Patient Characteristics Affect Hip Contact Forces during Gait

Enrico De Pieri *[1], David E. Lunn *[2,3], Graham J. Chapman [2,3], Kasper P. Rasmussen [4], Stephen J. Ferguson [1] **, Anthony C. Redmond [2,3]

* equally contributed to this work

[1] Institute for Biomechanics, ETH Zurich, Zurich, Switzerland

[2] Institute for Rheumatic and Musculoskeletal Medicine, University of Leeds, Leeds, UK

[3] NIHR Leeds Biomedical Research Centre, Leeds, UK

[4] AnyBody Technology A/S, Aalborg, Denmark

** corresponding author

Email addresses: enrico.depieri@hest.ethz.ch; D.Lunn@leeds.ac.uk; G.J.Chapman@leeds.ac.uk; kasperpihl.r@gmail.com; sferguson@ethz.ch; A.Redmond@leeds.ac.uk

Corresponding author contact details:

Prof. Dr. Stephen J. Ferguson

Address: Hönggerbergring 64, 8093 Zürich, Switzerland

Phone number: +41 44 633 93 30

Email: sferguson@ethz.ch

Running Headline: **Patient characteristics affect HCF**

Abstract

Objective: To examine hip contact force (HCF), calculated through multibody modelling, in a large total hip replacement (THR) cohort stratified by patient characteristics such as BMI, age and function.

Design: 132 THR patients undertook one motion capture session of gait analysis at a self-selected walking speed. HCFs were then calculated using the AnyBody Modelling System. Patients were stratified into three BMI groups, five age groups, and finally three functional groups determined by their self-selected gait speed. Independent 1-dimensional linear regression analyses were performed to separately evaluate the influence of age, BMI and functionality on HCF, by means of statistical parametric mapping (SPM).

Results: The mean predicted HCF were comparable to HCFs measured with an instrumented prosthesis reported in the literature. The regression analyses revealed a statistically significant positive relationship between BMI and HCF, indicating that obese patients are more likely to experience higher HCF during most of the stance phase, while a statistically significant relationship with age was found only during the late swing-phase. Patients with higher functional ability exhibited significantly increased peak contact forces, while patients with lower functional ability displayed a pathological flattening of the typical double hump force profile.

Conclusions: HCFs experienced at the bearing surface are highly dependent on patient characteristics. BMI and functional ability were determined to have the biggest influence on contact force. Current preclinical testing standards do not reflect this.

Keywords: Total hip replacement, Hip contact force, Stratification, Biomechanics, Gait

Introduction

Total hip replacement (THR) surgery is commonly regarded as one of the most successful elective orthopaedic surgeries of the 20th century [1]. It alleviates pain in patients suffering from debilitating hip osteoarthritis and improves function. However there is some lifetime risk of implants requiring revision, the rates of which are currently 4.4% at 10 years and 15% at 20 years[2]. Epidemiological studies have provided evidence to suggest that patient characteristics, such as age, BMI and gender are important factors in the survivorship of hip implants[2, 3]. One in three patients undergoing THR at < 50 years of age are expected to require revision surgery during their lifetime, with risks of one in five for patients 50 to 59 years, one in ten for patients 60 to 69 years, and one in 20 for patients ≥ 70 years [4]. The revision risk for younger patients is consistently higher than for older patients at all time-points i.e. 5, 10, 15 and 20 years and gender also seems to affect risk [2]. Men aged younger than 70 years old have an increased revision risk compared to female patients, and at the age of 50 years females have a 15% lower chance of revision compared to their male counterparts. BMI also contributes to lifetime revision risk, with obese patients having twice the risk of revision at 10 years compared to healthy weight and overweight patients, and it has been suggested by Culliford *et al.* [5] that for every unit increase of BMI, there is a 2% increased risk of revision of a THR.

The precise reason for these differences in revision rates between patient sub-groups is not clear, however the variations in revision rates suggest that the demands placed on the implant likely differ between patient groups. Due to the relatively small sample sizes typically employed in biomechanical studies of THR cases, few studies have explored how patient characteristics can differentially influence function post THR, and ultimately how those characteristics might affect what demand is placed on the implant.

In these few studies age and BMI have been shown to influence function in THR patients. In one analysis of a larger sample of patients from multiple retrospective studies, Foucher *et al.*[6] found that older patients had limited hip sagittal ROM and hip power generation compared to younger patients who recovered better post-operatively. When stratifying gait function by age in a large cohort (n=134) of THR patients, Bennett *et al*.[7, 8] reported that gait kinematics and kinetics were not influenced by age, except for a reduced ROM exhibited in an 80 years and over age group, a finding also consistently observed in healthy control patients of a similar age range[9]. Foucher *et al.* [6 10] reported that BMI plays a role in recovery, with higher BMI patients having a reduced hip range of motion (ROM) and hip abductor moment compared to healthy control participants. Furthermore, lower BMI was associated with higher postoperative values of sagittal ROM, adduction moments, and external rotation moments compared to THR patients with a higher BMI.

