LEEDS BECKETT UNIVERSITY

Exploring the Measurements Following Hip Arthroscopy and Dynamometry Testing of the Hip

MOHAMED A. HUSAIN

A thesis submitted in partial fulfilment of the requirements of Leeds Beckett University for the degree of Doctor of Philosophy

December 2017

Abstract

Background: Medium-term longitudinal studies considering postoperative hip arthroscopy (HA) changes are lacking. Objectively measuring hip muscle strength changes using a hand-held dynamometer (HHD) is essential following HA. However, HHD can be affected by the method and location of stabilisation. **Purpose:** (1) To investigate the clinical changes in MS*, ROM†, MHHS‡, and VAS§ following HA and the factors that affect these changes. (2) To evaluate the validity and reliability of two hip strength assessment platforms (GrAASP and STAPH**). (3) To determine the best location and method of stabilisation of a HHD. **Methods:** A secondary dataset of 316 HA patients (age: 41.4±14.4 years) was analysed. Two samples of healthy active participants (age: 22.9±4.4 and 19.7±1.7 years) were assessed to evaluate the limits of agreement (LoA) between GrAASP and an isokinetic device (IKD), and test-retest agreement in scores using GrAASP and STAPH. **Results:** The biggest improvements in hip MS, MHHS and VAS were between preoperative and 24 weeks postoperatively (all $p<0.001$). No change in hip ROM was found ($p=0.569$). Age and sex negatively affect postoperative measurements and athletes respond better to postoperative rehabilitation. The narrowest LoA tests were between a distally fixed method and IKD in 0° and 30° of abduction for hip abduction (HAB) and adduction (HAD) isometric testing. The narrowest test-retest LoA were found using a distally fixed method in HAB and HAD testing. The narrower test-retest LoA were obtained using the STAPH compared to the GrAASP in HAB and HAD testing. **Conclusions:** Clinicians should expect at least 8 weeks to observe significantly favourable changes in MS, MHHS and VAS postoperatively. The magnitude of clinical changes varies based on the patients' characteristics. Age, sex and athletic activity should be considered in the rehabilitation protocols and further studies should be performed on HA patients. The HHD should be stabilised distally to obtain more consistent results. **Limitations:** The secondary data had a lot of missing values which may have influenced the results. Several sources of error may have affected the data collection. The results of the second study are limited to young, healthy, and active populations.

---

* Muscle strength
† Range of motion: measured using a plastic goniometer.
‡ Modified Harris Hip Score: a questionnaire to measure function and pain.
§ Visual Analogue Scale: a 10-cm line to measure pain.
** The Groin Abductor-Adductor Strength Platform (GrAASP) and the Hip Strength Assessment Platform (STAPH).

Acknowledgments

I would like to express my deepest gratitude and thanks to my supervisors Dr Claire Griffiths, Prof Carlton Cooke and Prof Ernest Schilders for their continuous support, guidance and help. You have facilitated a smooth transfer in my career from a clinician to a researcher. It has been delightful to work with you, and I am looking forward to future collaboration.

Besides my supervisory team, I would like to thank my PhD colleagues and friends: Dr Jamie Matu, Dr James Nobles, Dr Katie Pickering, Dr Lorena Lozano, Dr Oliver Shannon and Antonis Elia, for their feedback and being there when I needed the support. I will miss our Yorkshire tea breaks and nights out.

To mum and dad, I would not have made it to this stage without your endorsement in every aspect. You are the reason for my success, and I will ever be grateful to you. To my sister and friend, you made my first year in Leeds an unforgettable memory. I still remember our frequent trips to London and York.

To Prof Mona Al-Ansari, former dean of the College of Physical Education and Physiotherapy at the University of Bahrain, I can't thank you enough for this opportunity. I hope I have made you proud.

Lastly, I would like to thank the Carnegie School of Sport administration team for their assistance and keeping up with my requests, and to the University Research Office for making this an easier journey than I thought.

Thank you all for everything.

Declaration

I hereby confirm that this thesis is my own work. The studies are almost entirely my own work unless stated otherwise. The collaborative contributions have been indicated clearly and acknowledged. Due references have been provided on all supporting literatures and resources.

I confirm that the thesis, or any part thereof, has not been previously submitted for any degree or a comparable academic award.

Table of Contents

Abstract ii

Acknowledgments iii

Table of Contents iv

List of Tables viii

List of Figure xii

List of Abbreviations xxiii

Chapter 1 Introduction 1

1.1. Hip Pain, Arthroscopy and Muscle Strength Assessment 1

1.2. Framework of the Thesis 2

Chapter 2 Literature Review 9

2.1. Hip Joint Anatomy 9

2.2. Anatomical Variations Based on Sex 12

2.3. Intra-articular Hip Pain 17

2.4. Hip Arthroscopy Procedures 18

2.5. Outcomes of Hip Arthroscopy Procedures 20

2.6. Hand-Held Dynamometry 55

2.7. Stabilisation and Strength Assessment Platforms (SAP) 72

2.8. Literature Review Key Points 80

Chapter 3 Study One: Hip Arthroscopy Data Secondary Analysis 87

3.1. Chapter Overview 87

3.2. Description of the Data 87

3.3. Research Significance, Contribution and Questions 89

3.4. Ethical Consideration 91

3.5. Methodology 92

3.6. Results 101

3.7. Discussion 114

3.8. Study Limitations 129

3.9. Conclusion 132

Chapter 4 Designing Process of the Assessment Platforms 135

4.1. Chapter Overview 135

4.2. Significance of the Platforms 135

4.3. Idea Development and Stakeholders 136

4.4. Development of the Platforms 137

4.5. Development of the Dynamometer Support 144

Chapter 5 Study Two: Methodology 149

5.1. Chapter Overview 149

5.2. Research Significance, Contribution and Questions 149

5.3. Ethical Considerations and Participation 153

5.4. Measurement Tools 156

5.5. General Testing Protocol 157

5.6. Phase One Testing Protocol 161

5.7. Phase Two Testing Protocol 168

5.8. Statistical Analysis 174

Chapter 6 Study Two: Results 180

6.1. Chapter Overview 180

6.2. Phase One Results 180

6.3. Phase Two Results 210

6.4. Combined Results of Phases One and Two of Testing Using the MicroFET2 in Testing the Hip Abductor and Adductor Muscles (RQ 9) 224

Chapter 7 Study Two: Discussion 229

7.1. Chapter Overview 229

7.2. Bilateral Torque Differences 229

7.3. Sex Differences 234

7.4. Angles and Speeds Differences 236

7.5. Muscle Strength Ratios 239

7.6. Agreement Between the Biodex and the MicroFET2 in Phase one 244

7.7. Agreement Between the MicroFET2 Sessions in Phase One 247

7.8. Agreement Between the MicroFET2 Sessions in Phase Two 249

7.9. Comparison Between the Two Assessment Platforms 251

7.10. Study Limitations 254

Chapter 8 Conclusion 263

8.1. Overview and Prospective Research 263

8.2. Novel Contribution and Main Outcomes 267

8.3. Overall Thesis Limitations 271

8.4. Recommendations and Clinical Applications 273

Appendix A Chapter 2 Supporting Material 277

Appendix B Chapter 3 Supporting Material 293

Appendix C Chapter 4 Supporting Material 301

Appendix D Chapter 5 Supporting Material 307

Appendix E Chapter 6 Supporting Material 326

Appendix F Chapter 7 Supporting Material 344

Appendix G Supplementary Material 347

References 356

List of Tables

Table 2.1 - Overview of the studies that provided normative values of lower extremities' muscle strength..... 15

Table 2.2 - Indications for hip arthroscopy in a sample of 47 patients (Byrd and Jones, 2001) 20

Table 2.3 – Overview of the studies investigating the changes in muscle strength following hip arthroscopy procedures. 27

Table 2.4 – Overview of the studies investigating the changes in joint range of motion following hip arthroscopy procedures. 29

Table 2.5 - Overview of the studies investigating mainly the changes in function, return to sport and pain following hip arthroscopy procedures. 34

Table 2.6 - Outerbridge classification (BlueCross, 2016)..... 45

Table 2.7 - Overview of the studies investigating the factors affecting the outcomes of hip arthroscopy procedures. 48

Table 2.8 - HHD test-retest (multiple day intra-rater) and inter-rater reliability (n=18) (Arnold, et al., 2010)..... 58

Table 2.9 - The relationship and differences between ID measurements (Biodex 3) and HHD measurements adjusted for leg length (n=17) (Arnold, et al., 2010)..... 58

Table 2.10 - The relative and absolute reliability of lower extremity strength measured by inter-rater reliability (Lu, et al., 2011) 61

Table 2.11 - Overview of the studies investigating the validity and reliability of mainly lower extremities muscle strength testing using HHDs. 67

Table 2.12 - Overview of the studies investigating the differences of lever arm length in HHD testing. 71

Table 2.13 - Hip abductor torque (N·m) for the elderly and young patients in the first session (Kramer, et al., 1991). 72

Table 2.14 – Intra-class correlation coefficient of isometric muscle strength measurements using a HHD with and without a belt (Katoh & Yamasaki, 2009). 74
Table 2.15 - Overview of the studies investigating the types of stabilisation in HHD testing of the lower extremities muscles. 82
Table 2.16 - Overview of the studies that were performed using stabilising platforms or frames in lower extremities muscle strength testing. 84
Table 3.1 - Variables in the hip arthroscopy dataset (Count = 67) 88
Table 3.2 – The models used in multilevel modelling. 99
Table 3.3 - Characteristics of the hip arthroscopy patients (1) 101
Table 3.4 - Characteristics of the hip arthroscopy patients (2) 102
Table 3.5 - Repeated measures ANOVA of strength and range of motion comparing between weeks of measurement. 108
Table 3.6 - Repeated measures ANOVA of other variables comparing between weeks of measurement. 108
Table 3.7 - Summary of the effects of factors on strength, range of motion, function and pain postoperatively (based on Model 5). 111
Table 3.8 - Results (1) of multilevel modelling based on Model 5 showing the regression coefficients (and standard error) to estimate trends in means of measurement scores. 112
Table 3.9 - Results (2) of multilevel modelling based on Model 5 showing the regression coefficients (and standard error) to estimate trends in means of measurement scores. 113
Table 5.1 - The models used in multilevel modelling. 178
Table 6.1 - Number of complete and partial datasets in phase one. 181
Table 6.2 - Baseline characteristics of the participants in phase one. 181
Table 6.3 - Sports played by the participants in phase one. 182
Table 6.4 - Repeated measures ANOVA test results comparing between male and female measurements in phase one testing (dominant side only). 188

Table 6.5 - Repeated measures ANOVA test results comparing between male and female measurements in phase one testing (non-dominant side only). 189
Table 6.6 - Repeated measures ANOVA test results comparing between the hip angles of abduction in isometric testing and speeds in isokinetic testing of phase one testing. 190
Table 6.7 - Ratios of hip muscle strength obtained by the Biodex and MicroFET2 in phase one (dominant side only). 191
Table 6.8 – Limits of agreement tests between hip abduction measurements of the Biodex and the 1st MicroFET2 testing session in phase one (dominant side only).... 195
Table 6.9 - Limits of agreement tests between hip adduction measurements of the Biodex and the 1st MicroFET2 testing session in phase one (dominant side only).... 200
Table 6.10 - Pearson Correlation (r) between measurements of the Biodex and the 1st MicroFET2 session in phase one (dominant side only). 201
Table 6.11 - Limits of agreement tests between the results of the two MicroFET2 sessions in phase one (dominant side only). 205
Table 6.12 – The Coefficient of Variation between the two MicroFET2 sessions in phase one (dominant side only). 205
Table 6.13 - Summary of the effects of factors on abduction and adduction strength measurements of the second MicroFET2 session in phase one (based on Model 3). 207
Table 6.14 – Results of the sensitivity analysis using multilevel modelling based on Model 3 showing the regression coefficients (and standard error) to estimate trends in means of measurement scores between the two dynamometer sessions. 209
Table 6.15 - Baseline characteristics of the participants in phase two. 210
Table 6.16 - Sports played by the participants in phase two..... 211
Table 6.17 - Repeated measures ANOVA test results comparing between male and female measurements in phase two (Dominant side only). 216
Table 6.18 - Repeated measures ANOVA test results comparing between male and female measurements in phase two (Non-dominant side only). 217

Table 6.19 - Ratios of muscle strength obtained by the MicroFET2 in phase two (dominant side only). 218
Table 6.20 - Limits of agreement tests between the results of the two MicroFET2 testing sessions of phase two (dominant side only). 222
Table 6.21 – The coefficient of variation between the two MicroFET2 sessions in phase two (dominant side only). 222
Table 6.22 - Baseline characteristics of the participants in phases one and two. 225
Table 6.23 - Limits of agreement tests between of the measurement results obtained by the GrAASP [phase one] and the STAPH [phase two] (dominant side only). 228

List of Figures

Figure 1.1 – Framework of the thesis (HHD; hand-held dynamometry; green dashed boxes refer to the gap in literature and rationale of the studies; orange box refer to study one and the blue boxes refer to the chapters related to study two. 8

Figure 2.1 - The right hip joint showing the ligaments (Neumann, 2009b). 9

Figure 2.2 –Side view of the pelvis showing the acetabulum and labrum (Left); Ligaments of the hip joint (Right) (Retchford and Kemp 2013). 10

Figure 2.3 - Anatomical differences between the male and female pelvises (OpenStax, 2015). 12

Figure 2.4 - Labral tear (left) and multiple loose bodies (right) as observed during hip arthroscopy procedures (Wall, et al., 2012). 17

Figure 2.5 - Cam impingement (left) and pincer impingement (right) (Byrd, 2010)..... 18

Figure 2.6 - Portals of the central compartment of the hip (Dienst, 2005). 19

Figure 2.7 - Patrick's (FABER) test (Biggs, 2010)..... 25

Figure 2.8 – The factors affecting post-surgical outcomes (Green and red boxes refer to the factors resulting in good and worse outcomes respectively; BMI, body mass index; HA, hip arthroscopy; CHP, chondropathy; OA, osteoarthritis). 41

Figure 2.9 - Hip extension torque values (*, significant difference between examiner 1 and examiner 2 ($p < 0.001$) and examiner 3 ($p < 0.001$) (Krause et al., 2014)..... 64

Figure 2.10 - Hip external rotation torque values (*, significant difference between examiner 1 and examiner 3 in short-lever (seated) test only, $p = 0.001$) (Krause et al., 2014). 64

Figure 2.11 - Hip abduction torque values (*, significant difference between examiner 1 and examiner 3, $p = 0.001$) (Krause et al., 2014). 65

Figure 2.12 – A summary of the factors affecting hand-held dynamometry testing..... 66

Figure 2.13 - Strength measurement of the right extensors (Top) and the hip abductors (bottom) (Nadler et al., 2000) 75

Figure 2.14 – A participant positioned for abductor testing (Widler, et al., 2009). 75

Figure 2.15 - Hip abductor strength and gluteus medius EMG (mean and standard deviation (A) Maximal voluntary contraction (MVC) strength (normalized to body weight; *, $p < 0.05$; ***$p < 0.001$) (B) Contralateral-to-tested EMG ratio during maximal isometric hip testing. (**$p < 0.01$) (Widler, et al., 2009)..... 76

Figure 2.16 - The stabilising platform (frame) by Traina et al. (2010). 77

Figure 2.17 – A participant positioned for hip abduction/adduction strength testing (Tourville, et al., 2013). 78

Figure 2.18 - Hip abduction testing using the PVC pipe stabilisation device (Jackson, et al., 2016). 79

Figure 2.19 - Sequence of the stabilising platforms listed according to the authors and the muscle groups of interest (HFL, hip flexion; HEX, hip extension; HAB, hip abduction; HAD, hip adduction; HER, hip external rotation; KEX, knee extension; APF, ankle planter flexion). 80

Figure 3.1 - General design of the study showing the major steps from data collection to conclusion (Red boxes refer to the steps performed by Professor Schilders and his staff with no further information; Green boxes refer to the steps performed by the researcher; ADL, activities of daily living; DoS, director of studies; LREC, local research ethics coordinator; *, will be explained in a separate flow chart)..... 92

Figure 3.2 - Changes to the dataset after exclusion (red boxes) and dividing patients based on the operation side (black boxes)..... 95

Figure 3.3 - Structure of the repeated measures ANOVA analysis model (Preop, preoperative; Rt, right; Lt, left)..... 97

Figure 3.4 - Structure of the multilevel modelling analysis (Preop, preoperative; W, weeks; green boxes represent level 1 and the blue boxes represent level 2 of the model). 99

Figure 3.5 - Statistical tests used to analyse the secondary data (ANOVA, analysis of variance; MLM, multilevel modelling; orange boxes represent aim of the test reflecting the research questions; blue boxes represent the statistical tests used). 100

Figure 3.6 - Hip abductor muscle strength changes pre- and post-operatively based on the RM-ANOVA results and operation side (n=54) (With standard deviation bars and percentage of change from pre-operative measurements; *, statistically significant difference at 0.05 from pre-operative measurements). 103
Figure 3.7 - Hip adductor muscles strength changes pre- and post-operatively based on the RM-ANOVA results and operation side (n=53) (With standard deviation bars and percentage of change from pre-operative measurements; *, statistically significant difference at 0.05 from pre-operative measurements). 103
Figure 3.8 - Hip external rotation range of motion changes pre- and post-operatively based on the RM-ANOVA results and operation side (n=51) (With standard deviation bars and percentage of change from pre-operative measurements). 104
Figure 3.9 - Hip internal rotation range of motion changes pre- and post-operatively based on the RM-ANOVA results and operation side (n=51) (With standard deviation bars and percentage of change from pre-operative measurements). 105
Figure 3.10 - The Flexion-Abduction-External Rotation test distance changes pre- and post-operatively based on the RM-ANOVA results and operation side (n=60) (With standard deviation bars and percentage of change from pre-operative measurements; *, statistically significant difference at 0.05 from pre-operative measurements). 105
Figure 3.11 - The Visual Analogue Scale changes pre- and post-operatively based on the RM-ANOVA results (n=49) (With standard deviation bars and percentage of change from pre-operative measurements; VAS, visual analogue scale*, statistically significant difference at 0.05 from pre-operative measurements). 107
Figure 3.12 - The Modified Harris Hip score changes pre- and post-operatively based on the RM-ANOVA results (n=27) (With standard deviation bars and percentage of change from pre-operative measurements; *, statistically significant difference at 0.05 from pre-operative measurements). 107
Figure 4.1 – Drawing of the board version P1.0 (top view) 137
Figure 4.2 – Drawing of the board version P2.0 (top view). 138
Figure 4.3 – Cardboard model of version P2.1. 139

Figure 4.4 - Cardboard model of version P2.2. 140

Figure 4.5 - 3D SolidWorks drawing of version 2.3. 140

Figure 4.6 – The (GrAASP) showing the stabilising straps, head and back sponge support and the dynamometer support. 141

Figure 4.7 – A participant positioned to have his left hip tested. The left hip is aligned over the intersections between the horizontal and hip lines. 142

Figure 4.8 - The hip strength assessment platform (STAPH) – version P3.0..... 143

Figure 4.9 - A participant aligned in the supine, side lying and sitting positions to test hip abductors (left), extensors (centre) and internal rotators (right) consecutively (yellow arrows point at the direction of push). 144

Figure 4.10 - Concept 1 of the dynamometer support – version S1.0..... 144

Figure 4.11 – Version S2.0 of the dynamometer support. 145

Figure 4.12 – Version S2.1 of the dynamometer support. 145

Figure 4.13 – Version S3.0 of the dynamometer support. 145

Figure 4.14 - Stabilising the support system during abduction testing. 146

Figure 4.15 - The difference (yellow arrow) between the starting position based on 10° of abduction (green line) and the actual testing position (red line)..... 146

Figure 4.16 - Version S3.1 of the dynamometer support. 146

Figure 4.17 - The MicroFET2 support box (version 3.2) showing the mould and the brackets..... 147

Figure 4.18 – A summary of the platform designs and concepts (GrAASP, groin abductor-adductor strength platform; STAPH, hip strength assessment platform; red box represents concept 1; blue boxes represent concept 2; green box represents concept 3; P, platform). 147

Figure 4.19 – A summary of the stabilising system designs and concepts (GrAASP, groin abductor-adductor strength platform; STAPH, hip strength assessment platform; red box represents concept 1; blue boxes represent concept 2; green boxes represent concept 3; S, support). 148

Figure 5.1 – The MicroFET2 hand-held dynamometer (left) and the Biodex System 4 Pro (right). 156
Figure 5.2 - Stature measurement using a stadiometer. 157
Figure 5.3 - Waist circumference measurement. 157
Figure 5.4 – (From left to right) Placing a mark 5 cm proximal to the LM, ML, LFC and MFC. 159
Figure 5.5 - Side view showing the skin landmarks of phase one. 159
Figure 5.6 - The skin marks used in phase two of testing (blue crosses). 160
Figure 5.7 - Short lever arm measurement (left) and long lever arm measurement (right). 160
Figure 5.8 –Biodex set-up to test the right hip (left); A participant getting familiar with isometric testing at 0° (right). 161
Figure 5.9 – Biodex isometric testing: (A) at 0 degrees of abduction; (B) at 15 degrees of abduction; (C) at 30 degrees of abduction. 162
Figure 5.10 – Left hip alignment on the reference lines to test hip adduction strength while the opposite limb is stabilised (left); Left hip adduction strength measurement on a distal location using the manual method (middle); Left hip adduction strength measurement on a proximal location using the stabiliser (right) (The green and orange dashed lines represent the hip and limb reference lines respectively; the red dot represent the axis of motion; the yellow arrows represent the direction of push). 164
Figure 5.11 - General design of phase one testing (Orange boxes represent recruitment and screening phase; blue boxes represent testing phase; red box represents analysis phase; the dashed box represents pre-test preparation before each session). 165
Figure 5.12 - Biodex testing protocol in phase one (RPM, revolutions per minute; Rep, repetition; sec, second; green boxes represent recruitment and screening phase; orange boxes represent test preparation phase; blue boxes represent isokinetic tests; grey boxes represent isometric tests and red boxes represent rest periods). 166
Figure 5.13 – MicroFET2 testing protocol in phase one (GrAASP, groin abductor-adductor strength platform; RPM, revolutions per minute; rep, repetitions; Orange boxes

represent the preparation phase; Blue boxes represent the HDD testing stage; Red boxes represent the rest periods). 167
Figure 5.14 - Using the foam pad (left) and the wooden section (right) to achieve proper hip and knee position. 169
Figure 5.15 - Phase two testing of the hip abductor, adductor, flexor and extensor muscle groups (left to right) (yellow arrows represent the direction of push) 171
Figure 5.16 - Phase two testing of the hip internal rotator, external rotator and knee flexor muscle groups (left to right) (yellow arrows represent the direction of push)... 171
Figure 5.17 - General design of phase two testing (Orange boxes represent recruitment and screening phase; blue boxes represent testing phase; red box represents analysis phase; the dashed box represents pre-test preparation before each session). 172
Figure 5.18 - MicroFET2 testing protocol in phase two (STAPH, hip strength assessment platform; RPM, revolutions per minute; rep, repetitions; Orange boxes represent the preparation phase; Blue boxes represent the MicroFET2 testing stage; Red boxes represent the rest periods). 173
Figure 5.19 - Structure of the multilevel modelling (Orange circles represent level 1 and the blue boxes represent level 2 of the model) 178
Figure 5.20 - Statistical methods used to analyse the data of phases one and two (ANOVA, analysis of variance; orange boxes represent aim of the test reflecting the research questions; blue boxes represent the statistical tests used; GrAASP, groin abductor-adductor strength platform; STAPH, hip strength assessment platform). .. 179
Figure 6.1 - Isometric hip abductor muscle torque based on the side as measured by the Biodex in phase one (with percentage of difference and standard deviation bars). 183
Figure 6.2 – Isometric hip adductor muscle torque based on the side as measured by the Biodex in phase one (with percentage of difference and standard deviation bars). 183

