indicate that there are functional differences between the 2 patient groups. The different 3-dimensional loading pattern, combined with different absolute load magnitudes, which can be expected in association with weight differences between genders, could affect the implant behavior and play a role in differing implant survival rates particularly in younger male and female patients. It is notable that while there are gender-related differences in risk of revision for people undergoing surgery up to the age of 75 years [16], the risk is comparable for patients older than 79 years old, suggesting that failure rates are not constant and probably depend on a combination of factors, such as patient-specific kinematics [22,39]. This lack of clear understanding is highlighted in the failure rates for ASR implants which were unexpectedly higher in females [15]. One way to predict how these patient-level factors might affect outcomes would be through more representative preclinical testing. The current standardization of preclinical

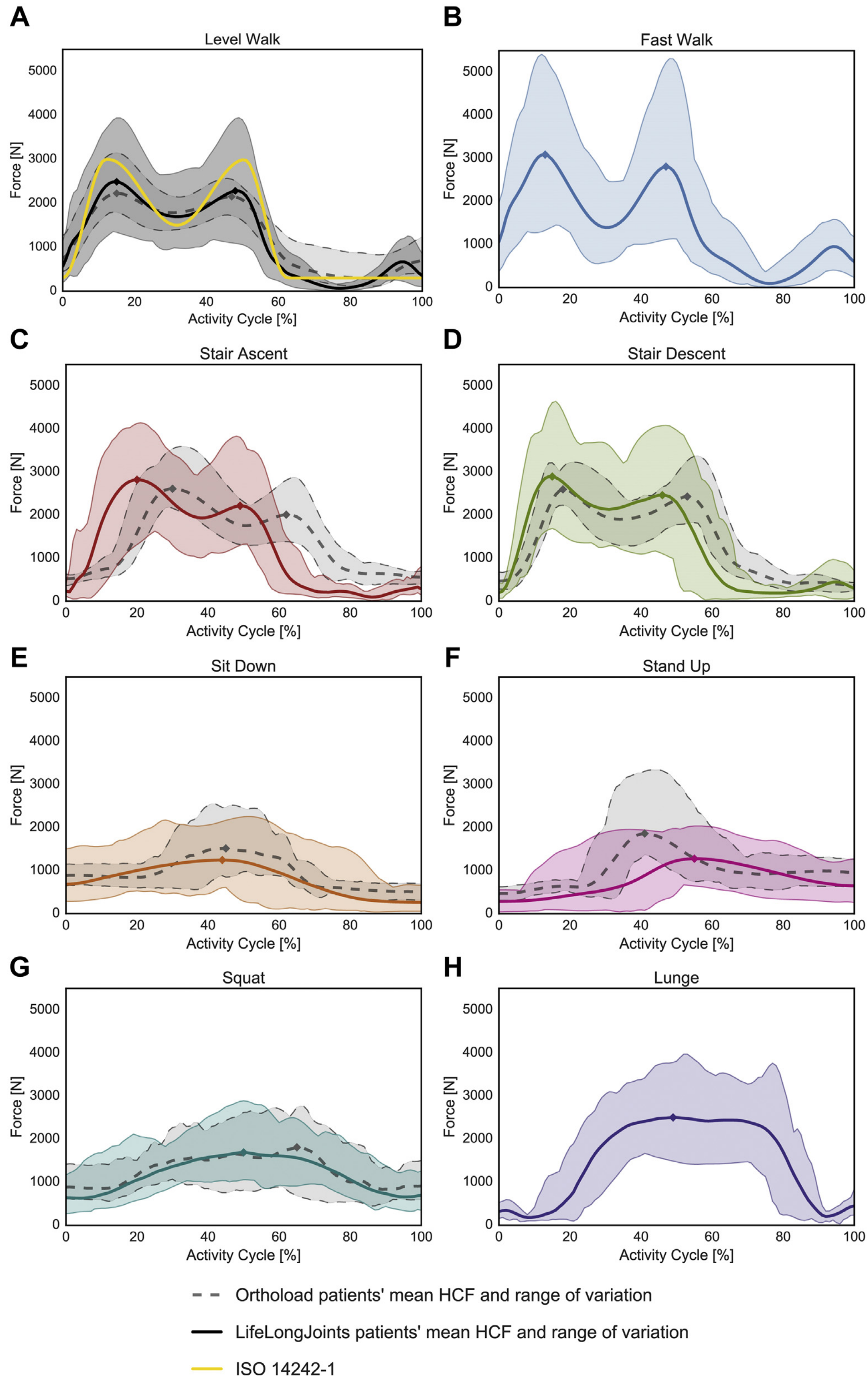

Fig. 2. Predicted resultant hip contact forces (HCFs) across the LifeLongJoints (LLJ) patients' cohort during different activities of daily living: level walking, fast walking, stair ascent, stair descent, sit down, stand up, squat, and lunge. HCFs are reported as mean across the cohort (colored solid line) with relative range of variation (colored shaded area). HCFs for the LLJ patients are compared to the measured HCF from the Orthoload dataset (https://orthoload.com/test-loads/standardized-loads-acting-at-hip-implants/) [24] for all the activities for which a comparison was available. Measured HCFs are also reported as mean (gray dashed line) with relative range of variation (underlying gray shaded area). Peak mean values are highlighted for each mean HCF profile. The axial load defined in the ISO 14242-1 preclinical testing standard is also reported (solid yellow line) for additional comparison against the load variability experienced in vivo by total hip arthroplasty (THA) patients.