As described above, real-world patient function [10] and survivorship of the hip implant [2] is affected by the characteristics of the patient, although this is not currently reflected in preclinical wear testing standards such as ISO 14242. Current preclinical testing protocols use a stylised waveform vaguely representing a 'standard' THR patient's walking cycle to test the wear properties of the implant. A recent study found that post-operative patient function accounts for 42% to 60% of wear, compared to surgical factors which account for 10% to 33% of wear [11], emphasising the importance of understanding how gait varies between different patient groups. No previous studies have tried to understand how patient characteristics affect the absolute forces at the bearing surface, forces which arguably will have the most influence on *in vivo* wear rates. Instrumented implants have been used to calculate contact force at the bearing surface [12, 13], however the data available from these implants is limited to small numbers of patients and extrapolating these data to the wider patient population is not appropriate. Modern computational models of the musculoskeletal system can be used to calculate joint contact forces and are becoming increasingly more clinically applicable [14]. These models have the capability to calculate accurate joint contact forces in THR patients [15], and can be used to predict and compare contact forces in stratified samples derived from a large patient cohort [16]. The primary aim of this study therefore, was to examine hip contact force (HCF), calculated through multibody modelling, in a large THR cohort when stratified by patient characteristics such as BMI, age and function.

Method

Patients

132 THR patients were recruited into the study through a clinical database of surgical cases. Inclusion criteria for the hip replacement group were; between 1-5 years THR post-surgery, older than 18 years of age, no lower limb joint replaced other than hip joint(s), fully pain free and not suffering from any other orthopaedic or neurological problem which may compromise gait. Ethical approval was obtained via the UK national NHS ethics (IRAS) system and all participants provided informed, written consent.

Data Capture

Lower limb kinematics and kinetics were collected using a ten camera Vicon system (Vicon MX, Oxford Metrics, UK) sampling at 100Hz, integrated with two force plates (AMTI, Watertown, MA, USA) capturing at 1000Hz in a 10m walkway. The operated limb (or most recently operated limb, in bilateral cases) was used for analysis. All patients were allowed a familiarisation period prior to completing 3-5 successful trials of each walking condition. A successful trial was defined as a clean foot strike within the boundary of the force plate. The CAST marker set was used to track lower limb segments kinematics in six degrees of freedom, with four non-orthogonal marker clusters positioned over the lateral thighs, lateral shanks and sacrum as described comprehensively elsewhere [17, 18]. Six retroreflective markers were positioned on the first, second and fifth metatarsophalangeal joints as well as the malleoli and calcanei. Participants wore a pair of tight-fitting shorts and a vest onto which reflective markers were affixed using double-sided tape at bony anatomical landmarks to determine anatomical joint centres. Before walking trials commenced, a static trial was collected in an anatomical reference position.

Data Processing

All markers were labelled and gap-filled using the spline fill function in Vicon Nexus 2.5 (Vicon MX, Oxford Metrics, UK), before the labelled marker coordinates and kinetic data were exported to Visual 3D modelling software (C-Motion, Rockville, USA) for further analysis. Kinematic data were filtered using a low-pass (6Hz) Butterworth filter. Ground reaction force (GRF) data were filtered using a low-pass Butterworth filter (25Hz) and heel strike and toe-off were determined using thresholds (>20N for heel strike and <20N for toe off) from the GRF.

Musculoskeletal modelling

Musculoskeletal simulations were performed using commercially available software (AnyBody Modeling System, Version 7.1, Aalborg, Denmark). A recently validated generic musculoskeletal model [19] was scaled to match the anthropometrics of each patient. The scaling of the model segments was based on the marker data collected during a static trial [20]. Marker trajectories and GRF data from each gait trial served as input to an inverse dynamics analysis, based on a 3rd order polynomial muscle recruitment criterion, to calculate muscle forces and HCFs. A total of 494 gait trials were processed and analyzed through the toolkit AnyPyTools (https://github.com/AnyBody-Research-Group/AnyPyTools).

The different components of HCFs, defined in a common femur-based reference frame [12] were computed for the operated limb over a gait cycle. The data were time-normalized from heel-strike (0%), through toe-off (60%), to heel strike (100%) and interpolated to 1% steps (101 points). An average per patient was then calculated based on the 3-5 trials collected.

Stratification by patient characteristics

Patients were stratified by into three groups based on their BMI. BMI scores were calculated as measured weight divided by measured height squared (kg/m^2). The three groups were; healthy weight (BMI ≤25 kg/m^2); overweight (BMI >25kg/m^2 to ≤ 30 kg/m^2) and obese (BMI > 30 kg/m^2)[21]. Patients were also stratified by age into five groups; 1) age 54 to 64 years, 2) 65 to 69 years, 3) 70 to 74 years, 4) 75 to 79, and 5) 80 years and over.