Figure 6.3 - Isokinetic hip abductor muscle torque based on the side as measured by the Biodex in phase one (with percentage of difference and standard deviation bars). 184
Figure 6.4 - Isokinetic hip adductor muscle torque based on the side as measured by the Biodex in phase one (with percentage of difference and standard deviation bars). 184
Figure 6.5 - Isometric hip abductor muscle torque based on gender and side as measured by the Biodex in phase one (with percentage of difference and standard deviation bars; DS, dominant side; NS, non-dominant side; *, significant difference between males and females at 0.05)..... 186
Figure 6.6 - Isometric hip adductor muscle torque based on gender and side as measured by the Biodex in phase one (with percentage of difference and standard deviation bars; DS, dominant side; NS, non-dominant side; *, significant difference between males and females at 0.05)..... 187
Figure 6.7 - Isokinetic hip abductor muscle torque based on gender and side as measured by the Biodex in phase one (with percentage of difference and standard deviation bars; DS, dominant side; NS, non-dominant side; *, significant difference between males and females at 0.05)..... 187
Figure 6.8 - Isokinetic hip adductor muscle torque based on gender and side as measured by the Biodex in phase one (with percentage of difference and standard deviation bars; DS, dominant side; NS, non-dominant side)..... 188
Figure 6.9 - Bland-Altman plot between abduction isometric at 0 and abduction fixed distal in phase one (pair 4 from Table 6.8; red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits). 193
Figure 6.10 - Bland-Altman plot between abduction isokinetic at 60°.sec^{-1} and abduction manual distal in phase one (pair 14 from Table 6.8; red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits; *, significant difference at 0.05). 194

Figure 6.11 - Bland-Altman plot between adduction isometric at 30° and adduction fixed proximal in phase one (pair 31 from Table 6.9; red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits). 197
Figure 6.12 - Bland-Altman plot between adduction isometric at 30° and adduction fixed distal in phase one (pair 32 from Table 6.9; red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits). 197
Figure 6.13 - Bland-Altman plot between adduction isokinetic at 60.sec^{-1} and adduction manual proximal in phase one (pair 33 from Table 6.9; Red line illustrates mean of difference; Dotted red lines illustrate 95% confidence interval of the mean of difference; blue lines illustrate agreement limits). 199
Figure 6.14 - Bland-Altman plot between adduction isokinetic at 120.sec^{-1} and adduction manual proximal in phase one (pair 37 from Table 6.9; Red line illustrates mean of difference; Dotted red lines illustrate 95% confidence interval of the mean of difference; blue lines illustrate agreement limits). 199
Figure 6.15 - Bland-Altman plot of abduction testing in phase one (proximal using the manual method) values obtained in the MicroFET2 sessions (pair 41 from Table 6.11; red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits; *, significant difference at 0.05). 203
Figure 6.16 - Bland-Altman plot of abduction testing in phase one (distal using the fixed method) values obtained in the MicroFET2 sessions (pair 44 from Table 6.11; red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits; *, significant difference at 0.05). 203
Figure 6.17 - Bland-Altman plot of adduction testing in phase one (proximal using the manual method) values obtained in the MicroFET2 sessions (pair 45 from Table 6.11; red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval

of the mean of difference; blue lines illustrate agreement limits; *, significant difference at 0.05)..... 204
Figure 6.18 - Bland-Altman plot of adduction testing in phase one (distal using the fixed method) values obtained in the MicroFET2 sessions (pair 48 from Table 6.11; red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits; *, significant difference at 0.05). 204
Figure 6.19 – Torque values of the hip extensor, hip and knee flexor muscles in the sagittal plane in phase two from MicroFET2 testing (with percentage of difference and standard deviation bars; DS, dominant side; NS, non-dominant side; S1, S2 first and second sessions). 212
Figure 6.20 - Torque values of the hip abductor and adductor muscles in the frontal plane in phase two from MicroFET2 testing (with percentage of difference and standard deviation bars; DS, dominant side; NS, non-dominant side; S1, S2 first and second sessions)..... 213
Figure 6.21 - Torque values of the hip internal and external rotator muscles in the horizontal plane in phase two from MicroFET2 testing (with percentage of difference and standard deviation bars; DS, dominant side; NS, non-dominant side; S1, S2 first and second sessions). 213
Figure 6.22 - Torque values based on gender and side as measured in the first session in phase two from MicroFET2 testing (with percentage of difference and standard deviation bars; DS, dominant side; NS, non-dominant side; *, significant difference between males and females at 0.05)..... 215
Figure 6.23 - Torque values based on gender and side as measured in the second session in phase two from MicroFET2 testing (with percentage of difference and standard deviation bars; DS, dominant side; NS, non-dominant side; *, significant difference between males and females at 0.05). 216
Figure 6.24 - Bland-Altman plot of hip internal rotation testing values obtained in the two MicroFET2 sessions in phase two (pair 5 from Table 6.20; red line illustrates mean of

difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits). 220
Figure 6.25 - Bland-Altman plot of hip abduction testing values obtained in the two MicroFET2 sessions in phase two (pair 3 from Table 6.20; red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits). 220
Figure 6.26 - Bland-Altman plot of hip adduction testing values obtained in the two MicroFET2 sessions in phase two (pair 4 from Table 6.20; red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits; *, significant difference at 0.05). 221
Figure 6.27 - Bland-Altman plot of knee flexor testing values obtained in the two MicroFET2 sessions in phase two (pair 7 from Table 6.20; red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits). 221
Figure 6.28 - Bland-Altman plot of hip abduction testing using GrAASP (fixed method on a distal location) (left) and using the STAPH (right) (red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits; *, significant difference at 0.05). 226
Figure 6.29 - Bland-Altman plot of hip adduction testing using GrAASP (fixed method on a distal location) (left) and using the STAPH (right) (red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits; *, significant difference at 0.05). 227
Figure 7.2 – The pattern of hip external rotator muscles weakness (Circles refer to gluteus maximus; triangles refer to deep external rotator muscles; green and yellow colours refer to stronger and weaker muscles respectively in comparison to the other side)..... 232
Figure 7.3 - Lower Crossed Syndrome (Janda, 2014)..... 232
Figure 7.3 - Hip abductor muscle torque-angle data for both sides (Neumann et al., 1988) 236

Figure 7.4 – The short (yellow) and long (red) lever arms. 240

Figure 7.5 – The updated timeline of the dynamometer stabilising platforms according to the authors and the muscle groups of interest showing the designed platforms by the author in the dashed box (HFL, hip flexion; HEX, hip extension; HAB, hip abduction; HAD, hip adduction; HER, hip external rotation; HIR; hip internal rotation; KEX, knee extension; KFL, knee flexion; APF, ankle planter flexion). 252

List of Abbreviations

ACL	Anterior Cruciate Ligament
ADL	Activities of Daily Living
AFL	Australian Football League
AIIS	Anterior Inferior Iliac Spine
ASIS	Anterior Superior Iliac Spine
BMI	Body Mass Index
CHP	Chondropathy
CV	Coefficient of Variation
df	Degrees of Freedom
DoS	Director of Studies
FAI	Femoroacetabular Impingement
GrAASP	Groin Hip Abductor-Adductor Strength Platform
HA	Hip Arthroscopy
HAB	Hip Abduction
HAD	Hip Adduction
HAGOS	Copenhagen Hip and Groin Outcome Score
HER	Hip External Rotation
HEX	Hip Extension
HFL	Hip Flexion
HHD	Hand-held Dynamometry
HIR	Hip Internal Rotation
HOS	Hip Outcome Scale
ICC	Intra-Class Correlation
ICU	Intensive Care Unit
iHOT	International Hip Outcome Tool
IKD	Isokinetic Device
ITBS	Iliotibial Band Syndrome
KEX	Knee extension
KFL	Knee flexion
Lt	Left
LFC	Lateral Femoral Condyle
LM	Lateral Malleolus
LREC	Local Research Ethics Coordinator
MCID	Minimal Clinically Important Difference
MD	Mean of Difference
MFC	Medial Femoral Condyle
MHHS	Modified Harris Hip Score
MM	Medial Malleolus
MMT	Manual Muscle Testing
MRI	Magnetic Resonance Imaging
MVC	Maximum Voluntary Contraction
N	Newton
N.m	Newton Metre
NAHS	Non-Arthritic Hip Score

NRS	Numeric Rating Scale
OA	Osteoarthritis
OR	Odds Ratio
PA	Physical Activity
PASS	Patient Acceptable Symptomatic State
PFPS	Patellofemoral Pain Syndrome
PT	Physiotherapy
QoL	Quality of Life
RA	Rheumatoid Arthritis
Rt	Right
RQ	Research Question
SD	Standard Deviation
SEM	Standard Error of Measurement
SRD	Smallest Real Difference
SS	Sports Specific
STAPH	Hip Strength Assessment Platform
THA	Total Hip Arthroplasty
TMG	Tensiomyography
VAS	Visual Analogue Scale
-2LL	-2 Log Likelihood

Chapter 1 Introduction

1.1. *Hip Pain, Arthroscopy and Muscle Strength Assessment*

The hip joint, which is a weight-bearing joint, is affected by large forces during mobility (Neumann, 2009b). Consequently, pain and various pathologies can develop in the adolescent, young and older populations (Domb, et al., 2015b; Samora, et al., 2011; Wall, et al., 2012). About 10% of the general population complains of hip pain, while 10% to 20% of the injuries found in football and ice hockey players are related hip and groin incidents, and 70% in recreational runners (Quinn, 2010; Reiman & Thorborg, 2014).

Physiotherapy and rehabilitation are the first choices to treat hip and groin pain (Quinn, 2010). When conservative treatment fails, hip arthroscopy can be performed as an alternative to open surgery to treat intra-articular conditions (Safran & Hariri, 2010). The procedure is applied to treat sources of internal pain such as femoroacetabular impingement (FAI), labral pathology and dysplasia (Wall, et al., 2012).

FAI, which is divided into cam-type or pincer-type or both, is one of the most common indications of hip arthroscopy. It is caused by an alteration in the bony morphology of the femoral head-neck junction and the acetabular rim, resulting in altered kinematics and a mechanical conflict between these two structures (Mardones, et al., 2010). Symptomatic FAI can have a direct impact on physical activity, sports, and activities of daily living. Moreover, FAI patients often experience reductions in muscle strength and functional capacity (Casartelli, et al., 2014; Freke, et al., 2016). This condition has been identified as a common cause of non-arthritic pain in active patients, which can negatively affect their ability to engage in physical exercises (Brunner, et al., 2009).

The postoperative success of hip arthroscopy procedures is quantified based on the results of different outcome measures. The impact on function, physical activity and sports are assessed using several questionnaires such as the modified Harris Hip Score

(MHHS), the Hip Outcome Score (HOS) and the Copenhagen Hip and Groin Outcome Score (HAGOS) (Sansone, et al., 2015). The range of motion is typically assessed using the conventional goniometers or the Flexion-Abduction-External Rotation (FABER) test (Magee, 2008; Nho, et al., 2011).

Changes in the hip and thigh muscle strength is also an essential factor that has been investigated. Muscle strength is traditionally evaluated using Manual Muscle Testing (MMT), which is based on a numerical grading system from 0 to 5 (Hislop & Montgomery, 2007). Even though the technique has concurrent validity, it can overestimate muscle strength (Bohannon, 1986). Moreover, the scoring grade could vary by 25% between the examiners (Bohannon, 2002). Isokinetic devices can also be used to measure muscle strength with high reliability (Sole, et al., 2007). However, they are expensive, complicated and require a lot of space and training (Jackson, et al., 2016).

Hand-held dynamometry (HHD) is an easily applied muscle strength technique that has gained popularity in the past years (Bohannon, 2012), and has shown sufficient validity and reliability to measure hip muscle strength (Kawaguchi & Babcock, 2010; Lu, et al., 2011). It is a widely used method to quantify hip muscle strength across a healthy population, athletes and groin pain patients (Fulcher, et al., 2010; Kelln, et al., 2008; Thorborg, et al., 2011). In the two studies forming this thesis, HHD is the method of choice for assessing hip muscle strength, which further explained in section 1.2.

1.2. *Framework of the Thesis*

Study One: Hip Arthroscopy and Hip Muscle Strength Changes (Chapter 3)

Hip arthroscopy is a promising procedure to resolve the sources of pain associated with nonarthritic internal hip pain, preserve the joint, and restore function (Wall, et al., 2012). Performing hip arthroscopy leads to positive changes in hip movement quality, muscle strength and function for an extended period (Casartelli, et al., 2014; Byrd & Jones,

2009). Most patients reported significant improvements in hip muscle strength (Kemp, et al., 2014b; Nepple, et al., 2015), range of motion (Brunner, et al., 2009; Philippon, et al., 2009), and function (Byrd & Jones, 2009; Dippmann, et al., 2014). It preserves the soft tissues, shorten the rehabilitation phase, and speeds-up the return to physical activity (Casartelli, et al., 2014). The procedure can be performed for a variety of patient groups ranging from young patients to high-level athletes (Amenabar & O'Donnell, 2013; Nho, et al., 2011).

Most arthroscopy follow-up studies were performed on the knee joint, which cannot be extrapolated to the hip joint due to the differences in anatomy and biomechanics. Also, the pattern and magnitude of the clinical changes following hip arthroscopy can vary based on the treated condition and sample of patients (McCarthy, et al., 2011; McCarthy & McMillan, 2013). The short- to medium-term follow-up studies of hip arthroscopy patients are lacking in the literature (McCarthy, et al., 2011). The current emphasis is on long-term studies that examine the changes in function and returns to physical activity after one year from the arthroscopy (Meftah, et al., 2011; Polesello, et al., 2014).

The previously published long-term studies where limited by small sample sizes, a homogenous group of patients diagnosed with a particular pathology, or selected outcome measures (Casartelli, et al., 2014; Daivajna, et al., 2015; Kemp, et al., 2015). There is a massive need for short- to medium-term studies based on an extensive sample of patients that examine various aspects of the postoperative outcomes and the factors influencing it (McCarthy, et al., 2011).

The first study of this thesis aimed to fill some of the voids in the literature about the short- to medium-term longitudinal studies. The researcher investigated the postoperative changes that occurred in the patients' physical and functional status following hip arthroscopy. The measurements included in the analysis consisted of muscle strength, the range of motion, function and pain of a large sample of hip arthroscopy patients. These measurements were obtained at four-time points extending over a year after the surgery (Figure 1.1).

The importance of this part of the study is to look at the pattern of recovery regarding strength, the range of motion, function and pain. It will also allow the researcher to quantify the magnitude of change in these measurements. Thus, the patients' recovery time can be estimated to give the clinician an idea of the necessary changes that need to be done to the rehabilitation program to achieve the desired goals. Another aim of the study was to analyse the effects of several factors such as age, sex, and level of athletic activity to examine their impact on postoperative progression (Figure 1.1). The results will aid in the policy and protocol making process regarding the eligibility for surgery, postoperative progression and prognosis. For example, any differences between sedentary patients and athletes will be taken into consideration in designing the rehabilitation program, exercise progression and pre-competition screening (Tijssen, et al., 2016).

Study 2: Hip Muscle Strength Assessment and the Stabilising Platforms (Chapters 5-7)

The hip muscles ensure a smooth transfer of movement from the trunk to the lower extremities and are considered one of the three systems that maintain the joint's integrity and microstability (Dangin, et al., 2016). The strength of the lower limb muscles impacts the activities of daily living, walking velocity, mortality, and the quality of life (Blacklock, et al., 2007; Brown, et al., 1995; Goodpaster, et al., 2006; Rasch, et al., 2010). Even though the hip muscles apply tremendous amounts of forces on the hip joint, which are equivalent to 3-6 times the body weight, any abnormality in the function of these muscles results in pain, weakness and possible hip pathology (Wang, et al., 2017).

Reduction in hip and thigh muscle strength has been found in patients with groin injuries (Tyler, et al., 2010), patellofemoral syndrome (Taylor-Haas, et al., 2014), iliotibial band syndrome, and anterior cruciate ligament rupture (Bittencourt, et al., 2016). Moreover, lower limb injuries based on the muscle strength differences between males and

females have also been documented (Willson, et al., 2012). Hip arthroscopy patients were also found to have reduced hip muscle strength before and after the procedure (Casartelli, et al., 2011; Nepple, et al., 2015), which puts them at a higher risk of developing further pathologies if the issue was not identified and addressed properly (Anderson, et al., 2001; King, et al., 2015; Tyler, et al., 2010). However, the changes in postoperative muscle strength have not been extensively studied before.

Muscle strength assessment is an essential component in the clinical evaluation of healthy individuals, patients with general neuromuscular pathologies, and hip arthroscopy patients to document weakness and strength progression (Magnusson, et al., 1990). Since muscle strength plays a vital role in physical activity and quality of life (Blacklock, et al., 2007), and strength changes are seen postoperatively in hip arthroscopy patients (Casartelli, et al., 2014), it is necessary to objectively evaluate and measure these changes using a quick and reliable method (Figure 1.1).

Hand-held dynamometry (HHD) is becoming a popular objective measurement tool to evaluate muscle strength. However, the location of placing the HHD and the method of stabilisation can directly affect the strength measurements (Kim, et al., 2014; Krause, et al., 2007). Therefore, several stabilisation methods and platforms were introduced to enhance the reliability of HHD testing and to overcome the high testing error encountered with manual stabilisation (Tourville, et al., 2013; Traina, et al., 2010; Widler, et al., 2009). The current stabilising platforms are limited in their capabilities due to their design, size and weight (section 2.7). Also, the validity of some of the platforms was not assessed (Fenter, et al., 2003b). Lastly, there were no follow-up studies to improve the design of these platforms (Figure 1.1).

The second study of this thesis is an extension of the first study focusing on muscle strength testing using a HHD. It is based on two rationales: First, that a HHD was utilised to collect the muscle strength measurements of the hip arthroscopy patients in the first study, and many factors such as the placement location and method of stabilisation can impact these measurements. Second, there is a need for a new platform to stabilise the

HHD, improve the reliability of testing various hip muscles, and overcome the strength differences between the therapist and patient. Thus, the second study aimed to determine which stabilisation method (manual or fixed) and which placement location (proximal or distal) are superior and provide more consistent results. Another aim was to describe the features of two, new, hip muscle strength assessment platforms designed by the author. Also, to evaluate the validity and reliability of these platforms in testing the muscle strength of healthy individuals (Figure 1.1). Future follow-up studies will be performed to look at the validity and reliability of the platforms in assessing hip arthroscopy and groin-pain patients.

The outcome of the second study will be useful to determine where to place the HHD on the lower limb during the clinical assessment. In addition, it will help in deciding if stabilisation using a platform can give more reliable results in assessing the hip muscle strength of healthy active individuals. When the results of the study are applied clinically, it will reduce the inter-tester variabilities in placing the HHD, improve the reliability and objectivity of testing and reduces measurement error. It will also set the ground for new research ideas related to lower and upper limbs muscle strength testing, and innovative stabilising platform designs.

Other

Chapter 2 of this thesis is a literature review covering the relevant hip joint anatomy and hip pain. It also includes the related studies about the outcome of hip arthroscopy procedures, the validity and reliability of hand-held dynamometry testing and the stabilisation platforms. The first study is described in chapter 3, which is about secondary data analysis of hip arthroscopy patients. The following three chapters contain the designing process of the two muscle strength assessment platforms, the methodology and the results of testing. The methodology and results chapters are divided into phase one for the first platform and phase two for the second platform.

Chapter 7 contains the combined discussion of both phases. A general view of the thesis (synopsis), novel contribution, general limitations, recommendations and perspective research are summarised in chapter 8. The framework of this research is illustrated in Figure 1.1.