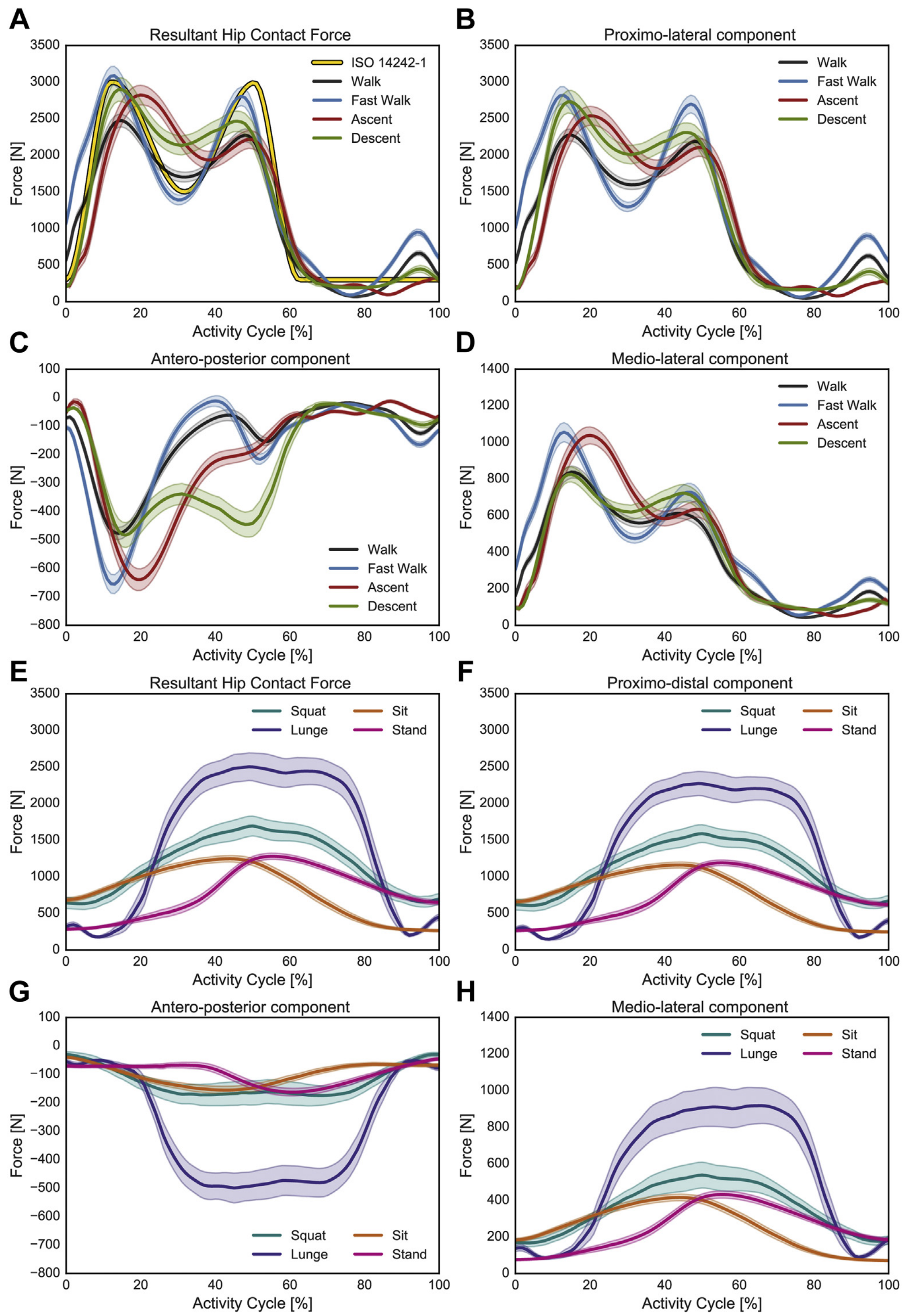

Fig. 3. Predicted HCFs across the patients' cohort during different activities of daily living: level walking, fast walking, stair ascent, stair descent, sit down, stand up, squat, and lunge. Mean resultant HCF and mean individual force components are reported separately in different panels for locomotor and nonlocomotor tasks with relative 95% confidence intervals. The axial load defined in the ISO 14242-1 preclinical testing standard is also reported (solid yellow line) for additional comparison against the load variability experienced in vivo by THA patients.

wear tests does not allow any assessment of the influences that interpatient variability, specifically in terms of loading, would have on the implant performance in vivo.

The analyses of the ADLs have revealed a large range and a large variability in forces experienced at the hip during locomotor and nonlocomotor activities. During the locomotive activities (walk, fast walk, and stair negotiation), there were similarities in the waveform shapes. During the nonlocomotive ADLs (lunge, sit to stand, stand to sit, and squat), the waveform of the resultant force was, as expected, different to the locomotive activities exhibiting a more unimodal and less dynamic loading pattern. Additionally, the individual force components also displayed large differences across activities. Higher posterior loads throughout the weight-bearing phase of the activity characterize stair descent when compared to