Stratification by functional ability

A widely used alternative measure of overall functional ability is gait speed [22, 23]. There is some negative overall correlation between chronological age and gait speed [24], although age has been shown to only explain 30% of the variance in gait speed [25], suggesting that gait speed itself might be a unique differential indicator of function compared to age. Furthermore in a recent study [26] suggested that patients walking at a higher gait speed is representative of the high functioning patients compared to slower patients who would represent the low functioning patients. Therefore, in the main analysis, in addition to the stratification by age, patients were also stratified into three functional strata determined by their self-selected gait speed. To define the functional strata, the mean and standard deviations (SD) of the gait speeds for the whole cohort were determined. All patients with a gait speed falling within 1SD of the mean were defined as normally functioning (NF).

Patients with a gait speed greater than 1SD above the mean were defined as high functioning (HF), and those with a gait speed more than 1SD below the mean were defined as low functioning (LF).

Data Analysis

Comparisons were made initially between the HCFs derived from the AnyBody model and the measured HCFs from the Bergmann Orthoload literature [12]. This was to compare absolute values and ranges between the two populations and to test the validity of the computational model outputs. Stratified mean peak values and 95% confidence intervals for the resultant force and the three force components are also reported.

Statistical Parametric Mapping (SPM) analysis

The computed HCFs were analysed using Statistical Parametric Mapping [27] (SPM, www.spm1D.org, v0.4, in the Python programming language, www.python.org). Independent linear regression analyses were performed to evaluate the influence of function, age, and BMI on the magnitude of the HCFs, as well as on the individual force components. For each linear regression analysis, the t statistic was computed at each point in the time series, thereby forming the test statistic continuum SPM{t}, technical details are provided elsewhere [28-30]. Significance level was set at $\alpha=0.01$, and the corresponding t^* critical threshold was calculated based on the temporal smoothness of the input data through Random Field Theory. Finally, the probability that similar supra-threshold regions would have occurred from equally smooth random waveforms was calculated. This analysis is based on the assumptions of random sampling and homology of data [30], as well as normality in the data distribution. Adherence to the latter assumption was tested by comparing the above-mentioned parametric linear regression analyses with their non-parametric counterparts [30]. The good agreement between the two types of analysis, in terms of number, temporal extent, and size of the supra-threshold clusters, supports the validity of the assumption of data normality.

The results of the three independent, 1-dimensional linear regression analyses from SPM were further verified by means of 0-dimensional multiple regression analyses. The additional analyses were run in SPSS (IBM SPSS Statistics for Windows, Armonk, NY, USA) at specific time points during the gait cycle, corresponding with the peak loads during stance and the local minimum during mid-stance (15, 32, and 48% of the gait cycle). The force values for the 132 patients at each of these time points, as well as the investigated predictor variables (BMI, age, and gait speed) were normally distributed. Variance inflation factor (VIF) and Tolerance statistics revealed no multi-collinearity in the data, while Durbin-Watson statistics confirmed no autocorrelation between residuals. The assumptions of homoscedasticity and normal distributions of the residuals were also met.

Results

Patient Demographics

132 patients took part in the study and the demographics can be found in Table 1.

- **Insert Table 1 here** -

Musculoskeletal Model Simulations

The predicted contact forces showed comparable trends and values with measured hip contact force data. The mean values were comparable with those in the Orthoload published data and the ranges were generally wider as might be expected from a larger dataset [12] (Figure1 and Table 2).

- **Insert Table 2 and Figure 1 here** -

Peak Hip Contact Forces

Stratified mean peak values for the resultant force and the three force components are reported in full as supplementary data (Supplementary File Table 1).

Statistical Parametric Mapping

The results of the comparator multiple linear regression analyses were in agreement with the outcome of the SPM analysis, confirming a statistically significant positive relationship for both BMI and gait speed with HCF during both the 1st peak and 2nd peak of the stance phase, and a statistically significant positive relationship for BMI and a negative one for gait speed during the mid-stance valley. For the SPM analysis, only differences which were statistically significant for more than 2% of the gait cycle are discussed.

BMI

There was a statistically significant relationship between BMI and the magnitude of the total HCF (Figure 2a). Obese patients demonstrated significantly increased HCF throughout the loaded stance phase (8.8 – 53.8%), mid-swing (74.6 – 79.3%), and terminal swing (88.7 – 100%). All the supra-thresholds clusters exceeded the critical threshold $t^*=3.676$ with associated p-values <0.001, 0.003, and <0.001 respectively.

The same trends were observed for the proximo-distal component (Figure 2b), for which the test statistics similarly exceeded the upper threshold t*=+3.678 at 5.4 – 54.3% (p<0.001), 73.5 – 79.2% (p=0.001), 88.4 – 100% (p<0.001).