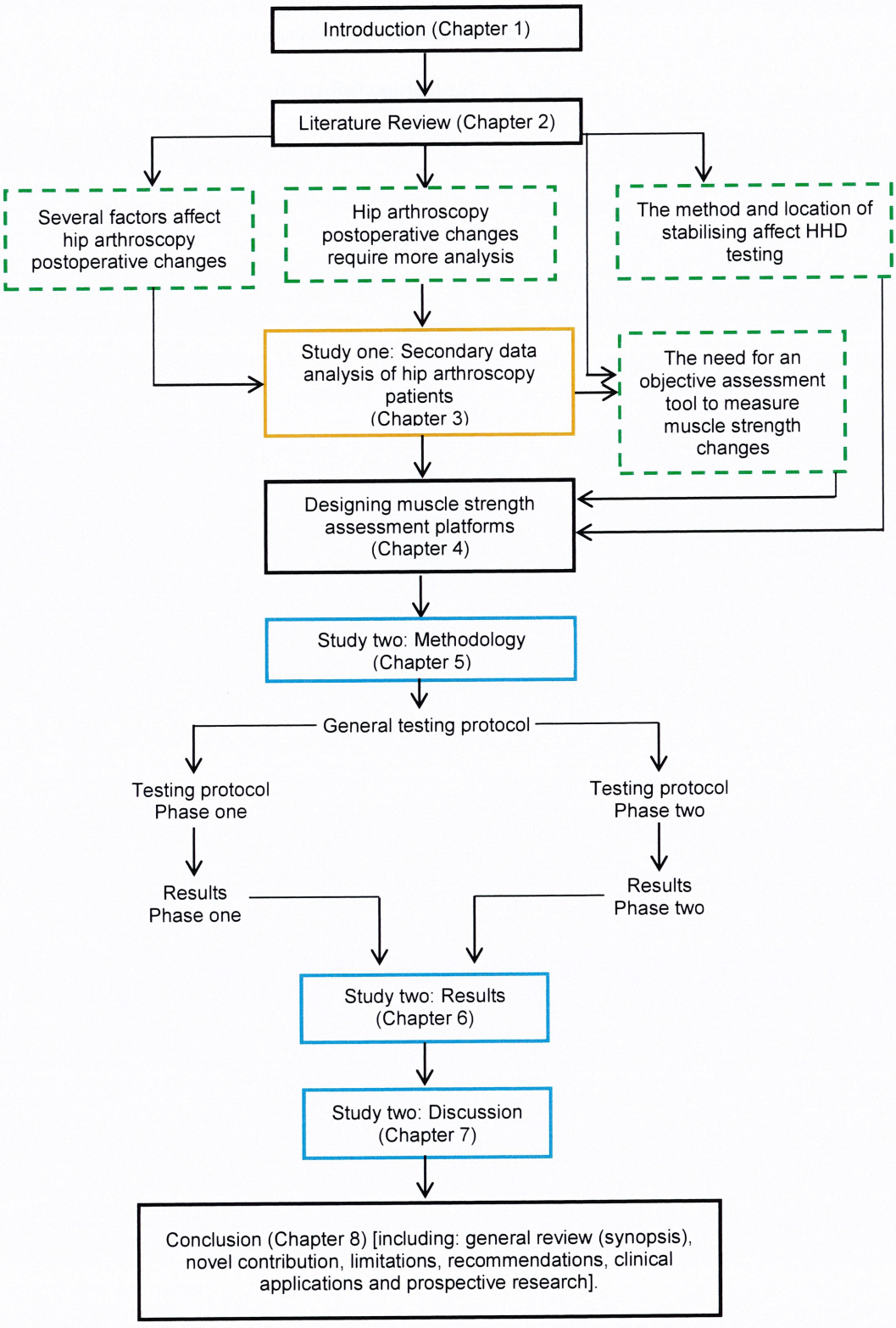


Figure 1.1 – Framework of the thesis (HHD; hand-held dynamometry; green dashed boxes refer to the gap in literature and rationale of the studies; orange box refer to study one and the blue boxes refer to the chapters related to study two.

Chapter 2 Literature Review

2.1. *Hip Joint Anatomy*

2.1.1. Introduction

Understanding the hip joint anatomy is essential to evaluating the importance of hip arthroscopy procedures in resolving intra-articular pain. Thus, the hip joint morphology and structure will be reviewed in this section.

2.1.2. Joint Structure

The hip joint, which is also known as the coxafemoral or acetabulofemoral joint, is formed between the head of the femur and the acetabulum of the pelvis. The pelvis is formed of the ilium, ischium and pubic bones of both sides with the sacrum bone in the middle. The acetabulum, which is the concave socket-like structure, is located just below the ala of the ilium (the large fan-shaped structure of the pelvis). It serves as an attachment point between the head of the femur and the pelvis forming the hip joint (Figure 2.1) (Neumann, 2009b).

The ball-and-socket configuration permits movement in three planes around three axes. Therefore, the joint can move into flexion, extension, abduction, adduction, external rotation and internal rotation. Large forces pass through this joint during mobility. However, the osseous structure, the thick capsule and the relatively large muscles enhance stability and ensure the safe transfer of such forces (Neumann, 2009b).

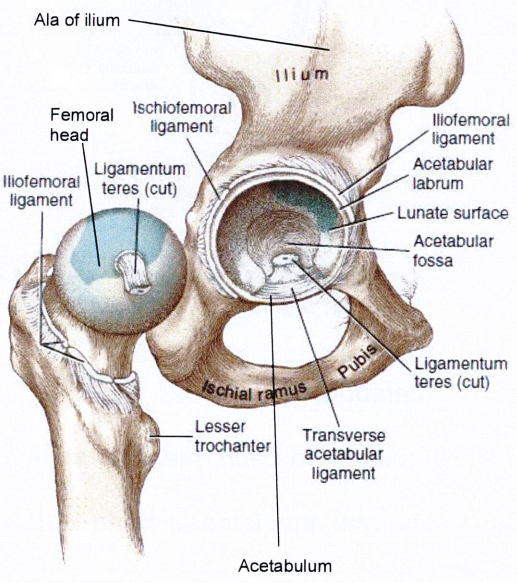


Figure 2.1 - The right hip joint showing the ligaments (Neumann, 2009b).

Arthroscopic exploration of the hip joint reveals two distinct compartments

(Dorfmann & Boyer, 1999). The central compartment, which is the space between the surface of the femoral head and the acetabulum, contains the lunate articular surface, ligamentum teres, and the acetabular fossa. The peripheral compartment consists of the remaining areas of the hip joint within the capsule (Dienst, 2005; Wall, et al., 2011).

A synovial membrane covers the inner surface of the hip joint and produces synovial fluid to maintain a smooth gliding of the surfaces of the joint (Neumann, 2009b). From the outside, a dense capsule composed of the iliofemoral, ischiofemoral, and pubofemoral ligaments surrounds the whole joint.

The proximal third of the femur consists of the femoral head and shaft, which are connected by the femoral neck at the intertrochanteric crest. Distally, a prominent and palpable bony structure known as the greater trochanter acts as an attachment point for the hip abductors. Medially, the lesser trochanter lies deep and serves as an attachment point for the hip flexors (Neumann, 2009b).

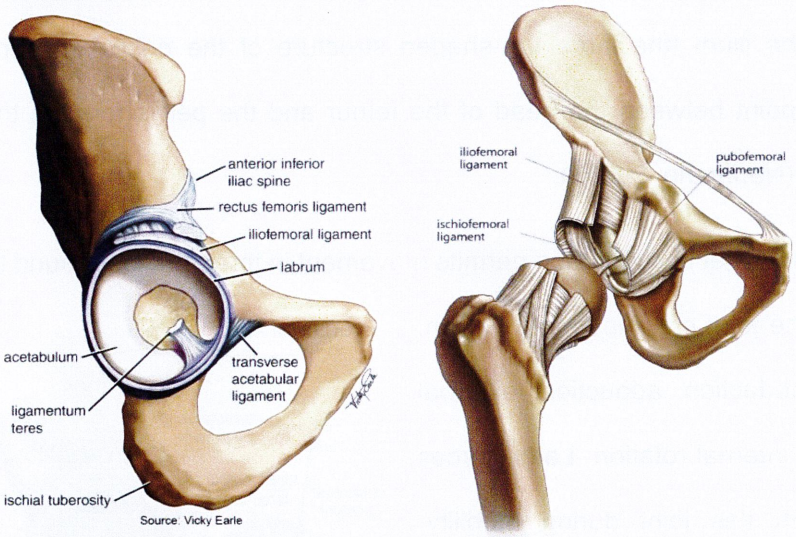


Figure 2.2 –Side view of the pelvis showing the acetabulum and labrum (Left); Ligaments of the hip joint (Right) (Retchford and Kemp 2013).

2.1.3. Acetabulum

The acetabulum (in *Latin,* meaning vinegar cup), is the cup-like structure in the pelvis that meets with the femoral head. The femoral head comes into contact with the acetabulum in the horse-shoe shaped lunate surface, which is covered with a thick

cartilage (Dvorak, et al., 1990). The cartilage slightly flattens to dissipate the force applied on the hip during physical activity, which is reported to increase to 300% of body weight during the mid-stance phase of walking (Neumann, 2009b).

A flexible fibrocartilage known as the labrum surrounds the outer rim of the acetabulum. The labrum transforms into the transverse acetabular ligament near the acetabular notch. A cross section through the labrum shows that it is triangular in shape with the apex pointing outward for about 5 mm towards the femoral head (Neumann, 2009b).

Due to the concavity of the acetabulum, it only covers certain areas of the femoral head. The centre-edge angle is one of the most common measurements to determine the amount of coverage of the femoral head by the acetabulum and to diagnose hip impingement. In a frontal-plane x-ray radiograph of the hip, the intersection between a vertical reference line passing through the femoral head and an acetabular reference line passing through the lateral edge of the acetabulum and centre of the femoral head forms the centre-edge angle (35° in adults). A reduced centre-edge angle to about 15° can increase force application by about 50% due to the reduced surface area, which may lead to osteoarthritis in the long term. On the other hand, increased coverage can result in hip impingement (explained in section 2.2) (Neumann, 2009b). Further description of the anatomy can be found elsewhere (Fernandez & Gastaldi, 2012; Moore, et al., 2006; Neumann, 2009b).

2.2. *Anatomical Variations Based on Sex*

Anatomy of the pelvis is different between males and females due to the skeletal, hormonal, and anthropometric variations. The female pelvis is characterised by a larger transverse pelvic inlet diameter (~13 cm) and a wider subpubic angle (90°-100°). In comparison, the male pelvis is steeper, has a narrower pelvic inlet diameter (20-25% smaller than females) and a reduced subpubic angle (60-70°) (Figure 2.3). Therefore, the ilia and acetabuli of females are oriented more downward and posterior in comparison to an upward and anterior orientation in males (Marani & Koch, 2014; OpenStax, 2015).

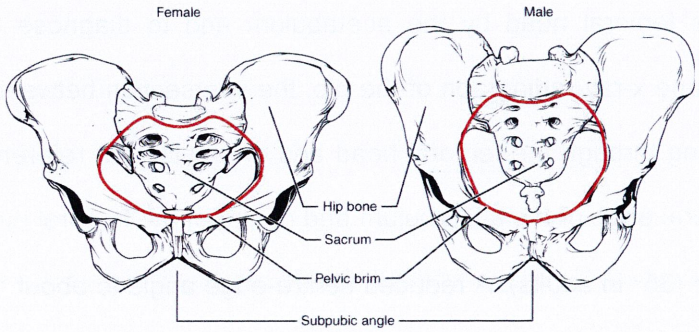


Figure 2.3 - Anatomical differences between the male and female pelvises (OpenStax, 2015).

Tensiomyography (TMG) has been used to examine the variation in muscular activity between males and females. It is a non-invasive diagnostic tool that is applied to test the contractile capacity, rigidity, and mechanical characteristics of the superficial muscles. One hundred and sixty-six volleyball players volunteered to have their thigh muscles tested. The results showed a significant difference in the normalised response time (Vrn)* between males and females in the biceps femoris ($F[1] = 18.57$, $p < 0.001$, $\eta p^2 = 0.12$) and the vastus medialis tests ($F[1] = 6.30$, $p = 0.01$, $\eta p^2 = 0.04$). The results also showed a higher Vrn in the biceps femoris of the males (55.28 $mm.s^{-1}$) compared to the females (43.10 $mm.s^{-1}$, $p < 0.001$). However, the females had slightly higher Vrn of the vastus medialis (56.67 $mm.s^{-1}$) than the males (52.99 $mm.s^{-1}$, $p = 0.013$)

* Described by Rodriguez-Ruiz et al. (2014) as "the relationship between the difference in the radial displacement between 10% and 90% of the muscle belly maximum radial displacement and the increase in the muscular contraction time between these same values".

(Rodriguez-Ruiz, et al., 2014). The findings indicate that the response time of the knee muscles differs between males and females. Therefore, some of the muscular properties can vary based on sex.

Hormonally, the production of testosterone by the testes influences muscular growth and increases strength by 15% to 53% in males compared to females. In general, females were found to be one-third less strong than males, with the lower limbs of females being 66% as strong as males. The direct association between the muscle cross-sectional area and strength development is the most logical reason to explain this variation (Pheasant & Haslegrave, 2006).

The variation in force production between males and females has been associated with alteration in the hip and knee kinematics. For instance, female basketball players have a wider apparent knee flexion angle (males= 23.17° ± 4.29°, females= 26.94° ± 4.30°, p =0.006), a wider knee valgus angle (males= 4.86° ± 4.49°, females= 5.23° ± 4.14°, p >0.05), and an increased internal rotation angle (males= -6.01° ± 6.12°, females= -0.38° ± 6.12°, p =0.02) at initial contact compared to the male players (Suzuki, et al., 2015). During running, females can produce a greater apparent hip adduction of 4.0° (females= 11.2° ± 3.9°, males= 7.2° ± 2.8°, p =0.001, effect size= 1.2), and an increased apparent knee valgus of 2.8° (females= 3.4° ± 4.0°, males= 0.6° ± 2.5°, p =0.01, effect size= 0.87) at initial contact in comparison to the male runners (Willson, et al., 2012).

The lack of sufficient hip muscle strength during mobility has been linked with pathokinesiological findings such as increased adduction and internal rotation. Female basketball players have significantly weaker hip abductor (males= 21.65 ± 3.82 %BW, females= 18.48 ± 4.79 %BW, p =0.03) and weaker hip external rotator muscles in the seated position (males= 24.18 ± 3.90 %BW, females= 18.73 ± 3.75 %BW, p <0.01). They also demonstrate a negative correlation between the apparent knee valgus angle at initial contact and hip abduction strength (r_s= -0.46, p =0.03) (Suzuki, et al., 2015).

Muscle peak activation is also not similar between males and female. Female runners show a 30% increase in the muscle peak activation level of the gluteus maximus compared to males (males= 61.2 ± 25.5, females= 87.7 ± 41.8, p =0.028, effect size= 0.79), which was reported previously in females with patellofemoral syndrome (Willson, et al., 2012).

Increased hip adduction motion is associated with iliotibial band syndrome (ITBS), patellofemoral pain syndrome (PFPS) and tibial stress fracture in adult female long-distance runners. Similarly, increased internal rotation is linked with ITBS in adult male long-distance runners (Taylor-Haas, et al., 2014). The smaller knee flexion angle caused by the hip extensors weakness and the larger knee valgus caused from hip abductors weakness during landing have been recognised as risk factors for developing anterior cruciate ligament (ACL) injuries (Suzuki, et al., 2015). Cross country runners showed that peak isokinetic hip extensor torque is inversely correlated with hip movement in the transverse plane (r = -0.387, p =0.026), and hip abductor torque was inversely correlated with hip motion in the frontal plane while running (Taylor-Haas, et al., 2014).

The reduction in hip muscle strength can have serious consequences. Patients with weak hip abductors and hip extensors would be more prone to knee and ankle injuries due to abnormal mechanics and force distribution. The females, in particular, demonstrate increased knee valgus and reduced gluteus medius activity during landing. When combined, these two factors could result in PFPS and ACL injuries. Therefore, more light should be shed on identifying the individuals with hip muscle weakness that could be at risk. Moreover, a valid and a reliable strength testing method should be used as part of the clinical assessment protocols. A summary of these values can be found in Table 2.1.

Table 2.1 - Overview of the studies that provided normative values of lower extremities' muscle strength.

First author (year)	Study Details		Sample Details		Results	Main Findings and Comments
	Device	Muscles tested	Number and sex	Age in (M±SD) or range in years		
(Belhaj, et al., 2016)	Isokinetic (Cybex Norm)	HAB and HAD (peak torque)	12 males (asymptomatic)	23.2±3.9	Dominant at 60°.sec^{-1} : HAB= 129.17±22.87 and HAD= 111.25±23.97. Non-dominant at 60°.sec^{-1} : HAB= 131.83±19.14 and HAD= 115.50±36.77 (in Newton.metre).	HAB showed higher torque values than HAD.
(Bohannon, et al., 2011)	Belt-fixed HHD (MicroFET) and an isokinetic device (Biodex).	KEX (peak torque)	71 males and 113 females (n=184).	46.3±22.7	MicroFET: KEX Rt=138.9±63.4 and KEX Lt=135.6±59.2. Biodex: KEX Rt=172.6±70.6 and KEX Lt=171.2±72.7 (in Newton.metre).	The MicroFET underscored the participants KEX strength compared to the Biodex.
(Krause, et al., 2007)	HHD (MicroFET2).	HAB and HAD (normalised peak torque to body weight and height).	12 males and 9 females (n=21).	Males: 24±3; females: 25±3.	HAB: long lever arm=10.7%±2.2% body weight [BW]x height [Ht] and short lever arm=7.1%±1.5% BWxHt (p<0.001). HAD: long lever arm with top limb supported on bench=11.9%±2.8% BWxHt, long lever arm with manual stabilisation of top limb=10.6%±2.5% BWxHt (p=0.001). HAD: short lever arm with top limb supported on bench=6.2%±1.4% BWxHt and short lever arm with manual stabilisation of top limb=6.9%±1.9% BWxHt (p<0.05)	Torque production in long lever arm is greater than short lever arm in hip abduction and adduction testing.
(Lourencin, et al., 2012)	Isokinetic (Cybex 6000).	HAB and HAD (peak torque).	Males (n=20).	21 to 30	Dominant at 60°.sec^{-1}: HAB=199.4±30.35 and HAD=206.3±58.31. Non-dominant at 60°.sec^{-1}: HAB=188.4±37.28 and HAD=189.15±60.46. Dominant at 120°.sec^{-1}: HAB=169.75±25.57 and HAD=167.2±53.04. Non-dominant at 120°.sec^{-1}: HAB=160.15±35.90 and HAD=157.1±48.84 (in Newton.metre).	Reduction in torque output as the speed increased.

(Muff, et al., 2015)	Isokinetic (Con-Trex MJ) and HHD (MicroFET2).	KFL and KEX.	20 males and 10 females (*n*=30).	32.8±11.5	Dominant at concentric 60°.sec^{-1}: KEX=179.1±52.2 and KFL=91.9±28.5. Non-dominant at concentric 60°.sec^{-1}: KEX=165.4±52.0 and KFL=85.1±25.5. Dominant at concentric 180°.sec^{-1}: KEX=132.3±39.4 and KFL=72.2±22.9. Non-dominant at concentric 180°.sec^{-1}: KEX=130.4±39.6 and KFL=71.8±24.1. Dominant at eccentric 60°.sec^{-1}: KEX=201.9±61.9 and KFL=123.1±44.1. Non-dominant at eccentric 60°.sec^{-1}: KEX=195.5.4±69.8 and KFL=116.7±36.8. Dominant at isometric 0°.sec^{-1}: KEX=177.2±57.3 and KFL=84.0±25.2. Non-dominant at isometric 0°.sec^{-1}: KEX=195.5±69.8 and KFL=116.7±36.8 (in Newton.metre). HHD Dominant: KEX=52.9±16.7 and KFL=26.7±6.0. HHD non-dominant: KEX=51.1±17.2 and KFL=26.9±53 (in kgf).	Torque production reduces when the speed of testing increases. The dominant side is stronger than the non-dominant side.
(Wollin, et al., 2016)	Platform with load cell (XTRAN).	KFL	Male Football players (*n*=21).	16.81±0.54	Rt KFL=450.46±59.13 and left KFL=433.03±54.18 (in Newton).	Less than 4% side-to-side hamstrings strength difference exists in healthy football players.

Infor., information; M, mean; SD, standard deviation; HHD, hand-held dynamometer; HAB, hip abductor; HAD, hip adductor; KEX, knee extensor; KFL, knee flexor; Rt, right; Lt, left.

2.3. *Intra-articular Hip Pain*

Hip pain is a common complaint among active individuals. Hip and groin injuries make up approximately 10% of the cases at sports medicine centres, while the annual hip injury rate in recreational runners is estimated at 70% (Quinn, 2010). The aetiology of hip pain includes femoracetabular impingement (FAI), ligamentum teres and labral tears, loose bodies, and hip instability (Figure 2.4) (Wall, et al., 2012). Further details of these conditions are described in the literature (Wilson & Furukawa, 2014).

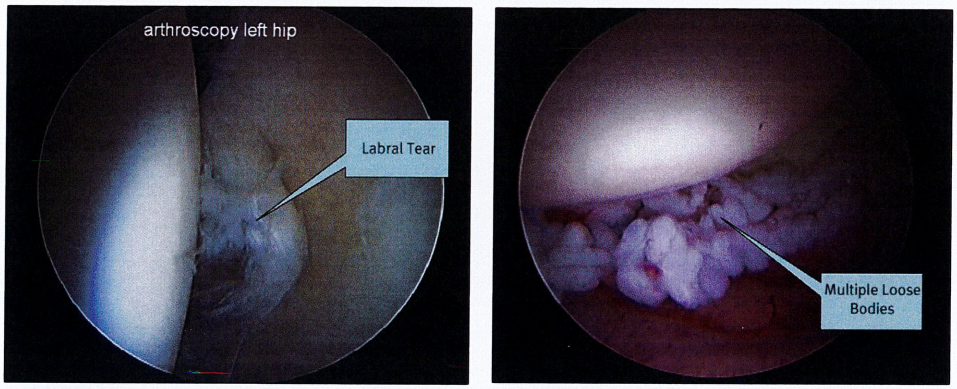


Figure 2.4 - Labral tear (left) and multiple loose bodies (right) as observed during hip arthroscopy procedures (Wall, et al., 2012).

Femoroacetabular impingement is a well-known source of intra-articular pain subdividing into cam or pincer impingement. Cam impingement occurs with asphericity of the femoral head-neck junction in the form of osseous abnormality and overgrowth typically present in young and active males (Figure 2.5). The condition can be easily recognised in a plane anteroposterior x-ray radiograph of the hip. Clinically, placing the hip into flexion, adduction and internal rotation elicits internal hip pain or discomfort, which is the quickest screening test. Due to the abnormally shaped head-neck junction, the acetabulum can be affected by the shearing force when it comes in contact with the femur, which damages the anterosuperior labrum and cartilage. Patients with cam impingement are usually young and active males (Byrd, 2010; Samora, et al., 2011).