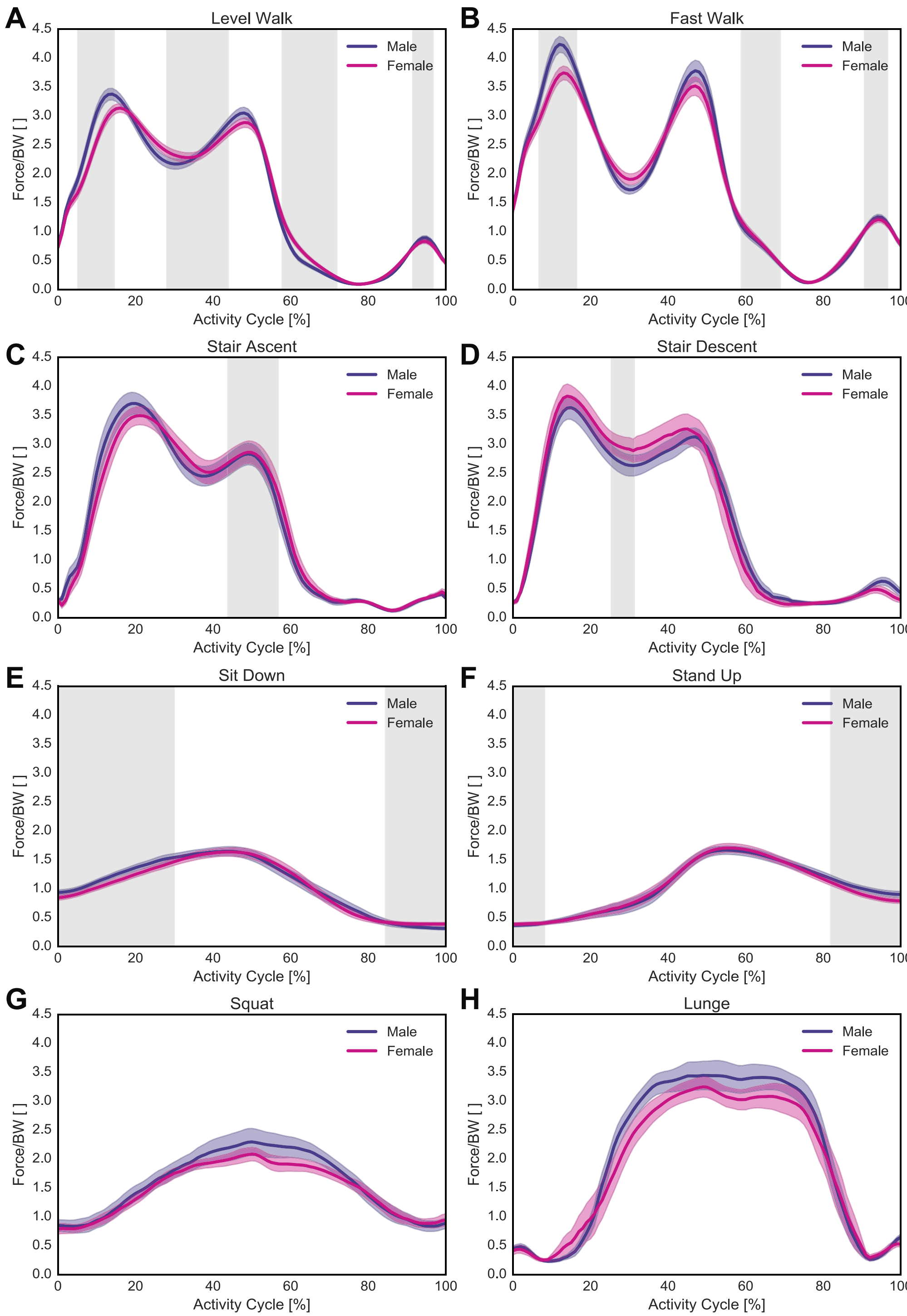

Fig. 4. Predicted resultant HCF across the patients' cohort during different activities of daily living: level walking, fast walking, stair ascent, stair descent, sit down, stand up, squat, and lunge when stratified by gender. Mean resultant HCFs with relative 95% confidence intervals are reported in blue for males and in magenta for females. The temporal spans in which significant differences were observed between the 2 groups are highlighted in the background (shaded gray areas). The statistical analysis was performed considering the 3 individual force components of each group as a vector-field, describing the 3-dimensional variance over time of the HCF vector trajectory. A 2-sample Hotelling's T^2 test, the vectorial analog of a scalar *t*-test [35], was carried out to evaluate the influence of gender on the contact forces. The results of the vector-field analysis are here presented on top of the resultant HCF magnitudes for ease of interpretation.

other locomotive activities, and these would be expected to alter the 3-dimensional loading pattern at the bearing surface and potentially lead to different wear behavior. These differences are profound when comparing individual activities and they could potentially be magnified when considering the much greater variety of activities that the wider THA population engages in [40,41]. Including contrived ADLs or adverse loading conditions in preclinical wear tests has previously produced higher levels of wear [42,43], and the interaction between patient kinematics and surgical factors such as cup placement [44] has demonstrated that wear is a multifactorial phenomenon. The complexity of the interplay between all these factors would be better explored through more extensive testing of implant performance, particularly under more demanding and clinically relevant conditions such as multiple ADLs [23,25].