In the anteroposterior direction (Figure 2c), statistically significant negative relationship was found during loading response to mid-stance (10.6 – 29.9%), terminal stance (45.4 – 55.3%), and from mid-swing phase (72.2 – 100%). The clusters exceeded the threshold t*= -3.667 with p-values <0.001. No significant difference was observed for the medio-lateral component (Figure 2d).

Age

There was a statistically significant negative relationship between age and the magnitude of the total HCF (Figure 3a), however this was limited to the terminal swing phase (90.7 – 98.7%), with the cluster exceeding the critical threshold t*=-3.660 with p<0.001. This indicates that younger patients are more likely to experience higher contact forces during this phase. The same trend was observed for the proximo-distal component, for which the test statistics similarly exceeded the lower threshold t*=-3.659 at 90.7 – 98.7% of the gait cycle, with an associated p-value <0.001 (Figure 3b), and for the medio-lateral component at 91.8 – 97.7% of the gait cycle (t*=-3.633, p=0.002) (Figure 3d). In the anteroposterior direction, no statistically significant relationship was found (Figure 3c).

Function

The mean gait speed for the functional ability stratum was 0.82 $m.s^{-1}$ (SD; ±0.08), 1.10 $m.s^{-1}$ (±0.09) and 1.37 $m.s^{-1}$ (±0.09) for LF, NF and HF, respectively. There was a statistically significant relationship between functional ability and the magnitude of the total HCF (Figure 4a). Patients with a higher function demonstrated significantly increased HCF during initial contact to loading response (0 – 16% gait cycle), terminal stance to initial swing (43.8 – 74.1%), and terminal swing (87.8 – 100%). A statistically significant negative relationship was instead found during mid-stance (27.9-34.9%). All the supra-threshold clusters exceeded the critical threshold t*=±3.668, with the chances of observing similar clusters in repeated random samplings being p<0.001.

The same trends were observed for the proximo-distal component (the dominant component in terms of magnitude), with the corresponding supra-threshold (t>t*=±3.666) areas spanning from 0 – 15.3%, 45.1 – 73%, 87.7 – 100%, and 27.4 – 35%, respectively (Figure 4b). In the anteroposterior direction, statistically significant negative relationship was found during initial contact to loading response (0.6 – 16.3%) and terminal swing (91.6 – 100%), indicating that higher function demonstrated a significantly increased posterior force during these phases (Figure 4c), while a statistically significant positive relationship was found during mid-stance (27.3 – 45.9%). All the clusters exceeded the critical threshold t*=±3.658 with p-values <0.001. Statistically significant positive relationships were observed for the medio-lateral component during initial contact to leading response (0-19.8%), terminal stance to mid-swing (43.8 – 75.4%), and late swing phase (91.6 – 100%) (Figure 4d).

Discussion

This is the first study to explore the effect of patient characteristics on joint loading through multibody modelling in a large cohort. We found that resultant HCF varies between different patient groups and identified systematic differences between strata for BMI and functional ability. The BMI strata displayed statistically significant differences in the resultant force throughout most of stance phase. Few differences were observed between the age strata, whereas the functional strata, represented by gait speed, displayed the greatest range of statistically significant differences across the time series (over approx. 60% of the whole gait cycle). Patients with a high functionality had increased peak loads during the stance phase of the gait cycle, while low functioning patients displayed a pathological HCF, with a flattening of the typical double hump (Figure 4a). These trends were similar when observing the difference in the proximo-distal component of the HCF, albeit unsurprisingly considering this is the main contributor to the resultant HCF. Our average peak HCF (2449N) was of a similar magnitude to the HCFs measured with instrumented implants by Bergmann *et al.* [12] (2225.7N) (Table 2). No past research has considered the effect of patient characteristics on HCF and comparison to previous literature is difficult. However, previous work has found that joint kinematics and forces acting around the joint are affected by different patient characteristics [6-8] and altered gait variables can affect the magnitude of joint contact forces [31], and therefore this variability in HCF would be expected.

BMI

We found a systematic trend for HCFs to increase with an increasing BMI, and this was expected due to the increase in body mass which has been previously reported to increase linearly with joint contact force [32]. These systematic changes in magnitude are a consistent finding in the literature comparing obese and healthy weight participants when force data are non-normalised, and the differences between BMI groups tend to disappear when normalised to body mass [33], which is common practice in the biomechanical literature exploring function. In our study we specifically chose not to normalise HCF to body weight, as we were interested in the absolute magnitude of the real world forces to which the bearing surface would be exposed. Analysing non-normalised HCFs may help to explain observed BMI dependant revision rates [2], as increased loads in preclinical hardware simulator testing has been shown to increase wear volume and wear particle size [34].