Pincer impingement is caused by increased acetabular coverage of the femoral head. The deepened acetabulum or a malformed acetabular rim results in over coverage of

the femoral head (Figure 2.5). Because of the abnormal load, fissures and ganglia develop in the labrum leading to ossification of the acetabular rim. The cycle is repeated over time resulting in degenerative changes and increased coverage. Pincer impingement can be diagnosed radiographically with the presence of the crossover sign, which is the overlap of the medial wall of the acetabulum over the ilioiscial line. The affected patients are typically middle-aged females complaining from pain with hip extension and external rotation (Byrd, 2010; Samora, et al., 2011). Mixed type impingement is the most common and combines the features of the cam and pincer impingements (Reichenbach, et al., 2010).

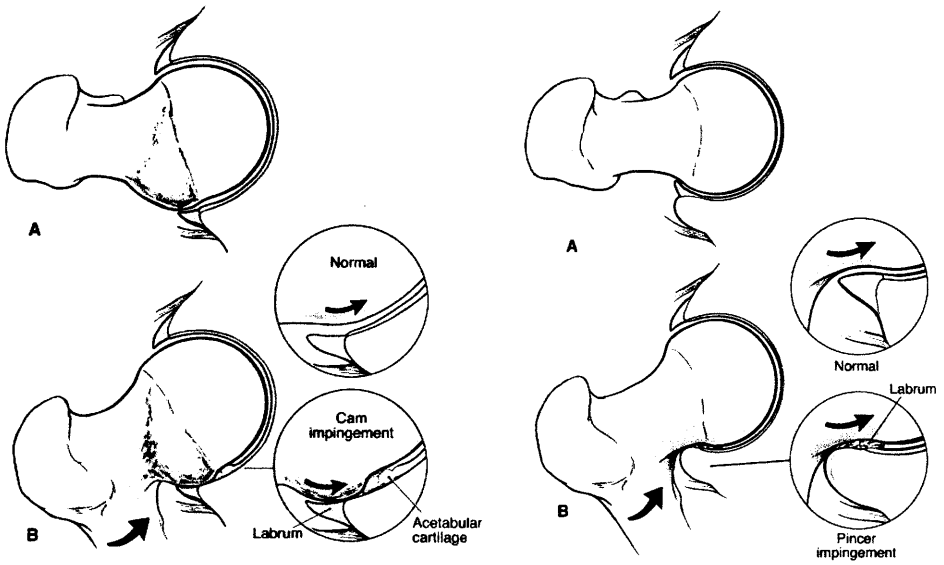


Figure 2.5 - Cam impingement (left) and pincer impingement (right) (Byrd, 2010).

2.4. *Hip Arthroscopy Procedures*

Hip arthroscopy (HA) is usually the last option after conservative treatment. Failure to respond to a minimum of six weeks of physiotherapy (PT), the absence of apparent arthritic changes, and the presence of related clinical and radiological findings indicate the need for arthroscopy. Total or partial hip arthroplasty (THA) is another solution that is often performed for patients with severe pain, osteoarthritis (OA) or rheumatoid arthritis (RA) (Philippon, et al., 2007a).

Dr Michael Burman first introduced the idea of hip arthroscopy in 1931 (Burman, 2001). The procedure was applied in the 1970s and further developed in the 1990s (Wall, et al., 2011). The surgical procedure of hip arthroscopy can be performed in a supine or side-lying positions based on the surgeon's preference and the associated pathology. Two or three portal systems (anterior, anterolateral and posterolateral) are used to visualise the central compartment of the joint and to perform the necessary excisions or repairs (Dienst, 2005).

The anterolateral portal is inserted through an incision about 2 cm anterior to the anterosuperior edge of the greater trochanter at 10° to 20° cranially and 20° to 30° posteriorly. It penetrates through the gluteus medius and tensor fascia latae muscles. The surgeon then inserts the posterolateral port at about 3 cm superior to the posterior edge of the greater trochanter aligned at 10° cranially and 30° anteriorly. This portal penetrates through the gluteus maximus, gluteus medius, gluteus minimus and the joint capsule to reach the joint. The anterior portal is inserted 4 to 6 cm distally to the anterior superior iliac spine at an angle of 30° medially and 30° to 45° cranially. When the third portal is used, it penetrates through the muscle bellies of the sartorius and rectus femoris. Penetrating these structure can affect the postoperative joint range of motion and muscle strength (discussed in section 2.5). Further explanation of the arthroscopy procedure is described elsewhere (Byrd, 2005; Dienst, 2005; Simpson, et al., 2010; Wall, et al., 2011).

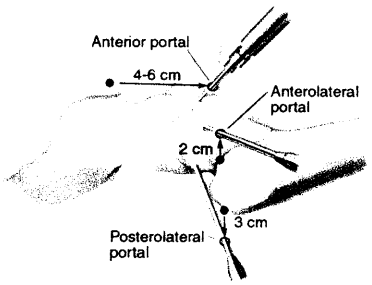


Figure 2.6 - Portals of the central compartment of the hip (Dienst, 2005).

The indications for HA procedures are many. Loose bodies, labral and chondral lesions, ligamentum teres lesions, and femoroacetabular impingements are the most common conditions treated with arthroscopy (Table 2.2) (Safran & Hariri, 2010). The results of a study based on 47 patients showed that labral pathology is the leading cause of having hip arthroscopy procedures (Byrd & Jones, 2001). Degenerative hip diseases cannot be

considered an indication for arthroscopy, and the presence of joint narrowing or arthritic changes could result in poorer prognosis (Byrd & Jones, 2001; Lynch, et al., 2013).

Table 2.2 - Indications for hip arthroscopy in a sample of 47 patients (Byrd and Jones, 2001)

Diagnosis	*n*
Labral pathology	27
Chondral damage	23
Ligamentum teres damage	11
Arthritic disorder	7
Dysplasia	5
Loose bodies	3
Impinging osteophyte	2
Avascular necrosis	1
Synovitis	1
Perthes disease	1
Total	47

2.5. *Outcomes of Hip Arthroscopy Procedures*

Hip arthroscopy is an alternative treatment to open surgery in resolving intra-articular hip pain with good to excellent satisfaction rates and outcomes (Byrd & Jones, 2009; Lynch, et al., 2013). The procedure is very safe compared to open hip surgery with an overall complication rate of 4% to 5.5% only, most of which are non-life threatening (Kowalczuk, et al., 2013; Lodhia, et al., 2015). Having the arthroscopy procedure results in marked improvements in hip kinematics, strength and pain in patients for more than 2.5 years (Casartelli, et al., 2014; Lodhia, et al., 2015). However, the postoperative prognosis can vary depending on the grade of arthritic changes, the labral tear size, and the number of anchors used. Also, extra-articular factors such as sex, age at the time of surgery, BMI, and back pain may also play a role in survivorship of the procedure (Redmond, et al., 2015b). Survivorship, which is discussed in section 2.5.6, is the extent of success of the hip arthroscopy procedure before re-operation or shifting to total hip arthroplasty (THA).

2.5.1. *Outcome Measurement Tools*

Hip arthroscopy can also result in marked improvements in the function and physical activity levels of the patients, which are measured using validated questionnaires. Pain also showed significant reduction following the procedure, as reported using outcome measures such as the visual analogue scale.

Several hip outcome measures such as the modified Harris Hip Score (MHHS), the Hip Outcome Score (HOS) and the Non-arthritic Hip Score (NAHS) were developed to measure the change in function and pain (Domb, et al., 2015b). Other commonly used outcome measures include the International Hip Outcome Tool (iHOT) and the Copenhagen Hip and Groin Outcome Score (HAGOS) (Sansone, et al., 2015).

The MHHS is a modified version of the original Harris Hip Score questionnaire, which was developed by William Harris (Kemp, et al., 2012; Byrd & Jones, 2000). The new version does not include the physical assessment part (Appendix A. 3) (Kemp, et al., 2012). Although the construct validity of the MHHS has been established (Kemp, et al., 2012), the reliability has not been determined yet on arthroscopy patients (Martin, 2005; Thorborg, et al., 2010b). The HOS is considered a valid tool to measure the activities of daily living (ADL) and activities of sport in patients with labral tears (Martin, et al., 2006b). The NAHS is also regarded as a valid and a reliable outcome measurement tool in young and active patients (Christensen, et al., 2003).

Chahal et al. (2014) calculated the minimal clinically important difference (MCID)* and the patient acceptable symptomatic state (PASS)† of the MHHS in assessing hip arthroscopy patients. The MCID was 13.0, 9.0, and 20.0 points at 3, 6, and 12 months following hip arthroscopy. Regarding the PASS, the value at 1-year follow-up was 84. The patients with lower baseline scores were less able to reach the MCID of the MHHS

* Measures the smallest change in the patient's status that is considered clinically important. In other words, the lowest point where a patient feels better or worse (Chahal et al., 2014).
† Measures the level of symptoms that separates being well (or feeling good) from being unwell (ibid)

at 1-year follow-up (Odds ratio (OR)* =0.28) (Chahal, et al., 2014). The study results indicate that those who started with lower scores and had worse conditions are more likely to take more time to achieve the PASS postoperatively. Therefore, baseline MHHS scores can be used along with other parameters to predict postoperative prognosis.

2.5.2. Changes in Muscle Strength

Research studies looking into the postoperative changes following HA showed general improvement in hip muscle strength (Casartelli, et al., 2014), even though, only a small number of studies reported muscle strength changes. Patients with FAI had preoperative strength deficits of 46% and 42% of the hip abductor and flexor muscle groups respectively. They also had bilateral strength deficits of 8.7% and 8% in the same muscle groups (Nepple, et al., 2015). Another study showed statistically significant reductions in the preoperative strength of the hip flexor (34%, p <0.01) and adductor (31%, p <0.05) muscle groups in the FAI patients compared to healthy individuals (Table 2.3) (Casartelli, et al., 2014). These two studies indicate that patients with intra-articular hip pain have coexisting muscle weakness that may have developed due to muscle inhibition or lack of activity.

Postoperative strength changes vary based on the muscle group. In a case study of an American Football player who suffered from posterior hip subluxation, the preoperative clinical assessment showed reduced hip flexor, extensor and adductor muscle strength. Following HA and nine weeks of rehabilitation, muscle strength increased by 83% in the abductor, 80% in the adductor, 23% in the extensor, 65% in the flexor, 89% in the external rotator, and 57% in the internal rotator muscle groups compared with pre-operative measurements. By the end of 16 weeks of rehabilitation phase which lasted for 16 weeks, the hamstring muscles showed a remarkable increase of 356% due to

* OR>1 indicates positive outcome and OR<1 indicates negative outcome.

hypertrophy and absence of pain. The noticeable increase in strength was accompanied by an improvement in function and return to sport (Table 2.3) (Philippon, et al., 2009).

Some hip arthroscopy patients have also showed a lack of full strength recovery in the long term. In a study by Casartelli et al. (2014), the postoperative findings after 2.5 years showed statistically significant strength improvement in the abductor (9%, p <0.05), adductor (34%, p <0.05), internal rotation (59%, p <0.001), external rotation (37%, p <0.001), flexor (23%, p <0.05), and extensor (43%, p <0.01) muscle groups compared to pre-operative measurements. However, the hip flexor muscles were significantly weaker (-18%, p <0.05) in the patient group compared to the healthy individuals at follow-up (Table 2.3). In a similar study, the hip muscle strength and range of motion changes of hip arthroscopy patients (n=84) were compared with healthy participants (n=60). They were followed up for 12 to 24 months after the surgery. On the last follow-up, strength of the adductor (Mean of difference (*MD*)= 0.27 $N{\cdot}m.kg^{-1}$, 95% CI [0.14, 0.41], p <0.001), abductor (women only, *MD*= 0.46 $N{\cdot}m.kg^{-1}$, 95% CI [0.27, 0.65], p =0.001), extensor (*MD*= 0.25 $N{\cdot}m.kg^{-1}$, 95% CI [0.10, 0.41], p =0.001), flexor (*MD*= 0.31 $N{\cdot}m.kg^{-1}$, 95% CI [0.19, 0.42], p <0.001) and external rotator muscle groups (*MD*= 0.09 $N{\cdot}m.kg^{-1}$, 95% CI [0.00, 0.17], p =0.044) were significantly lower in the patients compared to the healthy participants (Table 2.3) (Kemp, et al., 2014b).

Sufficient isometric muscle strength of the hip and thigh muscles should be achieved by the end of the rehabilitation phase. Strong muscles are particularly important to reduce the compressive forces on the hip joint and avoid developing further pathologies. In a recent study, the single leg squat performance of 34 patients was assessed 1 to 2 years post-arthroscopy and compared to the results of a control group. At peak squat, the patients had significantly higher apparent hip adduction (*MD*= 2.7°, 95% CI [0.7°, 4.8°], p =0.008) and apparent knee valgus (*MD*= 4.0°, 95% CI [1.0°, 7.1°], p =0.01) on both sides compared to the controls. During single leg stance, the operated side showed a significant increase in pelvic obliquity compared with the opposite side (*MD*= 1.2°, 95% CI [0.1°, 2.3°], p =0.045) (Charlton, et al., 2016).

For a few weeks after the surgery, it is very common for the patients to experience difficulty in recruiting the iliopsoas muscle. Consequently, they tend to over activate the secondary hip flexors such as the rectus femoris and tensor fasciae latae (Casartelli, et al., 2014; Edelstein, et al., 2012). Over-activity in these muscles may eventually progress to tendinopathy*, in which reductions of physical activity and the rehabilitation exercises are warranted (Enseki, et al., 2010).

Nonetheless, the deactivated hip flexor and extensor muscles can become stronger 1 to 2 years post-arthroscopy. However, increased activity over time does not necessarily indicate an increase in strength. Over-activity in these relatively strong and big muscles could also be a mechanism to compensate for the inhibited deep hip external rotator muscles, which act as the joint stabilisers (Charlton, et al., 2016). Similar findings of decreased deep segmental stabilisation and increased superficial muscular activity have been seen in patients with chronic neck, low back and hip pain (Retchford, et al., 2013; Falla, et al., 2004; Hodges & Richardson, 1996). A summary of these studies is listed in Table 2.3.

Summary

The development of hip muscle strength to achieve the best functional outcome after surgery is one of the major objectives of rehabilitation. Hip muscle strength may take up to a year or more to return to preoperative levels. Even though, some patients still experience muscle weakness for longer periods. Future studies should focus on analysing muscle strength changes immediately after the procedure and follow-up for up to a year. Such studies may improve our understanding of the phases the patients go through regarding gain in muscle strength.

* Pathology in the tendon most commonly in the form of tendon inflammation.

2.5.3. Changes in Range of Motion

The range of motion assessment of the hip joint includes goniometry and muscle length tests. Goniometry testing of the hip and other joints has been extensively explained before (Reese & Bandy, 2016). Patrick's or the FABER (flexion-abduction-external rotation) is a widely applied test to detect internal hip and sacroiliac joint pathology (Magee, 2008; Ross, et al., 2003; Zimny, 1998).



Figure 2.7 - Patrick's (FABER) test (Biggs, 2010).

Instead of being used as a diagnostic test, it can be used to monitor the overall progression of the hip range of motion. The FABER test is performed by passively placing the lateral malleolus of the affected limb just above the patella (knee cap) of the non-affected side. Then, the knee is allowed to passively lower towards the table while the examiner stabilises the opposite aspect of the pelvis (Figure 2.7) (Magee, 2008). The FABER test distance is the measurement between the lateral border of the patella to the table (Leard, et al., 2009; Ross, et al., 2003). However, the validity and reliability of the FABER distance test in assessing the flexibility of the hip following hip arthroscopy procedures are not established yet.

The changes in the hip range of motion after hip arthroscopy varied. In a case series study on 64 FAI patients, Nho et al. (2011) reported no significant changes in hip external rotation (HER) range of motion (ROM) (48.7° ± 10.1° to 49.5° ± 2.9°, p >0.05) and hip internal rotation (HIR) ROM (21.7° ± 10.8° to 19.6° ± 10.4°, p >0.05) postoperatively, even though, hip flexion (HFL) ROM showed significant improvement (116.5° ± 10.4° to 120° ± 0°, p =0.02). Philippon et al. (2009) found no changes in HIR and HER passive range of motion (PROM) at follow-up. However, a marked improvement of 13° was observed in HFL and hip adduction (HAD). Hip abduction (HAB) showed less change with 4° degrees only. The side-to-side comparison showed equal PROM for HFL and HAD, while there were slight differences of 10° and 5° in HIR and

HER respectively. On the other hand, the FABER test distance have also showed a 10-cm improvement postoperatively (Table 2.4).

The patients who had hip arthroscopy were also compared to healthy individuals. Kemp et al. (2014b) concluded that the hip arthroscopy patients had reduced HIR ROM of 5°, 95% CI [2, 8] (p=0.004) and increased hip extension (HEX) ROM of 5°, 95% CI [1, 9] (p=0.019) compared to the control group. When based on sex, the female patients with chondropathy* showed reduced HER ROM compared to the males (*MD*= 12°, 95% CI [7°, 16°] (p <0.001) (Table 2.4). The study results show that those who had HA may still have reduced HIR compared to the healthy individuals, which should be taken into consideration when analysing postoperative ROM findings.

While some studies showed no significant improvement in ROM, the results of Brunner et al. (2009) were different. When the patients were followed up for an average of 2.4 years after the HA, HIR ROM improved from 6° (range -20° to 45°) to 19° (range -5° to 45°) (p<0.001) and HFL ROM improved from 107° (range 60° to 130°) to 122° (range 70° to 145°) (p<0.001). A summary of these studies can be found in Table 2.4.

In the studies by Nho et al. (2011) and Philippon et al. (2009), the methods used to test hip ROM may have been different compared to the method employed by Brunner et al. (2009). In the latter study, HFL and HIR ROM were assessed from a supine position, which is a strange position to test HIR ROM. Furthermore, the surgical technique, amount of osteoplasty, and capsular release or tightening may have varied.

Summary

The magnitude of change in ROM can vary based on the performed procedure, patients' characteristics and the assessment method. While the ROM of patients generally improve, those with chondropathy may not show equivalent improvement.

* Disease of the joint cartilage.

Table 2.3 – Overview of the studies investigating the changes in muscle strength following hip arthroscopy procedures.

First author (year)	Study Details		Sample Details		Results	Main Findings and Comments
	Descr.	Muscles	Number and sex	Age in (M±SD) or range in years		
(Casartelli, et al., 2014)	Cohort study - symptomatic FAI.	HFL, HAD, HIR and HER.	3 males and 5 females (n=8).	29±10	Preoperative: significant difference in strength of HFL= -34% (p<0.01) and HAD= -31% (p<0.05) compared to controls. Postoperative: significant increase in strength of HAD=25% (p<0.05), HAB=9% (p<0.05), HIR=59% (p<0.001), HER=37% (p<0.001), HFL=23% (p<0.05), HEX=43% (p<0.01) compared to preoperative measures. Only HFL= -18% (p<0.05) was significantly lower compared to controls at follow up.	Significant improvement in hip muscle strength postoperatively.
(Charlton, et al., 2016)	Case study - 1 to 2 years following hip arthroscopy.	HFL and HEX.	17 males and 17 females (n=34).	36.7±12.6	HFL peak torque is correlated with increased pelvic obliquity (Spearman r_s=0.341, p<0.05), apparent hip adduction (r_s=0.496, p<0.05) and apparent knee valgus (r_s=0.391, p<0.05). HEX peak torque is correlated with apparent hip adduction (r_s=0.482, p<0.05) and apparent knee valgus (r_s=0.417, p<0.05) during single leg squat.	The hip flexor and extensor muscles peak torque can affect hip and knee kinematics during single leg squat.
(Kemp, et al., 2014b)	Cohort study - Chondrolabral pathology	HAD, HAB, HEX, HFL, and HER.	42 males and 42 females with hip chondrolabral pathology (n=84) and a control group.	36±10	Postoperative in comparison with control group (mean of difference [95% CI]): HAD=0.27 [0.14, 0.41] (p<0.001), HAB(women only)=0.46 [0.27, 0.65] (p=0.001), HEX=0.25 [0.10, 0.41] (p=0.001), HFL=0.31 [0.19, 0.42] (p<0.001) and HER=0.09 [0.00, 0.17] (p=0.044) lower in the chondrolabral group (in $N \cdot m.kg^{-1}$).	Reduction in hip muscle strength of the chondrolabral group compared to the control healthy group was observed 12-24 months post hip arthroscopy.
(Nepple, et al., 2015)	Cross section - Before hip arthroscopy	HAB and HFL.	32 male and 18 females with FAI (n=50).	18.1 to 49.8	46% and 42% of the patients had HAB and HFL strength deficits respectively. Bilateral deficits were 8.7% in HAB and 8% in HFL.	Hip strength deficits are present in patients with FAI.

(Philippon, et al., 2009)	Case study - posterior hip subluxation	HER, HAB, HAD, HFL, HIR, HEX, and HAM.	An American football male player.	25	At follow-up (week 9 postoperatively): There was increase in strength of HER=89%, HAB=83%, HAD=80%, HFL=65%, HIR=57%, HEX=23% and HAM=356%. There was a decrease in combined strength of HAB with HER=0.6%.	Improvement in muscle strength was accompanied with improvement in function and return to sport.

Descr., description; M, mean; SD, standard deviation; FAI, femoroacetabular impingement; HAB, hip abduction; HAD, hip adduction; HER, hip external rotation; HIR, hip internal rotation; HFL, hip flexion; HEX, hip extension; HAM, hamstring; CI, confidence interval.

Table 2.4 – Overview of the studies investigating the changes in joint range of motion following hip arthroscopy procedures.