Our data have shown large patient-specific and activity-related variations in the forces experienced at the hip joint, which differ from the standardized loading waveform currently used in preclinical testing standards, such as ISO 14242-1. Preclinical testing of implants and other orthopedic implants have come under scrutiny lately both from within the industry, with initiatives such as Beyond Compliance (https://www.beyondcompliance.org.uk/), and through external pressures, such as the recent release of articles by the International Consortium of Investigative Journalists (ref https://www.icij.org/investigations/implant-files/). While more evidence is required to confirm whether using representative waveforms would produce more realistic wear patterns compared to retrievals [45], further debate about the suitability of current standards is warranted. Future testing protocols should also consider other in vivo loading conditions not studied in the current cohort such as microseparation [46] edge loading or adverse events [23], which could be incorporated into computational models. In the interim using more realistic loading waveforms such as the ones identified in this work for preclinical hardware simulation would be a progressive step.

To conclude, the LLJ cohort has shown that the testing of hip implants under the current required standard of ISO 14242 does not represent accurately the in vivo loads, even under a limited set of ADLs. There is a case that implant industry could be more demanding in its requirements for preclinical testing before introducing a new implant to market and further work is obviously needed to explore the consequences of the altered loading patterns on wear and ultimately the success or otherwise of a hip arthroplasty. As a first step, the motion-capture dataset underpinning this and related studies is available as a public repository at https://doi.org/10.5518/319, while the associated musculoskeletal models can be obtained through Zenodo.org under the DOI, https://doi.org/10.5281/zenodo.1254286.

Acknowledgments

This study was supported by the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. GA-310477 LifeLongJoints and by the Leeds Experimental Osteoarthritis Treatment Centre which is supported by Arthritis Research UK (grant no. 20083). This research is also supported by the National Institute for Health Research infrastructure at Leeds. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research, or the Department of Health.

Appendix A. Supplementary Data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.arth.2019.10.006.