Age

When stratified by age there were very few differences observed in HCF in our patient cohort, with statistically significant differences only found during the terminal swing phase in the proximo-distal and resultant forces (90.7 – 98.7%) and medio-lateral component (91.8 – 97.7%), where the hip is relatively unloaded. Differences in terminal swing phase may be related to the capacity for individuals to energetically drive the limb forward. Compared to the functional strata, the temporal range of significance was much less, indicating that grouping patients by age, as a measure of function, does not differentiate well between patients. No other study has considered the effect of age on HCF measures specifically, however in a gait study using conventional motion capture analysis, Bennett *et al.* [7, 8] observed little kinematic or kinetic differences between age groups in THR patients. As noted previously, the absolute risk of revision in younger patients, can be up to ten times higher than in older patients [2] and it is likely that other factors such as overall activity level in younger patients being higher or younger patients undertaking more demanding adverse loading activities may contribute more than age-related variability in loads during normal walking.

Functional ability

Our results suggest the functional capability of the patient, identified by biomechanical characteristics, best identifies differences between patient groups. When stratifying patients by gait speed, not only were peak forces increased in the HF group, but the waveform in the LF group displayed pathological patterns with a flattening of the transition phase between the two peaks of axial forces (Figure 4a). A trend was also observed in joint contact forces derived at different walking speeds, with the slower walking speeds exhibiting a reduced force during the transition between the peaks [35]. This GRF/HCF waveform has been associated with pathological symptoms in patients with OA or other neurological pathologies [36], suggesting that amongst our patient cohort, all of whom during screening had self-reported as well-functioning, were patients who were indeed pathological, identified by different HCF waveforms. Furthermore, those with higher walking speeds exhibit increased GRFs and joint moments [37], a trend also observed in our HCFs in the function strata. Patient characteristics such as age and BMI are often controlled for in preclinical testing, whereas the real-world functional capability of THR patient is frequently overlooked. Our results suggest that the functional capability of patients could be the most influential factor in determining forces at the bearing surface.

Limitations

Previous work has identified that simulating different activities in preclinical testing also leads to increased wear volume [38]. In the current study we only analysed walking and in reality patients perform a number of other daily tasks which can change the overall loading conditions [39] . Walking is the most commonly performed daily task [40] however, and it is reasonable to suggest that walking would have a clinically relevant impact on implant performance post-surgery. Within the multibody modelling, a number of simulations were run from scaled generic models, and a certain level of error associated with soft-tissue artefacts and the lack of subject-specific bone geometry and muscle physiology information might persist. These models have been previously validated against in-vivo data from different subjects however [14, 15, 19] with good agreement. The overall agreement with the range of measurements from instrumented patients further supports the validity of the current models' predictions.

It could be expected that follow up time could have an effect on patient gait and hip contact force and short-term follow up has shown as much [31, 41]. However, patients were recruited between 1-5 years post operatively in an attempt to avoid abnormalities due to post-surgery recovery and patients mean follow-up time were similar in all groups (Table 1).

Finally, as this study was exploratory in nature we did not analyse any interactions between the strata. It would be expected that there could be some interactions, for example, between age and function [23], which could potentially be more clinically relevant. However the analysis of interactions is not possible in *spm1D* and therefore we decided to keep the focus of the paper on the temporal analysis in the individual strata, as this is relevant for other applications where full waveform data is required, such as preclinical testing.

In conclusion, we have found that the HCF predicted at the bearing surface is highly dependent on the characteristics of the patient. Conversely, current preclinical laboratory testing standards reflect only one loading scenario while our study has shown systematic differences in loading patterns between patient groups (Figures 2-4). To our knowledge these differences are also not considered in any *in-silico* wear prediction models, although more complex waveforms, compared to ISO, have resulted in greater predicted differences wear volume[42, 43]. By extension, if future modelling included patient variability, our data suggest that it is possible that differences in wear rates would also be predicted. We have to accept that failure of an implant is multi factorial and patient factors and surgical factors need to be taken into consideration. However if pre-clinical testing were robust enough to check how implants would perform in different types of patients then patient-dependant failures could potentially be better predicted. Importantly, patient variability is not considered at all in current preclinical hardware simulator testing, which determines whether a device new to market is fit for purpose. It was beyond the realm of this work to test this experimentally in full, but if the loading profiles generated in this study were used in preclinical hardware tests, it would be expected that the variability between patient groups found in this study would also be seen in experimental wear testing [44]. There is certainly a movement towards using different/updated testing procedures with a number of authors suggesting wear testing under more adverse loads is warranted [44]. Improved preclinical testing, both *in silico* and *in vitro*, using more patient stratified waveforms would highlight where and in whom failures are more likely to occur, allowing for better implant design and more informed decision making at the time of THR planning for surgeons. Future work should focus on using patient specific waveforms for *in vitro* testing to check whether the differences observed in this study influence experimental wear rates.

Acknowledgements

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 (NIHR) 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.

We would like to thank AnyBody Technology A/S for all the technical support, and particularly Morten E. Lund, for developing the toolkit *AnyPyTools* and helping with automatizing such a large-scale analysis.