First author (year)	Study Details		Sample Details		Results	Main Findings and Comments
	Descr.	ROM	Number and sex	Age in (M±SD) or range in years		
(Brunner, et al., 2009)	Case series - FAI treatment.	HIR and HFL.	41 males and 12 females (n=53).	17 to 66	[Improvement from preoperative to an average of 2.4 years at follow-up] HIR ROM improved from 6° (range -20° to 45°) to 19° (range -5° to 45°) (p<0.001). HFL ROM improved from 107° (range 60° to 130°) to 122° (range 70° to 145°) (p<0.001).	Significant improvements in ROM. However, hip ROM was measured from a supine position not prone or sitting positions, which is an uncommon position to measure HIR.
(Kemp, et al., 2014b)	Cohort study - Chondrolabral pathology	HIR and HEX.	42 males and 42 females with hip chondrolabral pathology (n=84) and a control group.	36±10	Postoperative ROM in comparison with control group (mean of difference [95% CI]): Reduced HIR=5° [2, 8] (p=0.004) and increased HEX=5° [1, 9] (p=0.019) in the chondrolabral group.	The deterioration of ROM is similar to patients with hip OA which may be caused by morphological changes and weakness in the hip extensors.
(Nho, et al., 2011)	Case series study - FAI treatment.	HFL, HER and HIR.	34 males and 34 females with FAI (n=64).	22.8±6.2	HFL ROM changed from 116.5°±10.4° preoperatively to 120°±0° at follow-up (p=0.02). HER ROM changed from 48.7°±10.1° preoperatively to 49.5°±2.9° at follow-up (p>0.05). HIR ROM changed from 21.7°±10.8° preoperatively to 19.6°±10.4° at follow-up (p>0.05)	No statistically significant changes in HER and HIR ROM after hip arthroscopy.
(Philippon, et al., 2009)	Case study - posterior hip subluxation	HFL, HAD, HAB, HER, HIR and FABER.	An American football male player.	25	At follow-up (week 9 postoperatively): There was an increase of HFL=13°, HAD=13°, HAB=4°, HER=0° and HIR=0° PROM. Minimal difference of HFL=5°, HIR=10° and HER=5° between sides. Reduction in FABER distance=10 cm.	Improvement in most hip PROM and FABER distance. No change in HER and HIR PROM.

Descr., description; M, mean; SD, standard deviation; HAB, hip abduction; HAD, hip adduction; HER, hip external rotation; HIR, hip internal rotation; HEX, hip extension; HFL, hip flexion; PROM, passive range of motion; FABER, flexion-abduction-external rotation; CI, confidence interval.

2.5.4. Changes in Function and Return to Physical Activity

Hip arthroscopy can result in a distinct improvement in function as measured by the MHHS, HOS, and other questionnaires. The change in function after arthroscopy is predicted to last as long as other factors do not come into play such as OA or another pathology. Ninety-six consecutive patients were followed-up at six weeks, three months, six months, and yearly after the arthroscopy. After an average of 9.9 months, the MHHS significantly improved from 60.8 to 82.7 points* (p <0.001) (Table 2.5) (Larson & Giveans, 2008). In a case series study involving eighty-seven patients who had HA for labral repair, the postoperative assessment showed that the MHHS improved from 59.9 ± 1.9 to 74.5 ± 1.9 points at three months (p <0.001) and then to 80.1 ± 1.9 points at six months (p =0.004). However, the level of function reduced slightly between six (80.1 ± 1.9 points) and twelve months (78.7 ± 1.9 points, p >0.05) (Table 2.5) (Dippmann, et al., 2014).

Hip arthroscopy patients were also followed for longer periods to monitor the change in function. Kamath et al. (2009) found a 40% increase in the MHHS scores from 56.8 preoperatively to 80.4 points at 2-years postoperatively (p <0.001). Similar results were reported using the non-arthritic hip questionnaire (NAHS) that showed an increase from 54.4 points (range: 28.8-92.5) to 85.7 points (range: 47.5-100, p <0.001) points at 2.4 years postoperatively in 45 active patients (Table 2.5) (Brunner, et al., 2009). Polesello et al. (2014) also reported that pre-operative MHHS scores significantly improved when patients were followed up for 5-8 years (Pre= 62.4 ± 13.4 points, post= 90.6 ± 11.6 points, p <0.001) (Table 2.5). However, it was previously reported that a change in scores of 67% was seen after the first month, and maximal improvement was seen in the third month, after which the scores were maintained within 5-years of follow-up and plateau afterwards (Byrd & Jones, 2001).

* Further data are not available.

Athletes and active population have also shown significant improvements in function. In a group of mixed athletes (n=47), the MHHS scores increased from 68.6 ± 12.8 to 88.5 ± 17.7 points (p =0.002) and the HOS improved from 78.8 ± 11.3 to 91.4 ± 14.0 points (p =0.03) (Table 2.5) (Nho, et al., 2011). The function of Baseball players measured by the MHHS also improved from 81 to 94 points* postoperatively (p =0.002) (Byrd & Jones, 2015). Another study showed that Australian Football League (AFL) players had significant improvements in their MHHS scores from 83.6 to 98 points and NAHS scores from 85.3 to 97.1 points† postoperatively (p <0.05) (Amenabar & O'Donnell, 2013).

The gain in functional abilities has also been assessed across various age groups. In a sample of sixteen adolescents (age range: 11 to 16), the MHHS significantly improved from 55 (range 33 to 70) to 90 (range 70 to 100) postoperatively (p =0.005). Regarding the HOS, the ADL subscale increased from 58 (range 38 to 75) to 94 (range 74 to100) (p =0.001) and the sport specific subscale (SS) increased from 33 (range 0 to 78) to 89 (range 58 to 100) (p =0.001) (Table 2.5) (Philippon, et al., 2008). Another study also showed a significant increase in the MHHS in a group of sixty paediatric and adolescent patients when followed for two years (Table 2.5) (Philippon, et al., 2012a). Domb et al. (2015a) compared the functional development of the patients aged ≥50 years (age: 50 to 69) to patients aged ≤30 years (age range: 13 to 30). The older patients had significant improvements in MHHS scores (61.2 to 82.2 points), NAHS (59.9 to 80.4 points), HOS-ADL (63.9 to 83 points) and HOS-SS (41.2 to 64.6 points) (all p <0.001). On the other hand, the younger patients gained slightly more functional abilities based on the MHHS scores (62.9 to 84.2 points), NAHS (60.5 to 84.2 points), HOS-ADL (63.1 to 86.5 points) and HOS-SS (42.2 to 72.7 points) (all p<0.001) (Table 2.5).

The percentage of return to sport following arthroscopy in elite players is high. Approximately 95% (n=23 out of 24) of AFL players returned to play at a professional level with sustained improvements in function 1 to 4 years postoperatively (Table 2.5)

* Additional data are not available.
† Additional data are not available.

(Singh & O'Donnell, 2010b; Singh & O'Donnell, 2010a). Another study showed 96% (*n*=25) of AFL players returned to play with 62% (*n*=16) of them still playing after an average of 52.5 months from the surgery (Table 2.5) (Amenabar & O'Donnell, 2013). In a study by Philippon et al. (2007), about 93% (*n*=42 of 45) of the professional players returned to play with 78% (*n*=35 of 45) staying active at a professional level after an average of 1.6 years (range: 0.5-5.5 years). Two of the three players who did not return to playing sports after surgery had an advanced stage of OA (Philippon, et al., 2007a). In a later study by Philippon et al. (2010), all professional hockey players (*n*=28) returned to hockey practice after an average of 3.8 months (range: 1-5) (Table 2.5) (Philippon, et al., 2010).

The high rate of return to sport was also achieved in high school and collegiate athletes. In a case series of 33 mixed level athletes playing various sports, approximately 79% (*n*=26 of 33) were able to return to sports after 9.4 ± 4.7 months after surgery. Out of these players, 92.3 % (*n*=24 of 26) were able to get back to a competitive level. Two years after the surgery, 73% (*n*=24 of 33) of them were still playing sports. The highest percentage of return was 90% for high school athletes (*n*=9 of 10), merely due to the sample size. Professionals came in second with 83% (*n*=5 of 6) and 59% in collegiate athletes (*n*=10 of 17) (Table 2.5) (Nho, et al., 2011). Further information about the studies that examined the return to sport and physical activity can be found in Table 2.5.

2.5.5. Changes in Pain

The change in the level of pain is another important aspect of the patients' follow-up. Pain is commonly assessed using the Visual Analogue Scale (VAS) or Numeric Rating Scale (NRS). The VAS is widely used in hip arthroscopy studies (Dippmann, et al., 2014; Brunner, et al., 2009; Lynch, et al., 2013) and is considered valid and reliable in several clinical situations (McCormack, et al., 1988). The VAS has sufficient construct validity with moderate to high test-retest reliability based on the literacy level of the patient (Bijur, et al., 2001; Hawker, et al., 2011b). The test-retest reliability of the VAS was measured

at several time points at 0 and 1, 30, 31, 60, 61, 90, 91, 120, and 121 minutes. The intra-class correlation for all paired VAS scores was 0.97, 95% CI [0.96, 0.98] (Bijur, et al., 2001). The detailed application of the VAS is described elsewhere (Hawker, et al., 2011b).

Hip arthroscopy can have an essential role in reducing intra-articular pain due to labral repair, osteoplasty* and enhanced stability of the joint. In a study by Dippmann et al. (2014), pain, as measured using the VAS, improved from 5.79 ± 2.42 cm preoperatively to 3.00 ± 2.42 cm at three months (p <0.001), then to 2.26 ± 2.42 cm at six months (p =0.017). However, the pain did not show any significant improvement between 6 and 12 months (2.26 ± 2.42 cm to 2.79 ± 2.42 cm, p >0.05). Similar findings were demonstrated by Brunner et al. (2009), who indicated a significant reduction in pain scores 5.7 (range: 1-9) to 1.5 (range: 0-6) points at an average of 2.4 years of follow-up using the NRS (p <0.001) (Table 2.5). It is also supported by the results of another study that showed a significant drop in pain using the NRS from 5.0 ± 2.7 points to 3.7 ± 3.9 points (p <0.001) after 2 to 5 years (Table 2.5) (Nielsen, et al., 2014). Further studies that assessed postoperative changes in pain can be found in Table 2.5.

Summary

The studies reported significant improvements in function and pain across adolescents, adults and older adults. Hip arthroscopy procedures have been shown to resolve non-arthritic hip pain resulting in a high return to sport or physical activity. Among the players who failed to get back to playing sport, progressive arthritic changes were the main reason.

* Removal of bone tissue.

Table 2.5 - Overview of the studies investigating mainly the changes in function, return to sport and pain following hip arthroscopy procedures.

First author (year)	Study Details		Sample Details		Results	Main Findings and Comments
	Descr.	Meas.	Number and sex	Age in (M±SD) or range in years		
(Amenabar & O'Donnell, 2013)	Case study - Australian football league players undergoing HA.	Return to sport / function.	Males (n=27).	16 to 30	Return to sport: 96% (n=25) of the players went back to sport. After an average of 52.5 months from the HA, 62% (n=16) were still playing and 38% (n=10) retired. Function: MHHS improved from 83.6 to 98 points postoperatively and NAHS improved from 85.3 to 97.1 points postoperatively (p<0.05)	High return to professional level of playing with marked improvement in function.
(Bizzini, et al., 2007)	Case series - surgical decompression (open surgery) of ice hockey professional players.	Return to sport	Ice hockey players (sex NA, n=5).	20 to 22	Return to sport: Full return to training was achieved at 6.7 months (range 5.5 to 9.5 months) and participation in the first competitive game was achieved at 9.6 months (range 7 to 14 months). Three players were able to play in the Swiss ice hockey professional league and two players failed to achieve that level.	Decompression surgery is a safe choice with less return to high level competitive playing. Arthroscopy is still a better option to reduce the recovery time.
(Brunner, et al., 2009)	Case series - FAI treatment.	Function / pain / return to sports	41 males and 12 females (n=53).	17 to 66	[Improvement from preoperative to an average of 2.4 years at follow-up] Function: NAHS improved from 54.4 to 85.7 points (p<0.001). Pain: VAS improved from 5.7 to 1.5 points (p<0.001). Return to sports: Increased activity in hiking (p=0.039), jogging (p=0.031), biking (p=0.001) and aerobics (p=0.008).	Significant improvements in function, pain, and return to sports.
(Byrd & Jones, 2009)	Case series	Function / return to sport / survivorship	11 males and 4 females (n=15).	14 to 70	[Improvement from preoperative to 120 months postoperatively] Function: MHHS improved from 51 to 96 points (p=0.004). Return to sport: 87% returned to previous level of sport. Survivorship: 1 patient (6.7%) had 2 subsequent procedures within 18 months, and 5 patients (33.3%) had THA at an average of 73 months.	Successful outcome of hip arthroscopy procedures with high return to sport.

(Byrd & Jones, 2015)	Case series - Baseball players.	Function / return to sports / survivorship	Males (n=41).	18 to 34	Function: MHHS improved from 81 preoperatively to 94 points postoperatively (p=0.002). Return to sport: 95% (n=42 procedures) returned to Baseball after an average of 4.3 months. Approximately 90% went to play at previous level of competition. Survivorship: 3 players (7.3%) had repeated arthroscopy.	Hip arthroscopy has a high success rate in helping Baseball players in returning to play.
(Casartelli, et al., 2014)	Case series - symptomatic FAI	Pain	3 males and 5 females (n=8).	29±10	Preoperative: median hip score= 2mm - 14mm. Postoperative: significant reduction in hip pain (p<0.05). Follow-up: no pain and 4 patients (50%) felt better and were able to return to sport.	Significant improvement in hip pain postoperatively.
(Daivajna, et al., 2015)	Case series study – patients with advanced OA.	Function	63 males and 14 females (total n=77)	19 to 64	Function: MHHS improved from 58 (range 28-87) to 66 (range 28-91) (p=0.003) and NAHS improved from 64 (range 27-93) to 74 (range 30-100) (p=0.0001) at the first year. MHHS improved to 72 (range 47-91) (p=0.002) and NAHS to 77 (59 to 95) (p=0.0003) at the second year.	Patients with advanced OA show good progress postoperatively.
(Dippmann, et al., 2014)	Case series - patients with labral repair for FAI	Function / pain	32 males and 55 females (n=87).	Males: 17 to 63, and females: 15 to 59.	[Improvement from preoperative to postoperative] Function: MHHS improved from 59.9±1.9 to 74.5±1.9 points at 3 months (p<0.001), from 74.5±1.9 at 3 months to 80.1±1.9 points at 6 months (p=0.004), and no significant improvement between 6 and 12 months (p>0.05). Pain: VAS improved from 57.9±2.6 to 30.0±2.6 mm at 3 months (p<0.001), from 30.0±2.6 at 3 months to 22.6±2.6 mm at 6 months (p=0.017), and no significant improvement between 6 and 12 months (p>0.05).	Function and pain significantly improved within the 1 year after hip arthroscopy. Marked changes are seen typically up to 6 months postoperatively.

(Domb, et al., 2015a)	Cohort study - patients aged ≥50 years old (group 1) vs. patients aged ≤30 years old (group 2).	Function / survivorship	18 males and 34 females (each group *n*=52).	Group 1: 50 to 69 and group 2: 13 to 30.	[Improvement from preoperative to an average of 2 years at follow-up] Function: Group 1: MHHS= 61.2 to 82.2 points, NAHS=59.9 to 80.4 points, HOS-ADL=63.9 to 83 and HOS-SS=41.2 to 64.6 points. Group 2: MHHS= 62.9 to 84.2 points, NAHS=60.5 to 84.2 points, HOS-ADL=63.1 to 86.5 and HOS-SS=42.2 to 72.7 points (all improvements in both groups $p<0.001$). Survivorship: Group 1=82.7% and group 2=98.1%.	Both groups showed marked improvement postoperatively with less survivorship rate in the patients aged ≥50 years.
(Kamath, et al., 2009)	Retrospective case series study looking at the factors that affect return to activity.	Function / survivorship	20 males and 32 females (*n*=52).	25 to 76	Function: 40% improvement in MHHS from 56.8 preoperatively to 80.4 postoperatively ($p<0.001$) with 56% showing good or excellent outcomes. Survivorship: Three patients (5.8%) (BMI=30.8, DoS=22 months, aged 44 to 54 years) had THA, two of them had osteoarthritis prior to the procedure.	Several factors can positively or negatively affects the return to sport after hip arthroscopy.
(Larson & Giveans, 2008)	Case series - FAI treatment.	Function / pain / survivorship	54 males and 42 females (*n*=96).	34.7±NA	[Improvement from preoperative to an average of 9.9 months at follow-up] Function: MHHS=60.8 to 82.7 points and Short form 12=60.2 to 77.7 points. Pain: VAS=6.74 to 1.88 cm (all $p<0.001$). Survivorship: 3 hips (3.1%) had THA.	Hip arthroscopy results in significant improvements in function and pain.
(Lindner, et al., 2013)	Case series – patients older than 50 years	Function	17 male and 40 female in group 1 and group 2 each (total=57 each).	Group 1: 50 to 76. Group 2: 14 to 30.	[Improvement from preoperative to two year follow-up] Group 1: MHHS from 64 to 89, NAHS from 60 to 88, HOS-ADL from 67 to 90, HOS-SS from 42 to 77 (all $p<0.05$). Group 2: MHHS from 55 to 84.5, NAHS from 53 to 80, HOS-ADL from 58 to 83, HOS-SS from 35 to 62 (all $p<0.05$). Differences between the two groups $p=0.03$.	Patients aged ≥50 years showed similar improvement in function to patients aged ≤30 years.
(Londers & Van Melkebeek, 2007)	Retrospective case series study.	Satisfaction	37 males and 19 females (*n*=56).	17 to 59	Satisfaction: 80% (*n*=45) of the patients were satisfied and would have the procedure again.	High satisfaction rate after hip arthroscopy.

(Nho, et al., 2011)	Case series - FAI treatment.	Function / return to sports	34 males and 13 females (n=47).	22.8±6.2	Function: MHHS improved from 68.6±12.8 preoperatively to 88.5±17.7 postoperatively (p=0.002). HOS improved from 78.8±11.3 preoperatively to 91.4±14.0 postoperatively (p=0.03). Return to sports: 79% returned to play with 92.3% of them playing at the same level. At 2 years postoperatively, 73% were able to maintain playing sport with 90% high school athletes, 59% collegiate athletes and 83% professional athletes returning to sports. Five patients were unable to return to play due to hip pain.	Significant improvement in function and ability to return to sport after the arthroscopy procedure.
(Nielsen, et al., 2014)	Case series study - FAI treatment.	Function / pain / survivorship	48 males and 69 females (n=117).	15 to 70	[Improvement from preoperative to an average of 13 months at follow-up] Function: MHHS improved from 72.1±16.8 to 83.1±16.9 points (p<0.001), HOS improved from 71.4±17.0 to 83.4±17.4 points (p<0.001). Pain: NRS improved from 5.0±2.7 to 3.7±2.9 points (p<0.001). Survivorship: 4.3% (n=5) had THA (mean age 48) with advanced cartilage damage.	Significant improvement in function following hip arthroscopy.
(Philippon, et al., 2007a)	Case series - athletes with FAI.	Return to sport / survivorship	42 male and 3 female athletes (n=45).	17 to 61	Return to sport: 93% (n=42) returned to professional level of playing after the arthroscopy, while 7% (n=3) did not return to play due to presence of OA. Overall, 78% (n=35) remained active at a professional level after an average of 1.6 years after the arthroscopy. Survivorship: Re-operation was performed in 11% (n=5) of the athletes, with two of the five athletes not returning to play because of the OA.	A high return of athletes to playing hockey, golf, American football and other sports after hip arthroscopy.

(Philippon, et al., 2008)	Case series - adolescents with FAI.	Function / satisfaction	2 male and 14 female adolescents (n=16).	11 to 16	[Improvement from preoperative to an average of 1.36 years at follow-up] Function: MHHS improved from 55 (range 33 to 70) to 90 (range 70 to 100) (p=0.005). HOS-ADL improved from 58 (range 38 to 75) to 94 (range 74 to100) (p=0.001). HOS-SS improved from 33 (range 0 to 78) to 89 (range 58 to 100) (p=0.001). Satisfaction: satisfaction score was 9 (range 9 to 10).	Hip arthroscopy can significantly improve function in adolescents.
(Philippon, et al., 2009)	Case study - posterior hip subluxation.	Return to sport	An American football male player.	25	Preoperative: 0 out of 20 in the Sport Cord Agility Test due to pain and poor mechanics. Postoperative: 20 out of 20 in the same test.	Significant improvement in function and ability to return to sport after the arthroscopy procedure.
(Philippon, et al., 2010)	Case series - hockey players with FAI.	Function / return to sport	Male Hockey players (n=28).	18 to 36	[Improvement from preoperative to an average of 24 months at follow-up] Function: MHHS improved from 70 (range 56 to 100) to 95 (range 74 to 100) (p=0.001). Return to sport: All players returned to hockey drills at 3.8 months (range 1 to 5).	High improvement in function and return to play in Hockey players after hip arthroscopy.
(Philippon, et al., 2012b)	Retrospective case series study.	Function / survivorship	Older patients (n=153).	≥50	[Improvement from preoperative to postoperative] Function: MHHS improved from 58 to 84 (p=0.001), HOS-SS improved from 42 to 72 (p=0.001) and the physical component of the Short Form-12 improved from 38 to 49 (p=0.001). Survivorship: 20% (n=31) had THA. At 3 years postoperatively, patients with >2mm of joint space had a survivorship rate of 90% (out of 64 patients) and those with ≤2mm had a survivorship rate of 57% (p=0.001).	Significant improvement in function in patients aged 50 years or more. Survivorship was lower with reduced joint space.
(Philippon, et al., 2012a)	Cohort study - paediatric and adolescent patients.	Function	19 male and 41 female paediatric and adolescent patients (n=60).	11 to 16	[Improvement from preoperative to a minimum of 2 years at follow-up] Function: MHHS increased from 57 (95% CI, [51, 62]) to 91 (95% CI, 88 to 94) (p<0.001).	Hip arthroscopy can significantly improve function in paediatrics and adolescent patients.