References

[1] Learmonth ID, Young C, Rorabeck C. The operation of the century: total hip replacement. Lancet 2007;370:1508–19.
[2] Pivec R, Johnson AJ, Mears SC, Mont MA. Hip arthroplasty. Lancet 2012;380:1768–77.
[3] Zagra L. Advances in hip arthroplasty surgery: what is justified? EFORT Open Rev 2017;2:171–8.
[4] Ethgen O, Bruyere O, Richy F, Dardennes C, Reginster JY. Health-related quality of life in total hip and total knee arthroplasty. A qualitative and systematic review of the literature. J Bone Joint Surg Am 2004;86–A:963–74.
[5] Rasanen P, Paavolainen P, Sintonen H, Koivisto AM, Blom M, Ryynanen OP, et al. Effectiveness of hip or knee replacement surgery in terms of quality-adjusted life years and costs. Acta Orthop 2007;78:108–15.
[6] Mariconda M, Galasso O, Costa GG, Recano P, Cerbasi S. Quality of life and functionality after total hip arthroplasty: a long-term follow-up study. BMC Musculoskelet Disord 2011;12:222.
[7] Evans JT, Evans JP, Walker RW, Blom AW, Whitehouse MR, Sayers A. How long does a hip replacement last? A systematic review and meta-analysis of case series and national registry reports with more than 15 years of follow-up. Lancet 2019;393:647–54.
[8] Kurtz S, Ong K, Lau E, Mowat F, Halpern M. Projections of primary and revision hip and knee arthroplasty in the United States from 2005 to 2030. J Bone Joint Surg Am 2007;89:780–5.
[9] Culliford D, Maskell J, Judge A, Cooper C, Prieto-Alhambra D, Arden NK. Future projections of total hip and knee arthroplasty in the UK: results from the UK Clinical Practice Research Datalink. Osteoarthritis Cartilage 2015;23:594–600.
[10] López-López JA, Humphriss RL, Beswick AD, Thom HHZ, Hunt LP, Burston A, et al. Choice of implant combinations in total hip replacement: systematic review and network meta-analysis. BMJ 2017;359:j4651.
[11] Nieuwenhuijse MJ, Nelissen RGHH, Schoones JW, Sedrakyan A. Appraisal of evidence base for introduction of new implants in hip and knee replacement: a systematic review of five widely used device technologies. BMJ 2014;349:g5133.
[12] Hart AJ, Muirhead-Allwood S, Porter M, Matthies A, Ilo K, Maggiore P, et al. Which factors determine the wear rate of large-diameter metal-on-metal hip replacements? Multivariate analysis of two hundred and seventy-six components. J Bone Joint Surg Am 2013;95:678–85.
[13] Hart AJ, Sabah SA, Henckel J, Lloyd G, Skinner JA. Lessons learnt from metal-on-metal hip arthroplasties will lead to safer innovation for all medical devices. Hip Int 2015;25:347–54.
[14] Medley JB. Can physical joint simulators be used to anticipate clinical wear problems of new joint replacement implants prior to market release? Proc Inst Mech Eng Part H 2016;230:347–58.
[15] Langton DJ, Jameson SS, Joyce TJ, Hallab NJ, Natu S, Nargol AV. Early failure of metal-on-metal bearings in hip resurfacing and large-diameter total hip replacement: a consequence of excess wear. J Bone Joint Surg Br 2010;92:38–46.