Author contributions

All authors were involved in the conception and design of the study. DEL and EDP performed data acquisition, data processing and analysis. All authors were involved in interpreting the data, revising the manuscript for critically important intellectual content and approved the final version to be submitted.

Role of the 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.

Competing interest statement

The authors have no competing interests to declare

Supplementary data

Supplementary data associated with this article can be found in the online version.

Data associated with this research, in C3d format, can be found at https://doi.org/10.5518/345. This data can be subsequently used with AnyBody Modelling software to calculate joint contact forces. Musculoskeletal models for all trials in the data repository have been implemented with the AnyBody Modelling software and are freely available at Zenodo (DOI: 10.5281/zenodo.1254286)

References

1. Learmonth ID, Young C, Rorabeck C. The operation of the century: total hip replacement. Lancet 2007; 370: 1508-1519.
2. Bayliss LE, Culliford D, Monk AP, Glyn-Jones S, Prieto-Alhambra D, Judge A, 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. The Lancet; 389: 1424-1430.
3. Towle KM, Monnot AD. An Assessment of Gender-Specific Risk of Implant Revision After Primary Total Hip Arthroplasty: A Systematic Review and Meta-analysis. The Journal of Arthroplasty 2016; 31: 2941-2948.
4. Abdel MP, Roth Pv, Harmsen WS, Berry DJ. What is the lifetime risk of revision for patients undergoing total hip arthroplasty? The Bone & Joint Journal 2016; 98-B: 1436-1440.

5. Culliford D, Maskell J, Judge A, Arden NK. A population-based survival analysis describing the association of body mass index on time to revision for total hip and knee replacements: results from the UK general practice research database. BMJ Open 2013; 3.

6. Foucher KC. Identifying clinically meaningful benchmarks for gait improvement after total hip arthroplasty. J Orthop Res 2016; 34: 88-96.

7. Bennett D, Humphreys L, O'Brien S, Kelly C, Orr JF, Beverland DE. Gait kinematics of age-stratified hip replacement patients--a large scale, long-term follow-up study. Gait Posture 2008; 28: 194-200.

8. Bennett D, Ryan P, O'Brien S, Beverland DE. Gait kinetics of total hip replacement patients-A large scale, long-term follow-up study. Gait Posture 2017; 53: 173-178.

9. Nigg BM, Fisher V, Ronsky JL. Gait characteristics as a function of age and gender. Gait & Posture 1994; 2: 213-220.

10. Foucher KC, Freels S. Preoperative factors associated with postoperative gait kinematics and kinetics after total hip arthroplasty. Osteoarthritis and Cartilage 2015; 23: 1685-1694.

11. 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: 2027-2042.

12. Bergmann G, Bender A, Dymke J, Duda G, Damm P. Standardized Loads Acting in Hip Implants. PLOS ONE 2016; 11: e0155612.

13. Bergmann G, Deuretzbacher G, Heller M, Graichen F, Rohlmann A, Strauss J, et al. Hip contact forces and gait patterns from routine activities. Journal of Biomechanics 2001; 34: 859-871.

14. Fregly BJ, Besier TF, Lloyd DG, Delp SL, Banks SA, Pandy MG, et al. Grand challenge competition to predict in vivo knee loads. J Orthop Res 2012; 30: 503-513.

15. Fischer MCM, Eschweiler J, Schick F, Asseln M, Damm P, Radermacher K. Patient-specific musculoskeletal modeling of the hip joint for preoperative planning of total hip arthroplasty: A validation study based on in vivo measurements. PLOS ONE 2018; 13: e0195376.

16. Saxby DJ, Modenese L, Bryant AL, Gerus P, Killen B, Fortin K, et al. Tibiofemoral contact forces during walking, running and sidestepping. Gait & Posture 2016; 49: 78-85.

17. Benedetti MG, Catani F, Leardini A, Pignotti E, Giannini S. Data management in gait analysis for clinical applications. Clinical Biomechanics 1998; 13: 204-215.

18. Cappozzo A, Catani F, Croce UD, Leardini A. Position and orientation in space of bones during movement: anatomical frame definition and determination. Clin Biomech (Bristol, Avon) 1995; 10: 171-178.

19. 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 In Press.

20. Lund ME, de Zee M, Andersen MS, Rasmussen J. On validation of multibody musculoskeletal models. Proc Inst Mech Eng H 2012; 226: 82-94.

21. Organization WH. Obesity and overweight. Fact sheet no. 311. Updated January 2015. World Health Organization.[Cited: 2015 November 20] Available from: http://www. who. int/mediacentre/factsheets/fs311/en 2015.