(Polesello, et al., 2014)	Case series.	Function / return to play / survivorship	21 males and 3 females (*n*=24).	34.6±9.5	[Improvement from preoperative to an average of 6.1 years at follow-up] Function: MHHS improved from 62.4±13.4 to 90.6±11.6 points (p<0.001). Return to play: 10 out of 14 active patients returned to play (71.4%) Survivorship: three patients (16.6%) had THA and one had another procedure. Satisfaction: all patients were satisfied with the procedure.	High return to play, satisfaction, and improved function at a minimum of 5 years following hip arthroscopy.
(Redmond, et al., 2015a)	Retrospective study - minimum of 2-year follow up after HA.	Survivorship	Patients (sex NA, *n*=30).	63.9±NA	Survivorship: 63% survivorship rate at 2-years follow up with 9 patient (30%) having THA after an average of 1.1 years and 2 patients having re-surgery (6.7%) (Total *n*=11). The 9 patients who had THA had lower preoperative MHHS scores (p=0.015), lower preoperative HOS-ADL (p=0.01), higher pain scores (p=0.05), higher acetabular inclination (p=0.023) and more sever chondral damage (p=0.033).	Reduced survival rate in patients aged >60 years old.
(Sansone, et al., 2015)	Case series - top-level athletes	Function / return to play / survivorship / satisfaction.	68 male and 17 female athletes (*n*=85).	25±5	[Improvement from preoperative to an average of 12.3 months at follow-up] Function: iHOT-12 from 42±18 to 73±24, HAGOS-ADL from 66±22 to 86±17, HAGOS-Sport from 39±21 to 75±23, HAGOS-PA from 27±28 to 70±30, and HAGOS-QoL from 34±21 to 67±26 points (all p<0.0001). Return to play: athletes who had symptoms of <12 months (87%) returned to sport on a higher rate compared to athletes with >12 months (61%) duration of symptoms (p<0.05). Survivorship: three procedures (4%) were performed again due to impingement (*n*=2) and adhesions (*n*=1). Satisfaction: 79 (93%) were satisfied with the procedure.	Excellent outcome of hip arthroscopy procedure in top-level athletes.

(Sansone, et al., 2016a)	Case series - patients with mild to moderate OA.	Function / survivorship / satisfaction	77 males and 23 females (*n*=100).	47±10	[Improvement from preoperative to 24 months at follow-up] Function: iHOT-12 from 42±18 to 65±27, HAGOS-ADL from 56±23 to 75±23, HAGOS-Sport from 40±21 to 61±30, HAGOS-PA from 33±26 to 56±33, and HAGOS-QoL from 31±16 to 55±28 points (all $p<0.0001$). Survivorship: 10 procedures were performed again with fascia latae release (*n*=1, 1.3%) and THA (*n*=9, 11.3%). Satisfaction: 56 of 68 (82%) patients were satisfied and 10 (14%) patients were dissatisfied with the procedure.	Hip arthroscopy is an effective procedure in improving the function of patients with mild to moderate OA.
(Sansone, et al., 2016b)	Case series - patients with FAI.	Function / survivorship / satisfaction.	190 males and 99 females (*n*=289).	37±13	[Improvement from preoperative to 25 months at follow-up] Function: iHOT-12 from 43±17 to 66±27, HAGOS-ADL from 60±22 to 78±22, HAGOS-Sport from 40±20 to 65±29, HAGOS-PA from 29±26 to 57±34, and HAGOS-QoL from 33±18 to 58±29 points (all $p<0.0001$). Survivorship: 17 reoperations (5%) and 14 THA procedures (4%) after the arthroscopy. Satisfaction: 236 of 289 (82%) patients were satisfied, 38 (13%) patients were dissatisfied with the procedure.	Good outcome of hip arthroscopy in patients with FAI.
(Tran, et al., 2013)	Case series - adolescent patients with cam FAI.	Function / satisfaction.	29 males and 6 females (*n*=35).	11 to 18	[Improvement from preoperative to follow-up at 1-2 years] Function: MHHS improved from 77.39±18.68 to 94.15±11.99 points ($p<0.0005$) and NAHS improved from 76.34±16.76 to 93.18±13.28 points ($p<0.0005$). Satisfaction: 88.2% (*n*=30) were satisfied and 91.2% (*n*=31) would have the procedure again.	Significant improvement in function in adolescent patients with high satisfaction rate.

Descr., description; Meas., measurement; M, mean; SD, standard deviation; FAI, femoroacetabular impingement; OA, osteoarthritis; MHHA, modified Harris Hip Score; VAS, visual analogue scale; NRS, numeric rating scale; NAHS, non-arthritic hip score; HOS, hip outcome score; iHOT, international hip outcome tool; HAGOS, Copenhagen hip and groin outcome score; -ADL, activities of daily living subscale; -PA, physical activity subscale; -QoL, quality of life subscale; -SS; sport specific subscale; THA, total hip arthroplasty.

2.5.6. *Factors Affecting the Outcomes of Hip Arthroscopy*

The outcomes of hip arthroscopy can vary among different populations. Surgeons and physiotherapists noticed that some factors could have an impact on the rate and quality of improvement following HA. The factors that play a role in postoperative progress include body mass index, surgical technique, age, chondropathy, and the presence of a degenerative joint disease (Figure 2.8).

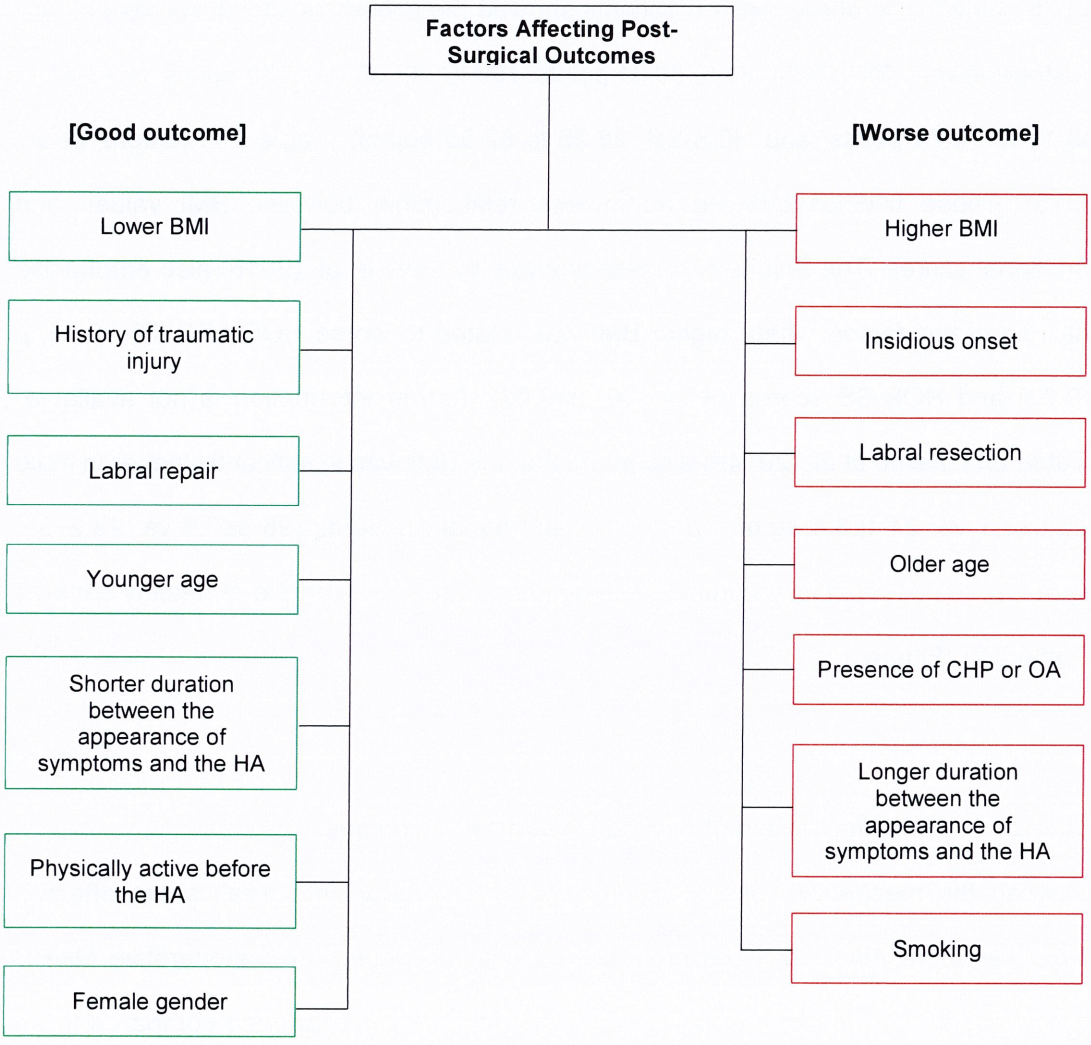


Figure 2.8 – The factors affecting post-surgical outcomes (Green and red boxes refer to the factors resulting in good and worse outcomes respectively; BMI, body mass index; HA, hip arthroscopy; CHP, chondropathy; OA, osteoarthritis).

2.5.6.1 Body Mass Index

A cohort study included three groups of non-obese (BMI <30 $kg.m^{-2}$), obese class-I (BMI 30 to 34.9 $kg.m^{-2}$) and obese class-II patients (BMI 35 to 39.9 $kg.m^{-2}$). Improvement from preoperative to follow-up at two years showed that the non-obese patients had high preoperative and postoperative function scores (MHHS: 63.41 to 83.81 points, NAHS: 60.86 to 83.62, HOS-ADL: 66.24 to 86.24 points, and HOS-SS: 44.01 to 73.26 points). In contrast, the obese class-I patients had lower scores (MHHS: 54.81 to 75.95 points, NAHS: 48.98 to 72.51 points, HOS-ADL: 53.22 to 72.99 points, and HOS-SS: 30.56 to 60.75 points). The obese class-II patients showed the lowest functional scores pre and postoperatively (MHHS: 50.81 to 80.01 points, NAHS: 42.36 to 72.50 points, HOS-ADL: 48.11 to 74.73 points, and HOS-SS: 28.25 to 62.56 points) (Table 2.7) (Gupta, et al., 2015). These findings indicate an inverse relationship between BMI values and functional scores. The results of a meta-analysis by Levy et al. (2016) also emphasise the role of this factor, where higher BMI was related to worse HOS-ADL (R^2 =0.35, p =0.02) and HOS-SS scores (R^2 =0.30, p=0.03) (further information is not available) (Table 2.7). Kemp et al. (2014b) also found that the BMI was significantly higher in male (25.6±3.1 vs. 24.1±3.4 $kg.m^{-2}$, p <0.001) and female patients (26.6±7.8 vs. 23.2±3.4 $kg.m^{-2}$, p <0.001) with joint cartilage damage compared to a sample of healthy controls (Table 2.7) (Figure 2.8).

2.5.6.2 Mechanism of Injury and Surgical Technique

Knowing the mechanism of injury can direct the surgeon into treating the affected structures more efficiently. The prognosis of patients based on postoperative MHHS scores is better if a history of traumatic injury to the hip was reported compared to an insidious onset (Byrd & Jones, 2001; Schilders, et al., 2011). Also, failure to identify and treat underlying FAI negatively impacts survivorship (McCarthy & McMillan, 2013).

The surgical technique and approach towards treating labral pathology were also found to influence the outcome. The prognosis and recovery of patients with labral resection

were significantly worse at two years postoperatively (only 28% had excellent outcomes) compared to the patients with labral fixation (80% had excellent outcomes). Moreover, radiographic signs of OA were statistically more evident in the resection group of patients at one year (p =0.02) and two years (p =0.009) compared to the fixation group (Table 2.7) (Espinosa, et al., 2006; Espinosa, et al., 2007). When patients without OA (Tönnis grade* <2) were excluded, those who had labral repair showed better postoperative improvement in the MHHS scores (preoperative= 60.23 [range 24-85]; postoperative= 93.59 [range 55-100]) compared to patients who had labral resection (preoperative= 62.78 [range 29-96]; postoperative= 88.84 [range; 35-100], p =0.036) (Table 2.7) (Schilders, et al., 2011). Therefore, preservation of labral function (reconstruction) has a positive effect on survivorship and functional scores (Figure 2.8) (McCarthy & McMillan, 2013; Domb, et al., 2014).

2.5.6.3 Age

Several studies linked age with survivorship† and conversion to total hip arthroplasty (THA) (Bogunovic, et al., 2013; McCarthy, et al., 2011; Philippon, et al., 2012b). In a cohort study by Domb et al. (2015b), patients aged ≥50 years (range 50 to 69) had a survivorship rate of 82.7% compared to 98.1% in patients aged ≤30 years (range 13 to 30) (Table 2.7). In another study, this rate was lower at about 70% in 30 patients aged >60 years old, with 30% conversion rate to THA after two years from the HA. Those who had THA had lower pre-operative MHHS scores (p =0.015), higher pain scores (p =0.05) and more severe chondral damage (p =0.033) compared to those who did not convert (Redmond, et al., 2015a). These conclusions are also supported by the findings of Nielsen et al. (2014), who found that patients aged >40 years had significantly different MHHS (p =0.02) and HOS (p =0.01) scores compared to patients aged <40 years old (Table 2.7) (Figure 2.8).

---

* Explained in section 2.5.

† Referring to the extent of success before re-operation or shifting to THA.

Preoperative age of the patient can be considered an indicator of postoperative outcome. McCormick et al. (2012) pointed out that being less than 40 years old is a predictor of good postoperative outcome (OR=7, 95% CI [2.9, 16.9], p <0.0001) (Table 2.7). Levy et al. (2016) performed a meta-analysis and concluded that increased age was significantly related to worse postoperative MHHS scores (R^2 =0.14, p <0.01), HOS-SS scores (R^2 =0.12, p =0.05) and rates of reoperation (R^2 =0.08, p =0.02) (Table 2.7). Despite this, other researchers believe that sex and age as independent factors do not significantly affect survivorship after HA (McCarthy & McMillan, 2013) (Figure 2.8).

2.5.6.4 Chondropathy and Osteoarthritis

The probability of having chondropathy (CHP) and OA increases with age, but are not thought to be causally related. In seventy-two out of 100 arthroscopy patients, mild chondropathy (Outerbridge I-II) was evident during the arthroscopy procedures of 44 patients (61%), while severe chondropathy (Outerbridge III-IV) was found in 28 patients (39%) (Table 2.6). Based on age, chondropathy was present in 61% of the patients aged between 18 and 35 years old, whereas it was present in 100% of the patients aged over 50 years (Table 2.7) (Kemp, et al., 2014a) (Figure 2.8).

The presence of pre-operative OA was a negative predictor of good postoperative outcome (OR =2.5, 95% CI [1.2, 5.3]) (McCormick, et al., 2012). Patients with pre-operative OA often don't return to high-level competitive sports; they are more likely to stick to low impact activities and convert to THA (Philippon, et al., 2007a; McCarthy & McMillan, 2013). Kemp et al. (2014a) revealed that patients with severe CHP had significantly worse functional outcomes compared to patients with mild CHP (HOOS-S, MD= 11, 95% CI [1, 2], p =0.02; HOOS-A, MD= 9, 95% CI [0, 18], p =0.042; HOOS-Q, MD= 17, 95% CI [4, 29], p =0.004; iHOT-33, MD= 12, 95% CI [1, 23], p =0.034). They also showed worse outcomes compared to the patients with moderate CHP (HOOS-S, MD= 12, 95% CI [3, 21], p =0.02; HOOS-P, MD= 11, 95% CI [3, 19], p =0.007; HOOS-A, MD= 12, 95% CI [5, 19], p =0.002; HOOS-Sp, MD= 19, 95% CI [8, 30], p =0.001;

HOOS-Q= 19, 95% CI [7, 31], p =0.002; iHOT-33, MD= 15, 95% CI [5, 25], p =0.004) (Table 2.7) (Figure 2.8).

Table 2.6 - Outerbridge classification (BlueCross, 2016).

Grade	Description
0	Normal
I	Cartilage with softening and swelling
II	A partial-thickness defect with fissures on the surface that do not reach subchondral bone or exceed 1.5 cm in diameter
III	Fissuring to the level of subchondral bone in an area with a diameter more than 1.5 cm
IV	Exposed subchondral bone

Mild to moderate OA will not necessarily progress to a more severe form, and the extent of success is not necessarily influenced by age (Mardones, et al., 2010; Daivajna, et al., 2015). Hip arthroscopy is still a valid option for pain relief caused by degenerative labral tears considering that the intra-articular damaged structures could be a factor in progressing the degenerative process (Mardones, et al., 2010). However, the presence of CHP and OA will reduce the percentage of success, and a THA may be required within a few years.

2.5.6.5 Combined Factors and Predictors

The outcome of a study aimed at examining failed hip arthroscopy procedures that required revision* arthroscopy or THA showed that the presence of multiple factors could affect the prognosis. The revision arthroscopy patients were younger (age: 28 ± 10 vs. 50 ± 7 years old, p <0.005) and had lower BMI (25 ± 6 vs. 28 ± 4.5 kg.m^{-2}, p =0.034) in comparison to the patients who had THA. They also had a lower mean duration between the arthroscopy and the second procedure (25 months, range 2.9-84 vs. 31 months, range 1.9-70, p >0.05) and were predominantly females (79% vs. 41%, p =0.009). The most common cause of revision arthroscopy was residual FAI (n=26 of 60), while it was osteoarthritis for THA (n=22 of 60) (Table 2.7) (Bogunovic, et al., 2013). Kamath et al.

* Performing a second hip arthroscopy procedure to treat unresolved pain.

(2009) determined the positive predictors of postoperative outcome to be female sex (OR =1.531, p =0.512), a left sided procedure (OR =1.406, p =0.619), duration of symptoms of >18 months (OR =1.228, p=0.749), and preoperative activity level (OR =2.534, p =0.002). Whereas the negative predictors of outcome were having a traumatic injury (OR =0.118, p =0.004), age of >45 years (OR =0.458, p =0.237), a BMI of >25 (OR =0.309, p =0.086) and smoking (OR =0.257, p=0.089) (Table 2.7). The presence of OA in the form of chondropathy can also be a predictor of worse MHHS scores postoperatively (OR =2.5, 95% CI, [1.2, 5.3], p =0.02). In contrast, younger age of <40 years can be a predictor of a good to an excellent outcome (OR =7, 95% CI, [2.9, 16.9], p <0.001) (McCormick, et al., 2012) (Figure 2.8).

2.5.6.6 Other Factors

The pre-operative duration of symptoms has been linked with the postoperative time and rate of returning to play and function. Philippon et al. (2010) found a correlation between the duration of symptoms and time to go back to play (r= 0.582, p =0.001). The players who had the procedure performed within one year of the injury returned to playing sport after three months postoperatively, while those who waited more than one year returned to playing sport at an average of 4.1 months (Table 2.7). Sansone et al. (2015) also concluded that athletes with a short duration of symptoms (<12 months) returned to sport at a higher rate (87%) compared to the athletes with a longer duration of symptoms (>12 months). A lower rate of 61% was reported in the latter group (p<0.05) (Table 2.7). In terms of function, the duration of symptoms was found to negatively correlate with the iHOT-12 (r= -0.189, p =0.012) and HAGOS-QoL (r= -0.209, p=0.012) (Table 2.7) (Sansone, et al., 2016b).

Being active with good functional scores before the arthroscopy can also affect the prognosis. The Sports Frequency Score (SFS) was found to correlate with the NAHS (r =0.392, p =0.003) and with the VAS (r =0.242, p =0.049) (Brunner, et al., 2009). Moreover, the amount of change in the scores of function such as the MHHS and NAHS

can be used as a predictor of outcome. Patients who score higher in the first few years after the surgery have a better chance of avoiding or delaying THA (Table 2.7) (Byrd & Jones, 2009). A summary of the studies looking into the factors that affect postoperative outcome can be found in Figure 2.8 and Table 2.7.

Summary

A variety of factors were found to impact postoperative improvements. Lower BMI, age and duration of symptoms can positively influence the surgical outcome. In contrast, performing labral resection and the presence of OA and chondropathy may result in a worse outcome. These factors should be considered when planning the best approach to treating intra-articular pain regarding labral preservation, performing arthroscopy or total hip arthroplasty.

Table 2.7 - Overview of the studies investigating the factors affecting the outcomes of hip arthroscopy procedures.

First author (year)	Study Details		Sample Details		Results	Main Findings and Comments
	Descr.	Factor	Number and sex	Age in (M±SD) or range in years		
(Bogunovic, et al., 2013)	Cohort study – patients who had revision arthroscopy and those who had THA.	Survivorship	19 males and 39 females.	15 to 71	Patients who had revision HA vs patients who had THA: younger in age (28±10 vs 50±7 years, $p<0.005$), had lower average BMI (25±6 vs 28±4.5 kg.m^{-2}, $p=0.034$) and had more females (79% vs 41%, $p=0.009$).	Patients who have revision HA are significantly different from those who have THA.
(Byrd & Jones, 2009)	Case series study.	DJD	11 male and 4 females (n=15).	14 to 70	When followed up for 120 months post-arthroscopy, patients with arthritis did much poorer ($p<0.006$) than the others and 5 patients had THA at an average of 73 months.	Pre-existing history of DJD negatively affects the outcome and survivorship of arthroscopy procedures.
(Daivajna, et al., 2015)	Case series study.	DJD	63 males and 14 females (total n=77)	19 to 64	[Improvement from preoperative to an average of 2.8 years at follow in patients with advanced OA] Survivorship: 44% (n=34) required THA at an average of 18 months after HA.	High percentage of transfer to THA in patients with advanced OA.
(Degen, et al., 2016)	Retrospective case series - simultaneous vs staged HA.	Function / surgical technique	81 patients (162 hips, staged bilateral HA) and 12 patients (24 hips, simultaneous bilateral HA).	NA	[Improvement between simultaneous versus staged HA at an average follow-up of 17.8 months] MHHS: 90.8±11 vs. 88.9±12.5 points ($p=0.662$), HOS-ADL=97.3±3.8 vs. 92.6±10.3 points ($p=0.057$), HOS-SS: 93.3±10.2 vs. 86.5±16.6 ($p=0.203$).	Similar improvement in function was seen in simultaneous and staged bilateral hip arthroscopy patients.