[16] Bayliss LE, Culliford D, Monk AP, Glyn-Jones S, Prieto-Alhambra D, et al. The effect of patient age at intervention on risk of implant revision after total replacement of the hip or knee: a population-based cohort study. Lancet 2017;389:1424–30.
[17] Towle KM, Monnot AD. An assessment of gender-specific risk of implant revision after primary total hip arthroplasty: a systematic review and meta-analysis. J Arthroplasty 2016;31:2941–8.
[18] Fawsitt CG, Thom HHZ, Hunt LP, Nemes S, Blom AW, Welton NJ, et al. Choice of prosthetic implant combinations in total hip replacement: cost-effectiveness analysis using UK and Swedish hip joint registries data. Value Health 2018;22:303.
[19] Sadoghi P, Liebensteiner M, Agreiter M, Leithner A, Bohler N, Labek G. Revision surgery after total joint arthroplasty: a complication-based analysis using worldwide arthroplasty registers. J Arthroplasty 2013;28:1329–32.
[20] Schmalzried TP, Shepherd EF, Dorey FJ, Jackson WO, dela Rosa M, Fa'vae F, et al. The John Charnley Award. Wear is a function of use, not time. Clin Orthop Relat Res 2000:36–46.
[21] Schmalzried TP, Huk OL. Patient factors and wear in total hip arthroplasty. Clin Orthop Relat Res 2004;418:94–7.
[22] Ardestani MM, Amenabar Edwards PP, Wimmer MA. Prediction of polyethylene wear rates from gait biomechanics and implant positioning in total hip replacement. Clin Orthop Relat Res 2017;475:2027–42.
[23] Fisher J. A stratified approach to pre-clinical tribological evaluation of joint replacements representing a wider range of clinical conditions advancing beyond the current standard. Faraday Discuss 2012;156:59–103.
[24] Bergmann G, Bender A, Dymke J, Duda G, Damm P. Standardized loads acting in hip implants. PLoS One 2016;11:e0155612.
[25] Fabry C, Herrmann S, Kaehler M, Woernle C, Bader R. Generation of physiological movement and loading parameter sets for preclinincal testing of total hip replacements with regard to frequent daily life activities. Bone Joint J Orthop Proc Suppl 2013;95(SUPP 15):194.
[26] De Pieri E, Lund ME, Gopalakrishnan A, Rasmussen KP, Lunn DE, Ferguson SJ. Refining muscle geometry and wrapping in the TLEM 2 model for improved hip contact force prediction. PLoS One 2018;13:e0204109.
[27] De Pieri E, Lunn DE, Chapman GJ, Rasmussen KP, Ferguson SJ, Redmond AC. Patient characteristics affect hip contact forces during gait. Osteoarthritis Cartilage 2019;27:895.
[28] Lunn DE, Chapman GJ, Redmond AC. Hip kinematics and kinetics in total hip replacement patients stratified by age and functional capacity. J Biomech 2019;87:19.
[29] Della Croce U, Bonato P. A novel design for an instrumented stairway. J Biomech 2007;40:702–4.
[30] Mulholland SJ, UP W. Activities of daily living in non-Western cultures: range of motion requirements for hip and knee joint implants. Int J Rehabil Res 2001;24:191–8.
[31] Carbone V, Fluit R, Pellikaan P, van der Krogt MM, Janssen D, Damsgaard M, et al. Tlem 2.0 - a comprehensive musculoskeletal geometry dataset for subject-specific modeling of lower extremity. J Biomech 2015;48:734–41.