22. Middleton A, Fritz SL, Lusardi M. Walking speed: the functional vital sign. J Aging Phys Act 2015; 23: 314-322.

23. Studenski S, Perera S, Patel K, et al. Gait speed and survival in older adults. JAMA 2011; 305: 50-58.

24. Bohannon RW, Williams Andrews A. Normal walking speed: a descriptive meta-analysis. Physiotherapy 2011; 97: 182-189.

25. Alcock L, Vanicek N, O'Brien TD. Alterations in gait speed and age do not fully explain the changes in gait mechanics associated with healthy older women. Gait & Posture 2013; 37: 586-592.

26. O'Connor JD, Rutherford M, Bennett D, Hill JC, Beverland DE, Dunne NJ, et al. Long-term hip loading in unilateral total hip replacement patients is no different between limbs or compared to healthy controls at similar walking speeds. Journal of Biomechanics 2018; 80: 8-15.

27. Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RS. Statistical parametric maps in functional imaging: a general linear approach. Human brain mapping 1994; 2: 189-210.

28. Pataky TC. Generalized n-dimensional biomechanical field analysis using statistical parametric mapping. J Biomech 2010; 43: 1976-1982.

29. Pataky TC. One-dimensional statistical parametric mapping in Python. Comput Methods Biomech Biomed Engin 2012; 15: 295-301.

30. Pataky TC, Vanrenterghem J, Robinson MA. Zero- vs. one-dimensional, parametric vs. non-parametric, and confidence interval vs. hypothesis testing procedures in one-dimensional biomechanical trajectory analysis. J Biomech 2015; 48: 1277-1285.

31. Wesseling M, de Groote F, Meyer C, Corten K, Simon JP, Desloovere K, et al. Gait alterations to effectively reduce hip contact forces. J Orthop Res 2015; 33: 1094-1102.

32. Sanford BA, Williams JL, Zucker-Levin AR, Mihalko WM. Hip, Knee, and Ankle Joint Forces in Healthy Weight, Overweight, and Obese Individuals During Walking. In: Doyle B, Miller K, Wittek A, Nielsen PMF Eds. Computational Biomechanics for Medicine. New York, NY: Springer New York 2014:101-111.

33. Lerner ZF, Browning RC. Compressive and shear hip joint contact forces are affected by pediatric obesity during walking. Journal of Biomechanics 2016; 49: 1547-1553.

34. 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. Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine 2006; 220: 279-287.

35. Georgios G, Ilse J, Mariska W, Sam VR, Sabine V. Loading of Hip Measured by Hip Contact Forces at Different Speeds of Walking and Running. Journal of Bone and Mineral Research 2015; 30: 1431-1440.

36. Perry J, Davids JR. Gait analysis: normal and pathological function. Journal of Pediatric Orthopaedics 1992; 12: 815.

37. Ardestani MM, Ferrigno C, Moazen M, Wimmer MA. From normal to fast walking: Impact of cadence and stride length on lower extremity joint moments. Gait & Posture 2016; 46: 118-125.

38. 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 Journal Orthopaedic Proceedings Supplement 2013; 95: 194-194.

39. Varady PA, Glitsch U, Augat P. Loads in the hip joint during physically demanding occupational tasks: A motion analysis study. Journal of Biomechanics 2015; 48: 3227-3233.

40. 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. Journal of Biomechanics 2001; 34: 873-881.

41. Colgan G, Walsh M, Bennett D, Rice J, O'Brien T. Gait analysis and hip extensor function early post total hip replacement. Journal of Orthopaedics 2016; 13: 171-176.

42. Liu F, Fisher J, Jin Z. Effect of motion inputs on the wear prediction of artificial hip joints. Tribology International 2013; 63: 105-114.

43. Gao L, Wang F, Yang P, Jin Z. Effect of 3D physiological loading and motion on elastohydrodynamic lubrication of metal-on-metal total hip replacements. Med Eng Phys 2009; 31: 720-729.

44. Zietz C, Fabry C, Reinders J, Dammer R, Kretzer JP, Bader R, et al. Wear testing of total hip replacements under severe conditions. Expert Review of Medical Devices 2015; 12: 393-410.

Figure Legends

Figure 1. Predicted HCF across the patients' cohort compared to the measured HCF from the Orthoload dataset (https://orthoload.com/test-loads/standardized-loads-acting-at-hip-implants/) [12]. Resultant force (blue) and single components – proximo-distal (red), antero-posterior (orange), medio-lateral (green) – are reported as mean across the cohort (solid line) and overall range of variation (shaded area) and compared to the corresponding mean and range of variations from the Orthoload measurements (in grey). Peak values reported in Table 2 are indicated in each plot.

Figure 2. Predicted hip contact forces across patients reported as a) resultant magnitude, and individual components: b) proximo-distal, c) antero-posterior, and d) medio-lateral component. The patients were stratified in *Healthy Weight* (blue), *Overweight* (purple) and *Obese* (red) according to

their BMI score. The upper panels report the averages for each patient strata (solid line) and their relative 95% confidence intervals. Additionally, the loading profile from the ISO14242-1 testing standard (dashed grey line) is compared to the proximo-distal forces for each group. The corresponding lower panels report the results of the SPM linear regression analysis. The significance α-level was set to 0.01 for each analysis and the corresponding threshold t* are reported (horizontal dashed lines). Whenever the test statistics continuum SPM{t} exceeds the threshold, significance is reached and the p-values associated with the supra-threshold clusters (shaded grey areas) are reported.