(Domb, et al., 2015b)	Cohort study - patients aged ≥50 years old (group 1) vs. patients aged ≤30 years old (group 2).	Age	Each group had 18 males and 34 females (each *n*=52).	Group 1: 50 to 69 and group 2: 13 to 30.	Survivorship: Group 1=82.7% and group 2=98.1%.	Less survivorship rate in the patients aged 50 years or above. Hip arthroscopy should be used to treat hip problems in patients with Tönnis arthritic grade of 0 or 1.
(Espinosa, et al., 2006)	Retrospective case series study looking at the surgical technique of HA.	Surgical technique	Group 1 (*n*=25): labral resection and group 2 (*n*=30): labral refixation (sex NA).	NA	Group 1 (resection group): 28% had excellent results, 48% had good results, 20% had moderate results, and 4% had poor results. Group 2 (refixation group): 80% had excellent results, 14% had good results, and 6% had moderate results. Clinical outcome: better outcome for group 2 at 1 year ($p<0.001$) and at 2 years ($p=0.01$). Radiographic signs of OA were more evident in group-1 at 1 year ($p=0.02$) and 2 years ($p=0.009$) postoperatively.	Performing labral refixation shows superior clinical outcome and results compared to labral resection with less radiographic signs of OA.
(Gupta, et al., 2015)	Cohort study on three groups.	BMI	Group 1 (non-obese, *n*=562, BMI <30 $kg.m^{-2}$), group 2 (obese class-I, *n*=94, BMI 30 to 34.9 $kg.m^{-2}$) and group 3 (obese class-II, *n*=24, BMI 35 to 39.9 $kg.m^{-2}$)	Group 1 (34.78±NA), group 2 (44.02±NA) and group 3 (39.33±NA)	[Improvement from preoperative to follow-up at 2 years] Group 1: MHHS=63.41 to 83.81, NAHS=60.86 to 83.62, HOS-ADL=66.24 to 86.24, HOS-SS=44.01 to 73.26. In group 2: MHHS=54.81 to 75.95, NAHS=48.98 to 72.51, HOS-ADL=53.22 to 72.99, HOS-SS=30.56 to 60.75. In group 3: MHHS=50.81 to 80.01, NAHS=42.36 to 72.50, HOS-ADL=48.11 to 74.73, HOS-SS=28.25 to 62.56.	The obese patients started with lower function scores and had lower postoperative scores as well compared to the non-obese patients. The higher the BMI, the lower the scores would be.

(Kamath, et al., 2009)	Retrospective case series study looking at the factors that affect return to activity.	Sex, side, duration of symptoms, activity level, trauma, age, BMI, DJD and smoking.	20 males and 32 females (n=52).	25 to 76	Positive predictors of outcome: Female sex (OR=1.531, p=0.512), left sided procedure (OR=1.406, p=0.619), duration of symptoms >18 months (OR=1.228, p=0.749), and preoperative activity level (OR=2.534, p=0.002). Negative predictors of outcome: Traumatic injury (OR=0.118, p=0.004), age>45 years (OR=0.458, p=0.237), BMI>25 (OR=0.309, p=0.086) and current smoker (OR=0.257, p=0.089) [OR>1 indicate positive outcome and OR<1 indicate negative outcome]. Two patients with OA and BMI=30.8 had THA due to failure of the arthroscopy.	Several factors can positively or negatively affects the return to sport after hip arthroscopy. Presence of DJD and high BMI reduces the chance of survivorship.
(Kemp, et al., 2014a)	Cohort study - Patients with chondropathy	CHP	Group 1 (no CHP, n=28): 16 males and 12 females. Group 2 (mild CHP, n=44): 23 males and 21 females. Group 3 (severe CHP, n=28): 12 males and 16 females.	Group 1= 31±10, group 2=35±12 and group 3=41±10	Group 3 had worse scores than group 1 in HOOS-S (MD and 95% CI) = 11[1, 2] (p=0.02), HOOS-A= 9[0, 18] (p=0.042), HOOS-Q= 17[4, 29] (p=0.004) and iHOT-33=12[1, 23] (p=0.034). Group 3 had worse scores than group 2 in HOOS-S= 12[3, 21] (p=0.02), HOOS-P= 11[3, 19](p=0.007), HOOS-A= 12[5, 19] (p=0.002), HOOS-Sp= 19[8, 30](p=0.001), HOOS-Q= 19[7, 31](p=0.002) and iHOT-33= 15[5, 25](p=0.004).	Severe chondropathy is associated with worse patients reported outcomes 24 months following hip arthroscopy.
(Kemp, et al., 2014b)	Cohort study - Patients with chondrolabral pathology.	BMI / sex	42 males and 42 females with hip chondrolabral pathology (n=84).	36±10	BMI: Chondrolabral group and males had higher body mass (p<0.001) wider waist girth (p<0.001) compared to healthy group and females. Sex (mean of difference [95% CI]): Females had lower HER ROM=12 [7, 16] (p<0.001), and lower strength of HAB=0.46 [0.27, 0.65] (p=0.001), HAD=0.36 [0.18, 0.54] (p<0.001), HEX=0.23 [0.03, 0.43] (p=0.033), HER=0.27 [0.15, 0.38] (p<0.001) and HIR=0.17 [0.08, 0.25] (p<0.001) (all in $N \cdot m.kg^{-1}$).	The deterioration of ROM is similar to patients with hip OA, which may be caused by morphological changes and weakness in the hip extensors.

(Levy, et al., 2016)	Meta-analysis	Age / BMI	81 studies (9,317 hips)	NA	Age: bivariate analysis showed that increased age is significantly related to worse postoperative MHHS (R^2=0.14, p<0.01), HOS-SS (R^2=0.12, p=0.05) and rates of reoperation (R^2=0.08, p=0.02). BMI: increased BMI was related to worse HOS-ADL (R^2=0.35, p=0.02) and HOS-SS (R^2=0.30, p=0.03).	Higher age and BMI are associated with worse postoperative outcome.
(Lindner, et al., 2013)	Case series – patients older than 50 years		17 male and 40 female in group 1 and group 2 each (total=57 each).	Group 1: 50 to 76. Group 2: 14 to 30.	[Patients were followed up for 2 years] Survival: An overall rate of 82% for group 1 with 18.4 months between HA and THA. Tönnis grade 0-1 had a survivorship rate of 71% (32 of 45) and 50% (6 of 12) for Tönnis grade 2.	Patients with Tönnis grades 0-1 had a higher survival rate compared to those at grade 2. The overall survivorship at two years after HA in patients aged >50 years is high.
(Mardones, et al., 2010)	Case series	CHP	15 patients (sex NA)	60 to 70	CHP: Complex acetabular chondral defects, Outerbridge grade III-IV and a subchondral cyst ≥5 mm were associated with poor outcome.	Evidence of cartilage damage results in poor prognosis.
(McCarthy, et al., 2011)	Retrospective case series study	Survivorship / DJD	47 males and 64 females (total n=111)	36±13	Survivorship: Average time to THA was 4.8±4.4 years. Estimated survivorship was 91% at 1-year, 75% at 5-years, and 63% at 10-years. OR=3.6 for patients >40 years and OR=20-58 for patients with Outerbridge grades III-IV to have THA.	Older age and evidence of DJD increases the odds of transferring to THA.
(McCormick, et al., 2012)	Retrospective case series study	DJD / age	Patients (sex NA, n=125).	17.3 to 62.8	DJD: presence of OA changes during the HA procedures was a predictor of negative outcome (OR=2.5, 95% CI [1.2, 5.3], p=0.02). Age: younger age (<40 years) is a predictor of good outcome (OR=7, 95% CI [2.9, 16.9], p<0.0001)	Presence of DJD before the HA (OA in this case) has a direct negative effect on postoperative outcome. Younger patients had better surgical outcome than the older patients.

(Nielsen, et al., 2014)	Case series - FAI treatment.	Age / CHP	48 males and 69 females (n=117).	15 to 70	Age: patients aged >40 years were significantly different from those aged <40 years in MHHS (p=0.02), HOS (p=0.01) but not in pain (p=0.25). Surgical technique: No difference between patients having labral refixation and those having labral resection in MHHS (p=0.60), HOS (p=0.28) and pain (p=0.72). CHP: Patients who had extensive cartilage damage had THA (4.3%, n=5).	Advanced age and extensive chondral damage do play a role in postoperative improvement and shifting to THA.
(Philippon, et al., 2007a)	Case series - athletes with FAI	DJD	42 male and 3 female athletes (n=45).	17 to 61	Four patients had extensive diffused DJD (one baseball player, one American football players, one hockey player and one golfer) at time of the arthroscopy but preferred to delay THA. Out of these four patients, two did not return to professional play due to the presence of extensive OA.	Presence of DJD before the HA (OA in this case) has a direct negative effect on returning to sport.
(Philippon, et al., 2010)	Case series - hockey players with FAI	Age / duration of symptoms / CHP	Male Hockey players (n=28).	18 to 36	Age: correlated with the number of games played (r= -0.48, p=0.009). Duration of symptoms: correlated with time to return to play (r= 0.582, p=0.001). Players who had the procedure performed within 1 year of the injury returned to sport after 3 months postoperatively, while those who waited >1 year returned to sport later at 4.1 months. CHP: time from the injury to surgery increased in patients with CHP (24 months vs 14 months, p<0.01).	Age negatively impacts the number of games played leading to reduced performance. Delaying the time of surgery can have a negative impact on returning to sport. Chondral damage can also delay the time of surgery.
(Philippon, et al., 2012b)	Retrospective case series study	DJD	Older patients (n=153).	≥50	DJD: At 3 years postoperatively, patients with >2mm of joint space had a survivorship rate of 90% (out of 64 patients) and those with ≤2mm had a survivorship rate of 57% (p=0.001).	Following hip arthroscopy, patients with reduced joint space have less survivorship rate.
(Philippon, et al., 2012a)	Cohort study - paediatric and adolescents	Age	19 male and 41 female paediatric and adolescent patients.	11 to 16	Age: Moderate but significant correlation between age and the alpha angle (r=0.324, p=0.02).	There is an association between age and the alpha angle.

(Redmond, et al., 2015a)	Retrospective study - minimum of 2-year follow up after HA.	CHP	Patients (sex NA, *n*=30).	63.9±NA	CHP: Chondropathy was more severe in the 9 patients who had THA after an average of 1.1 years (p=0.033). They also had higher acetabular inclination (p=0.023).	Presence of CHP before HA can negatively impact the survival rate after the procedure. THA procedures are eventually performed in 30% of the patients aged >60 years old.
(Redmond, et al., 2015b)	Cohort study – patients who converted to THA.	Sex / age / BMI / back pain / surgical technique	72 patients had THA (out of 792).	NA	Multiple factors: Those who converted had higher mean/frequency of male sex, age, BMI, back pain, previous back surgery, Tönnis grade >0, acetabular inclination angle, labral tear size, labral resection, number of anchors and capsular release.	Multiple factors can influence postoperative prognosis and converting to THA.
(Sansone, et al., 2015)	Case series - top-level athletes	Duration of symptoms	68 male and 17 female athletes (*n*=85).	25±5	Athletes with shorter duration of symptoms (<12 months) returned to sport at a higher rate (87%) compared to athletes with longer duration of symptoms (>12 months) with a lower rate (61%) (p<0.05).	Excellent outcome of hip arthroscopy procedure in top-level athletes.
(Sansone, et al., 2016a)	Case series - patients with mild to moderate OA	DJD	77 males and 23 females (*n*=100).	47±10	DJD: nine patients (11.3%) had THA within 2 years after the arthroscopy procedure due to the presence of OA.	The patients who had THA had reduced joint space even before the arthroscopy was performed. Moreover, THA was initially the best choice of surgery.
(Sansone, et al., 2016b)	Case series - patients with FAI.	Duration of symptoms	190 males and 99 females (*n*=289).	37±13	Duration of symptoms: it is negatively correlated with iHOT-12 (r= -0.189, p=0.012) and HAGOS-QoL (r= -0.209, p=0.012).	The duration of symptoms before the procedure can negatively affect the postoperative improvement.

(Schilders, et al., 2011)	Cohort study – HA patients with labral repair (group 1) or labral resection (group 2).	Surgical technique	76 hips in males and 25 hips in women (total *n*=101).	15 to 71	[Improvement in MHHS from preoperative to follow-up at an average of 2.44 years] Group 1: Preoperative= 60.23 [range 24-85]; Postoperative= 93.59 [range 55-100]. Group 2: Preoperative= 62.78 [range 29-96]; Postoperative= 88.84 [range; 35-100] (*p*=0.036).	Patients who had labral repair or fixation had better postoperative function scores than the patients who had labral resection.

Descr., description; M, mean; SD, standard deviation; NA, not available; HA, hip arthroscopy; DJD, degenerative joint disease; CHP, chondropathy; OA, osteoarthritis; BMI, body mass index; OR, odds ratios; HOOS, Hip Osteoarthritis and Disability Outcome Score; -S, symptoms and stiffness subscale of HOOS; -A, activities of daily living subscale of HOOS; -Sp, sport and recreation subscale of HOOS; -Q, quality of life subscale of HOOS; iHOT-33, International Hip Outcome Tool; THA, total hip arthroplasty.

2.6. *Hand-Held Dynamometry*

Skeletal muscle strength assessment is one of the most frequently used procedures in orthopaedic, physiotherapy and sport injury clinics. Determining which method of muscle strength assessment techniques to be used is based on the healthcare facility setting, available equipment, the assessor's knowledge, and nature of the problem.

Manual muscle testing (MMT) is the simplest form of muscle strength testing. Isokinetic devices (IKD), which are considered a gold standard in muscle strength testing, are also used in the clinics. However, IKD are found mainly in research centres and big orthopaedic clinics due to the high costs, size and complicated assessment protocols. Further information about the validity and reliability of MMT and IKD can be found in Appendix A. 4-8.

Hand-held dynamometry (HHD) was introduced as an inexpensive, quick and easy device to quantify muscle strength with a high measurement ceiling reaching up to 981 Newton (*N*) (Fulcher, et al., 2010; Bohannon, 2012). This new assessment method was adopted by clinicians over time, especially after the 1980s (Bohannon, 2006).

Commercial hand-held dynamometers are being used to assess hip and thigh muscle strength across healthy individuals, sports players and in patients with groin pain (Fulcher, et al., 2010; Kelln, et al., 2008; Thorborg, et al., 2011). On its own, manual muscle testing was able to detect a strength deficit between both sides in 78% of the cases. Nevertheless, when combined with manual muscle testing, the dynamometer could confirm that a strength deficit is present in more than 96% of the cases (Bohannon, 2005).

2.6.1. *Validity of Dynamometers*

The validity* of HHD testing has been compared to a gold standard in muscle strength testing. Isokinetic devices (IKD), such as the Biodex (Biodex Medical Systems, NY, USA) and the HUMAC NORM (Formerly known as the Cybex, Computer Sports Medicine, MA, USA) are considered a gold standard of muscle strength assessment.

In a study looking at the validity and reliability of using the HHD, eight participants were assessed by two examiners using the MicroFET2 HHD (Hoggan Health Industries, UT, USA) and the Biodex System 3 isokinetic device. Testing the dominant side using the MicroFET2 was more reliable in hip extension (ICC=0.93) compared to hip abduction testing (ICC=0.85). The correlation between the MicroFET2 and the Biodex was r =0.419 (p <0.05) for hip extension and r =0.834 (p <0.01) for hip abduction. The inter-tester correlations were r =0.951 (p <0.01) and r =0.789 (p <0.01) for hip extension and abduction respectively (Table 2.11) (Kawaguchi & Babcock, 2010). Although the results show high reliability and moderate to strong correlation between the MicroFET2 and the Biodex, the correlation between the two devices in hip extension testing accounts for only 16% of the variance. The variation in testing these two muscle groups could have resulted from the lack of a unified positioning in the MicroFET2 and Biodex testing. The participants were allowed to hold a chair while standing during Biodex testing, in which postural muscles may have been used to improve stability. Lastly, it was difficult to isolate the gluteus maximus muscle from the hamstring muscle in extension testing in the standing position.

Older adults were also considered in the comparison between the HHD and IKD. The quadriceps muscle strength of twenty participants was assessed using a HHD (model 01163; Lafayette Instrument Company, IN., USA) and the Biodex. The torque values obtained by the Biodex were higher (83.4 ± 28.0 $N{\cdot}m$) than those obtained by the HHD (68.9 ± 19.6 $N{\cdot}m$) with a strong correlation (r =0.91, p <0.0001) between the two devices.

---

* Refers to the extent of how much a device measures what it is supposed to measure (Cuthbert & Goodheart, 2007).

Despite this, the Bland-Altman plot of measurements between the two devices showed that the HHD undervalued quadriceps strength by a mean difference (MD) of 14.5 *N·m*, 95% CI [8.5, 20.6], which was noticed in stronger participants (Table 2.11) (Martin, et al., 2006a). Similar results were reported on 184 participants (mean age: 46.3 ± 22.7) in which a stabilised HHD underscored the quadriceps strength by 33.7 *N*·m on the right and 35.6 *N*·m on the left sides (p <0.001) (Table 2.15) (Bohannon, et al., 2011).

Arnold et al. (2010) measured the lower limb muscle strength of eighteen older adults (mean age: 74 ± 6.75 years). The participants were tested in two different sessions by a physiotherapist with twenty years of experience and a second-year physiotherapy student using a HHD (model 01163, Lafayette Instrument) and the Biodex System 3. The HHD showed high intra-rater reliability for the hip abduction testing in the supine (Right (Rt), ICC=0.85, 95% CI [0.65, 0.94]; Left (Lt), ICC=0.93, 95% CI [0.82, 0.97]) and standing positions (Rt, ICC=0.93, 95% CI [0.82, 0.97]; Lt, ICC=0.94, 95% CI [0.84, 0.98]). However, inter-rater reliability was slightly less in supine (Rt, ICC=0.84, 95% CI [0.33, 0.95]; Lt, ICC=0.85, 95% CI [0.60, 0.94]) and standing positions (Rt, ICC=0.94, 95% CI [0.84, 0.98]; Lt, ICC=0.89, 95% CI [0.72, 0.96]). In terms of the validity, moderate correlations were found between the HHD and the Biodex in Rt hip abduction (Biodex, 36.4 ± 14.5 *N·m*; HHD, 36.2 ± 16.1 *N·m*; r =0.69) and left hip abduction testing in the standing position (Biodex, 40.9 ± 15.0 *N·m*; HHD, 40.9 ± 17.6 *N·m*; r =0.86). Other muscle groups showed fair to high intra- and inter-rater reliability combined with varying degrees of correlations between the two devices in different testing positions (Table 2.8 and Table 2.9). The muscle torque values obtained by the two devices are very close to each other (Table 2.9). Unless the testing was performed with very high accuracy, proper stabilisation and strong examiners, the values obtained by the HHD are questionable when compared to the gold standard. Furthermore, the correlation test does not measure the extent of agreement between the HHD and the Biodex. Therefore, it is not the ideal statistical test for this type of study.

Table 2.8 - HHD test-retest (multiple day intra-rater) and inter-rater reliability (n=18) (Arnold,* et al., *2010).

Measurement		Left		Right	
		ICC (CI)	SEM (kg)	ICC (CI)	SEM (kg)
HF sitting	Intrarater	0.92 (0.80–0.97)	1.11	0.83 (0.61–0.93)	1.68
	Interrater	0.85 (0.64–0.94)	1.46	0.84 (0.62–0.94)	1.68
HF standing	Intrarater	0.91 (0.78–0.97)	0.87	0.84 (0.62–0.94)	1.81
	Interrater	0.88 (0.70–0.95)	1.21	0.82 (0.39–0.94)	1.63
HA supine	Intrarater	0.93 (0.82–0.97)	0.99	0.85 (0.65–0.94)	0.96
	Interrater	0.85 (0.60–0.94)	1.48	0.84 (0.33–0.95)	1.27
HA standing	Intrarater	0.94 (0.84–0.98)	1.03	0.93 (0.82–0.97)	1.07
	Interrater	0.89 (0.72–0.96)	1.31	0.94 (0.84–0.98)	0.96
KE	Intrarater	0.86 (0.65–0.95)	1.79	0.86 (0.54–0.95)	1.67
	Interrater	0.85 (0.65–0.94)	1.85	0.80 (0.21–0.94)	1.83
DF	Intrarater	0.84 (0.62–0.94)	1.03	0.76 (0.47–0.90)	1.36
	Interrater	0.62 (−0.09–0.88)	1.70	0.48 (0.11–0.81)	1.79
HE	Intrarater	0.84 (0.62–0.94)	1.69	0.82 (0.57–0.93)	1.89
	Interrater	0.94 (0.85–0.98)	1.10	0.91 (0.78–0.97)	1.45

HHD = handheld dynamometry; HF = hip flexion; HA = hip abduction; KE = knee extension; DF = dorsiflexion; HE = hip extension; ICC = intraclass correlation coefficient; CI = confidence interval; SEM = standard error of measurement.

Table 2.9 - The relationship and differences between ID measurements (Biodex 3) and HHD measurements adjusted for leg length (n=17) (Arnold,* et al., *2010)

Muscle action	ID Nm (SD)	HHD Nm (SD)	Pearson r	ICC value (CI)
Left HF standing	33.5 (20.0)	32.2 (11.5)	0.59	0.52 (0.06–0.80)
Right HF standing	38.4 (26.2)	33.1 (17.2)	0.77	0.70 (0.34–0.88)
Left KE	73.5 (22.3)*	52.5 (21.3)*	0.75	0.52 (−.10–0.83)
Right KE	68.3 (20.9)*	51.1 (20.0)*	0.58	0.44 (−.05–0.76)
Left HE	47.0 (14.7)	42.8 (17.5)	0.63	0.61 (0.22–0.84)
Right HE	47.7 (13.9)*	40.4 (15.9)*	0.57	0.52 (0.09–0.79)
Left HA standing	40.9 (15.0)	40.9 (17.6)	0.86	0.85 (0.64–0.94)
Right HA standing	36.4 (14.5)	36.2 (16.1)	0.69	0.70 (0.34–0.88)

ID = isometric dynamometry; HHD = handheld dynamometry; HF = hip flexion; HA = hip abduction; KE = knee extension; DF = dorsiflexion; HE = hip extension; ICC = intraclass correlation coefficient; CI = confidence interval; SD = standard deviation.
*Significant difference ($p < 0.05$) between ID and HHD measurements for that muscle action.

Systematic reviews have also been performed to look at the validity of the HHD against the IKD. In the literature review by Bohannon (2012), Pearson correlation between the two devices in the studies based on healthy participants ranged from r =0.52 to r =0.85 for the upper extremity testing and from r =0.74 to r =0.78 for the lower extremity testing. On the other hand, the studies involving patients had correlations of r =0.64 to r =0.90 and r =0.57 to r =0.81 for upper and lower extremity testing respectively (Table 2.11) (Kolber & Cleland, 2005).