[32] Lund ME, Andersen MS, de Zee M, Rasmussen J. Scaling of musculoskeletal models from static and dynamic trials. Int Biomech 2015;2:1–11.
[33] Lund ME, Rasmussen J, Andersen M. AnyPyTools: a Python package for reproducible research with the AnyBody modeling system. J Open Source Softw 2019;4:1108.
[34] Bergmann G, Deuretzbacher G, Heller M, Graichen F, Rohlmann A, Strauss J, et al. Hip contact forces and gait patterns from routine activities. J Biomech 2001;34:859–71.
[35] Pataky TC, Robinson MA, Vanrenterghem J. Vector field statistical analysis of kinematic and force trajectories. J Biomech 2013;46:2394–401.
[36] Andersen MS, Mellon S, Grammatopoulos G, Gill HS. Evaluation of the accuracy of three popular regression equations for hip joint centre estimation using computerised tomography measurements for metal-on-metal hip resurfacing arthroplasty patients. Gait Posture 2013;38:1044–7.
[37] Kepple TM, Sommer HJ, Siegel KL, Stanhope SJ. A three-dimensional musculoskeletal database for the lower extremities. J Biomech 1997;31:77–80.
[38] Ding Z, Tsang CK, Nolte D, Kedgley AE, Bull AM. Improving musculoskeletal model scaling using an anatomical atlas: the importance of gender and anthropometric similarity to quantify joint reaction forces. IEEE Trans Biomed Eng 2019. https://doi.org/10.1109/TBME.2019.2905956.
[39] Foucher KC, Hurwitz DE, Wimmer MA. Relative importance of gait vs. joint positioning on hip contact forces after total hip replacement. J Orthop Res 2009;27:1576–82.
[40] Zietz C, Fabry C, Reinders J, Dammer R, Kretzer JP, Bader R, et al. Wear testing of total hip replacements under severe conditions. Expert Rev Med Devices 2015;12:393–410.
[41] Morlock M, Schneider E, Bluhm A, Vollmer M, Bergmann G, Müller V, et al. Duration and frequency of every day activities in total hip patients. J Biomech 2001;34:873–81.
[42] Bowsher JG, Hussain A, Williams PA, Shelton JC. Metal-on-metal hip simulator study of increased wear particle surface area due to “severe” patient activity. Proc Inst Mech Eng Part H 2006;220:279–87.
[43] Williams S, Jalali-Vahid D, Brockett C, Jin Z, Stone MH, Ingham E, et al. Effect of swing phase load on metal-on-metal hip lubrication, friction and wear. J Biomech 2006;39:2274–81.
[44] Mellon SJ, Grammatopoulos G, Andersen MS, Pegg EC, Pandit HG, Murray DW, et al. Individual motion patterns during gait and sit-to-stand contribute to edge-loading risk in metal-on-metal hip resurfacing. Proc Inst Mech Eng Part H 2013;227(7):799–810.
[45] Walter WL, Insley GM, Walter WK, Tuke MA. Edge loading in third generation alumina ceramic-on-ceramic bearings: stripe wear. J Arthroplasty 2004;19:402–13.
[46] Partridge S, Tipper JL, Al-Hajjar M, Isaac GH, Fisher J, Williams S. Evaluation of a new methodology to simulate damage and wear of polyethylene hip replacements subjected to edge loading in hip simulator testing. J Biomed Mater Res B Appl Biomater 2018;106:1456–62.