Figure 3. Predicted hip contact forces across patients reported as a) resultant magnitude, and individual components: b) proximo-distal, c) antero-posterior, and d) medio-lateral component. The patients were stratified according to their age in five groups: 54:64 (orange), 65:69 (red), 70:74 (grey), 75:79 (blue) and ≥80 (green). The upper panels report the averages for each patient strata (solid line) and their relative 95% confidence intervals. Additionally, the loading profile from the ISO14242-1 testing standard (dashed grey line) is compared to the proximo-distal forces for each group. The corresponding lower panels report the results of the SPM linear regression analysis. The significance α-level was set to 0.01 for each analysis and the corresponding threshold t* are reported (horizontal dashed lines). Whenever the test statistics continuum SPM{t} exceeds the threshold, significance is reached and the p-values associated with the supra-threshold clusters (shaded grey areas) are reported.

Figure 4. Predicted hip contact forces across patients reported as a) resultant magnitude, and individual components: b) proximo-distal, c) antero-posterior, and d) medio-lateral component. The patients were stratified in Low Functioning (purple), Normal Functioning (blue) and High Functioning (green) according to their self-selected gait speed. The upper panels report the averages for each

patient strata (solid line) and their relative 95% confidence intervals. Additionally, the loading profile from the ISO14242-1 testing standard (dashed grey line) is compared to the proximo-distal forces for each group. The corresponding lower panels report the results of the SPM linear regression analysis. The significance α-level was set to 0.01 for each analysis and the corresponding threshold t* are reported (horizontal dashed lines). Whenever the test statistics continuum SPM{t} exceeds the threshold, significance is reached and the p-values associated with the supra-threshold clusters (shaded grey areas) are reported.

Table 1. Patient demographics for each classification strata. Values are reported as mean (SD) unless otherwise stated.

		Number of patients	Female:Male	Age (Years)	BMI (kg/m^2)	Post-surgery (Years)
All		132	66:66	71.6 (7.6)	28.2(3.8)	2.8 (1.4)
BMI	Healthy Weight	29	18:11	70.1(8.2)	23.4(1.2)	2.6(1.2)
	Overweight	67	31:36	73.2(7.2)	27.6(1.3)	2.8(1.4)
	Obese	36	17:20	69.7(7.0)	33.2(2.2)	3.0(1.6)
Age	54-64	22	11:11	60.4 (2.9)	28.5(5.3)	2.9(1.5)
	65-69	37	17:20	67.0(1.4)	28.9(3.4)	2.8(1.6)
	70-74	23	14:9	72.3(1.0)	27.8(4.2)	2.1(1.1)
	75-79	28	14:14	77.4(1.2)	28.2(3.0)	2.7(1.3)
	>=80	22	10:12	82.4(3.0)	27.1(2.7)	3.0(1.5)
Function	HF	18	7:11	69.3(6.1)	27.1(2.8)	3.6(1.4)
	NF	97	48:49	71.3(7.7)	28.2(3.8)	2.7(1.4)
	LF	17	11:6	75.8(6.3)	29.3(4.4)	2.7(1.2)

Table 2. A comparison of measured peak contact forces [12] and the calculated peak contact forces form our study. Values are reported as mean and ranges (min-max). The reported values are highlighted in the corresponding graphs in Figure 1.

Dataset	Peak resultant force 1st peak (R1) (min-max range)	Peak resultant force 2nd peak (R2) (min-max range)	Peak Proximal/Distal force 1st peak (PD1) (min-max range)	Peak Proximal/Distal force 2nd peak (PD2) (min-max range)	Peak posterior force(P1) (min-max range)	Peak Anterior forces (A1) (min-max range)	Peak Medial/Lateral force 1st peak (ML1) (min-max range)	Peak Medial/Lateral force 2nd peak (ML2) (min-max range)
LLJ dataset	2449.1 (1310.9 , 3913.5)	2279.0 (1093.8 , 3920.5)	2254.3 (1179.8 , 3694.4)	2197.3 (1030.8 , 3849.1)	-466.1 (-838.0 , -232.9)	-60.5 (-365.3 , 297.2)	826.0 (459.4 , 1353.5)	599.0 (273.2 , 1063.3)
Orthoload	2225.7 (1793.4 , 3147.0)	2149.9 (1721.2 , 2546.8)	2085.8 (1670.1 , 3006.5)	2073.6 (1643.8 , 2475.2)	-405.7 (-650.4 , -111.4)	23.5 (-193.0 , 211.7)	641.3 (366.7 , 819.5)	600.0 (341.1 , 807.2)