In the literature review by Bohannon (2012), the muscle torque values obtained by the HHD were either very close or less than those obtained by the IKD, especially if the testing was performed on healthy individuals or those with strong muscles. Since these

studies were not included in a meta-analysis, the overall reliability and validity figures cannot be obtained. The variation in correlation is due to the dissimilarities in the HHD placement location, the participants' position, the rate of force application and the muscles being tested. Further studies that are based on detailed and well-structured methodologies should be conducted. A summary of the reviewed studies looking into the validity of HHD testing can be found in Table 2.11.

The studies and literature reviews show that the HHD and IKDs have concurrent* validity in testing the upper and lower extremities of healthy individuals and those with known disorders or diseases. However, the HHD can underestimate strength when compared to the IKD. Therefore, the magnitude of difference, the measurement error, and the variation in scores should be taken into consideration.

2.6.2. Reliability of Dynamometers

The reliability† of HHD in testing the hip and thigh muscles strength was examined in various population groups. Bohannon and Andrews (1987) looked at the inter-rater reliability of HHD in testing cerebrovascular accidents, spinal cord injury, and other neurological and orthopaedic patients (n=30). The correlation between the two testers was strong in the upper extremity for the shoulder external rotators (r =0.88, p =0.03), the elbow flexors (r =0.94, p >0.05) and the wrist extensors (r =0.94, p =0.01). The correlation was slightly weaker in the lower extremity for the hip flexors (r =0.84, p >0.05), the knee extensors (r =0.90, p >0.05), and the ankle dorsiflexors (r =0.91, p >0.05). The highest reported difference between the two testers was 1.2 kg (11.8 *N)* (Table 2.11). Although the results show high inter-tester correlation, the intra-tester correlation between scores was not reported. Also, patients with such disorders often

* How well the measurements obtained by an assessment tool agree with another tool at almost the same time (Kolber & Cleland, 2005).

† The extent to which the results of an instrument or technique are consistent when repeated under identical circumstances and can be interpreted in various contexts and situations (Field, 2013).

have muscle strength scores of less than 3 (Fair) on the MMT grading scale, resulting in a reduced variation in scores. Wadsworth et al. (1987) showed test-retest correlation values of r =0.69 for the shoulder abductors, r =0.90 for the elbow extensors, r =0.88 for the wrist extensors, r =0.72 for the hip flexors and r =0.75 for the knee flexors (Table 2.11). Moreover, the HHD was more sensitive in detecting and discriminating small differences in the Fair to the Normal range as graded by MMT.

The relative* and absolute† reliabilities in testing the hip and thigh muscle strength have also been assessed. In a study involving nine healthy and physically active participants, the hip abductor, adductor, flexor, extensor, external and internal rotator muscle groups were tested in the supine, prone and side-lying positions (where applicable). The measurements variation using the standard error of measurement (SEM%‡) ranged between 3% and 12%. Supine hip abduction testing had the lowest SEM% of 2.9% (Intra-class correlation coefficient (ICC)=0.97, 95% CI [0.89, 0.99]), while it was 5.7% for the supine adduction testing (ICC=0.93, 95% CI [0.73, 0.98]). Hip extension testing in the prone position using a short lever arm had the highest SEM% of 12.1% (ICC=0.78, 95% CI [0.30, 0.95]) compared to the long lever arm with a SEM% of 0.86% (ICC=0.86, 95% CI [0.50, 0.97]) (Table 2.11) (Thorborg, et al., 2010a). A similar study on sixteen healthy young adults (mean age 22.7 $\pm$ 2.1 years) reported high reliability values and low measurement error in testing the hip abductors (ICC=0.89, 95% CI [0.83, 0.92], SEM=9.0), hip extensors (ICC=0.92, 95% CI [0.76, 0.97], SEM=7.6) and knee flexors (ICC=0.89, 95% CI [0.68, 0.96], SEM=13.3). The smallest real difference (SRD) was equivalent to 24.8, 29.8 and 36.7 N in the hip abductors, extensors and knee flexors respectively. Testing the knee extensors showed the least reliability values and highest measurement error (ICC=0.60, 95% CI [-0.16, 0.86], SEM=22.8, SRD=63.1). Moreover, the systematic bias (mean of difference) between raters when the hip flexors, abductors

* The consistency of the position in a group relative to others and is usually expressed using the intra-class correlation coefficient (ICC) (Weir, 2005).
† The consistency of scores of individuals and is expressed as the standard error of measurement (SEM) (Weir, 2005).
‡ Calculated by dividing the SEM with the average of the two tests values

and knee flexors were assessed was less than 5% (Table 2.11) (Lu, et al., 2011). The results of Thorborg et al. (2010a) and Lu et al. (2011) indicate excellent relative and absolute reliabilities of testing the hip and thigh muscles. Nevertheless, using the HHD to test the strong muscle groups such as the knee extensor muscles requires high forces to stabilise the HHD in place. The lack of sufficient strength may affect the consistency of the results and increase the measurement error.

Table 2.10 - The relative and absolute reliability of lower extremity strength measured by inter-rater reliability (Lu, et al., 2011)

Measurement	ICC	95% CI for ICC	$\bar{d}$ (| $\bar{d}$ | / mean)	LOA ($\bar{d} \pm 1.96$ *SD*)	*SEM*	*SEM*%	SRD	SRD%
Hip								
Flexors	0.83	0.50–0.94	−9.8 (4.1%)	−58.0 to −38.4	13.7	5.7	37.9	15.8
Abductors	0.89	0.68–0.96	−5.6 (2.0%)	−38.1 to 26.9	9.0	3.3	24.8	9.1
Extensor	0.92	0.76–0.97	12.6 (3.5%)	−15.8 to 41.0	7.6	2.1	29.8	8.4
Knee								
Flexors	0.89	0.68–0.96	2.1 (1.0%)	−46.3 to 50.5	13.3	6.3	36.7	17.4
Extensors	0.60	−0.16 to 0.86	23.6 (8.5%)	−46.4 to 93.6	22.8	8.2	63.1	22.8

$\bar{d}$ = mean difference of the average strength measured by rater A and rater B; $\bar{d}$ = the absolute mean difference ($\bar{d}$) divided by mean strength from rater A and rater B; LOA = limits of agreement = $\bar{d} \pm 1.96$ *SD*; ICC = intraclass correlation coefficient; *SEM* = standard error of measurement = $\sqrt{}$(total variance)(1−ICC); SRD = minimal detectable change = $1.96 \times \sqrt{2} \times$ *SEM*; *SEM*% = *SEM*/mean strength; SRD% = SRD/mean strength.

Two literature reviews were performed to investigate the reliability of HHD. Kolber and Cleland (2005) reviewed over 21 studies that were performed using the HHD on the upper and lower extremities. The upper extremity related studies showed that intra-tester reliability in testing the healthy participants was ICC=0.60 to 0.99 (r=0.83 to 0.96) for intra-session and ICC=0.64 to 0.98 (r=0.88 to 0.99) for inter-session studies. On the other hand, the reliability of HHD in testing the patients was ICC=0.89 to 0.99 (r=0.91 to 0.99) for intra-session and ICC=0.70 to 0.99 (r=0.69 to 0.99) for inter-session studies. However, the lower extremity HHD testing of the healthy participants showed reduced intra-session (ICC=0.16 to 0.98) and inter-session reliability (ICC=0.54 to 0.98, r=0.44 to 0.96). Comparatively, testing the patients showed slightly higher intra-session reliability of ICC=0.93 to 0.99 (r=0.84 to 0.99) and inter-session reliability of ICC=0.75

to 0.97 (r =0.74 to 0.99) (Table 2.11). The search, inclusion and exclusion criteria were not reported by the authors. Therefore, the results cannot be reproduced.

In a more recent literature review, approximately 184 of the 241 reviewed studies (76%) had a test-retest reliability of ICC ≥0.80. The reduced ICC values were reported in lower extremity muscle testing, young participants or the affected side in comparison to the upper extremity and the unaffected side. Furthermore, approximately 114 of 175 studies (65%) had inter-tester reliability values of ICC ≥0.80 (Bohannon, 2012).

In the two literature reviews by Kolber and Cleland (2005) and Bohannon (2012), the broad range of ICC values is due to the nature of the reviewed studies, heterogeneity of the recruited participants and the tested muscles. The stronger muscles of the healthy individuals may explain the reduced reliability in testing compared to the patients. The effects of testing strong muscles on the measurements are discussed later. A summary of the reviewed studies looking into the reliability of HHD testing can be found in Table 2.11.

2.6.3. Factors Affecting Dynamometry Testing

Testing error using the dynamometer could increase if the examiners were not strong enough. A study was performed to assess the influence of examiner strength characteristics on assessing three hip muscle groups. The assessment of muscle torque by the first examiner* was significantly less (p <0.001) than the torque measured by the second examiner† in testing the hip extensor and external rotation muscles. The first examiner also measured the hip abductor muscle torque significantly less (p <0.001) than the third examiner‡. However, there were no statistical differences (p >0.05) in measurements between the second and third examiners (Table 2.12) (Krause, et al.,

* A female; stature: 160 cm, body mass: 55.8 kg; chest press strength: 470 *N*; leg press strength: 1828 *N*.
† A female; stature: 172.7 cm; body mass: 70.3 kg; chest press strength; 675 *N*; leg press strength: 2274 *N*.
‡ A male; stature: 182.9 cm; body mass: 77.1 kg; chest press strength: 1062 *N*; leg press strength: 2787 *N*.

2014). If the generated forces were higher than the resistance applied by the examiner, the HHD would not be stable, leading to a reduction in the recorded force and high human testing error. In this case, a conventional isokinetic device should be used instead of the HHD (Kelln, et al., 2008; Krause, et al., 2014). These results indicate that the examiner's strength and possibly stature and body mass can affect the measurements (Figure 2.12).

The location of placing the dynamometer on the limb and the method used to stabilise it could also affect the measurements. Twenty-one participants had their hip abduction and adduction torque measured by two examiners. The inter-rater reliability of using a long lever* arm was higher in abduction testing (ICC=0.73, 95% CI [0.44, 0.88]) compared to a short-lever† arm (ICC=0.68, 95% CI [0.37, 0.86]). Regarding adduction testing, higher reliability was reported if the opposite side was stabilised on a bench using a long lever arm (ICC=0.82, 95% CI [0.61, 0.92]) compared to the short lever arm without bench stabilisation (ICC=0.74, 95% CI [0.42, 0.89]) (Table 2.12) (Krause, et al., 2007). The results of this study indicate that using a long lever arm is superior to a short lever arm in hip abduction and adduction torque tests.

In the study by Krause et al. (2014), intra-tester reliability using the long lever‡ arm was slightly better compared to the short lever§ arm in hip extensor (ICC_{short}=0.91 to 0.93, ICC_{long}=0.93 to 0.94), hip abductor (ICC_{short}=0.86 to 0.92, ICC_{long}=0.82 to 0.94) and hip external rotator muscles strength testing (ICC_{short}=0.95 to 0.96, ICC_{long}=0.93 to 0.97). Inter-rater reliability showed similar results in testing the strength of the hip extensor (ICC_{short}=0.88, ICC_{long}= 0.90), hip abductor (ICC_{short}=0.81, ICC_{long}=0.98) and hip external rotator muscles (ICC_{short}=0.95, ICC_{long}=0.97). Also, muscle torque values using the short lever arm were significantly less than the long lever arm in hip extensor muscles testing

* Distance between 5 cm proximal to the lateral malleolus and superior tip of the greater trochanter.
† Distance between 7 cm proximal to the lateral knee joint line and the superior tip of the greater trochanter.
‡ Distance between 5 cm proximal to the medial malleolus and the ASIS in the pelvis.
§ Distance between proximal to the knee joint and the ASIS in the pelvis.

(p <0.001, Figure 2.9). However, hip abduction torque values using the short lever arm were significantly greater than the long-lever arm (p <0.001, Figure 2.11). Likewise, hip external rotation values using the short lever arm in a seated position were significantly greater than the long lever arm in a prone position (p <0.001, Figure 2.10) (Table 2.12).

The discrepancy in the lever arm results between the hip extensors compared to the hip abductors and external rotators is because the short lever arm testing was performed in a prone position with the knee flexed to 90°, which may have reduced the activity of the hamstrings. On the other hand, the long lever arm testing was performed with the knee straight (in full extension), allowing the hamstring muscle to contract and aid the gluteus maximus muscle in extending the hip. The other two muscle groups showed reduced torque values in the long lever arm measurement due to the increased difficulty from a biomechanical perspective. A summary of the reviewed studies looking into the effect of lever-arm length in HHD testing can be found in Table 2.12 and Figure 2.12.

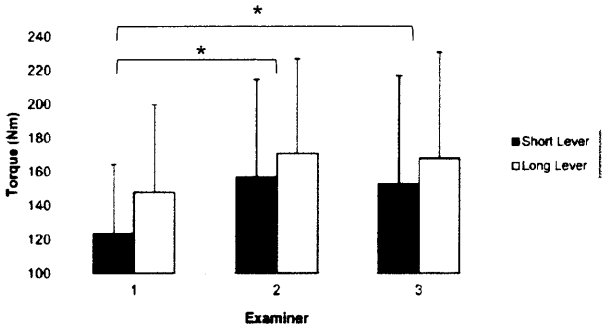


Figure 2.9 - Hip extension torque values (*, significant difference between examiner 1 and examiner 2 (p <0.001) and examiner 3 (p <0.001) (Krause et al., 2014).

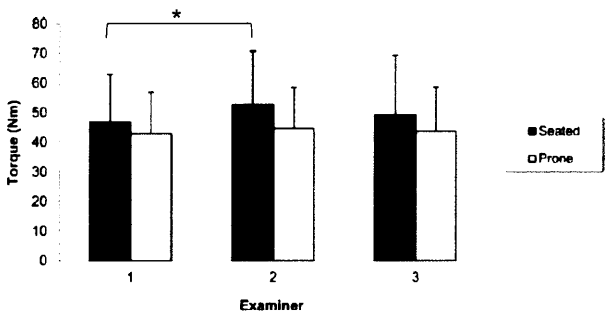


Figure 2.10 - Hip external rotation torque values (*, significant difference between examiner 1 and examiner 3 in short-lever (seated) test only, p =0.001) (Krause et al., 2014).

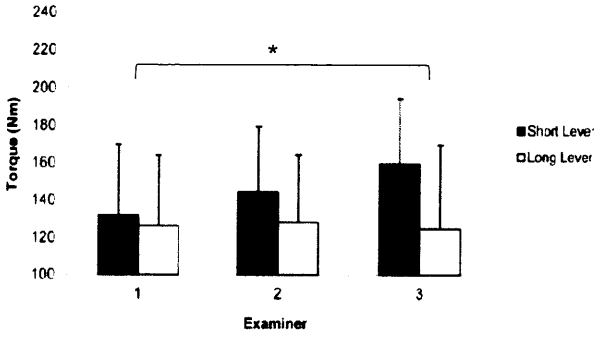


Figure 2.11 - Hip abduction torque values (*, significant difference between examiner 1 and examiner 3, p =0.001) (Krause et al., 2014).

Other factors have also been reported to affect strength measurements obtained by the HHD. The physiological responses and strength values varied among football players on a professional level compared to amateur level (Cometti, et al., 2001; Dellal, et al., 2011). Therefore, the results of HHD testing on a certain population cannot be generalised.

The most important factors that may influence dynamometry measurements were summarised in one study. Increasing age negatively affected muscle strength, with the weaker muscles reported in the older population. Similarly, the female participants showed reduced torque measurements compared to their male peers with the same age and level of athletic activity. On the other hand, the heavy, tall, and active participants scored higher than those who are skinnier, shorter, and less active. The authors pointed out that limb dominance and length of the lever arm during testing can affect the measurements (Keating & Matyas, 1996) (Figure 2.12). A summary of the discussed factors can be found in Figure 2.12.

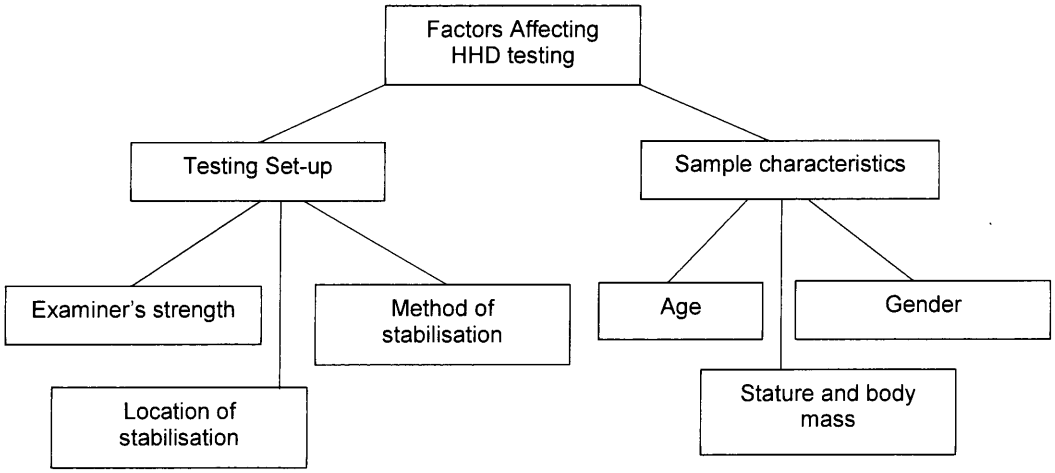


Figure 2.12 – A summary of the factors affecting hand-held dynamometry testing.

Summary

Muscle strength testing using a HHD is a reliable and a valid orthopaedic assessment technique across different age and lower limb muscle groups. However, the reliability values differ, and the testing outcome can be affected by the sample characteristics, the methodology, the location, the method of stabilisation and the strength of the examiner (Figure 2.12). Even though the device is still widely used by clinicians, the reported results of any study will be peculiar to the examined population and age group. Moreover, hand-held dynamometry testing can be affected by many factors. Lack of stabilisation and a unified testing protocol are two of the most important factors. These two factors need to be addressed in future research.

Table 2.11 - Overview of the studies investigating the validity and reliability of mainly lower extremities muscle strength testing using HHDs.

Study (first author, year)	Study Details				Sample Details		Results	Main Findings and Comments
	HHD Model	Muscles tested	Validity assessment	Reliability assessment	Number and sex	Age in (M±SD) or range in years		
(Arnold, et al., 2010)	Model 01163, Lafayette Company.	HFL, HEX, HAB, KEX and ADF.	Yes (Biodex)	Yes (intra- and inter-tester)	4 males and 4 females (*n*=8).	65 to 92	Intra-tester reliability (ICC) Rt side only: HFL SIP=0.83, HFL STP=0.84, HEX=0.82, HAB SUP=0.85, HAB STP=0.93, KEX=0.86 and ADF=0.76. Inter-tester reliability (ICC) Rt side only: HFL SIT=0.84, HFL STP=0.82, HEX=0.91, HAB SUP=0.84, HAB STP=0.94, KEX=0.80 and ADF=0.48. Validity (Pearson *r* [ICC]) Rt side only: HFL STP=0.77[0.70], HEX=0.57[0.52], HAB STP=0.69[0.70] and KEX=0.58[0.44].	High reliability of HHD testing of the lower limb. High correlation between the two devices.
(Bohannon & Andrews, 1987)	Chatillon	SER, EFL, WEX, HFL, and KEX.	No	Yes (inter-tester)	12 males and 18 females patients with chronic orthopaedic, neurological and medical conditions (*n*=30).	NA	Inter-tester reliability (Pearson *r*): SEX=0.88 (p=0.03), EFL=0.94 (p>0.05), WEX=0.94 (p=0.01), HFL=0.84 (p>0.05), and KEX=0.91 (p>0.05).	HHD of uppers and lower extremity muscles is reliable. The reduced reliability of lower extremity muscle testing is probably due to the high amount of force they produce.
(Denton, et al., 2014)	Nicholas	HAB, HAD, HFL, HEX and HIR.	No	Yes (inter-session)	5 males and 5 females (*n*=10).	25.4±6.0	Inter-session reliability (ICC): HAB=0.97, HAD=0.94, HFL=0.86, HEX=0.97, HIR=0.94 and HER=0.94.	All testing procedures were performed against gravity. Testing hip muscle strength with contra-lateral side stabilisation has excellent reliability.

(Lu, et al., 2011)	MicroFET2	HFL, HAB, HEX, KFL, KEX.	No	Yes (inter-tester)	8 males and 8 females (*n*=16).	22.7±2.1	Inter-tester reliability (ICC, 95% CI [lower, higher]): HAB=0.89, 95% CI [0.83, 0.92], SEM=9.0. HEX=0.92, 95% CI [0.76, 0.97], SEM=7.6, KFL=0.89, 95% CI [0.68, 0.96], SEM=13.3) and KEX=ICC=0.60, 95% CI [-0.16 to 0.86], SEM=22.8.	Testing hip and thigh muscle strength using a HHD has a high reliability except the knee extensors.
(Kawaguchi & Babcock, 2010)	MicroFET2	HEX and HAB.	Yes (Biodex)	Yes (intra- and inter-tester)	College students (sex NA)	18 to 22	Intra-tester reliability (ICC): Tester1 HEX Rt=0.96, HEX Lt=0.84, HAB Rt=0.95 and HAB Lt=0.94. Tester 2 HEX Rt=0.92, HEX Lt=0.96, HAB Rt=0.96 and HAB Lt=0.98. Inter-tester reliability (ICC): HEX Rt=0.93, HEX Lt=0.85, HAB Rt=0.85 and HAB Lt=0.92. Validity (Pearson *r*): HEX=0.419 and HAB=0.834.	HEX and HAB testing using the HHD is reliable and valid. Correlation between the MicroFET2 and Biodex is less in HEX testing due to positioning.
(Kelln, et al., 2008)	MicroFET2	HAB, HAD, HFL, HEX, HER, HIR, KFL, KEX, AIN, AEV, ADF.	No	Yes (intra-tester and inter-session)	9 males and 11 females (*n*=20).	28.1±7.4	Intra-tester reliability range (ICC): HAB=0.92-0.97, HAD=0.77-0.94, HFL=0.91-0.96, HEX=0.77-0.93, HER=0.90-0.96, HIR=0.84-0.96, KFL=0.91-0.94, KEX=0.82-0.93, ADF=0.87-0.94, AIN=0.82-0.94, AEV=0.91. Inter-session reliability range (ICC): HAB=0.82-0.91, HAD=0.62-0.80, HFL=0.66-0.79, HEX=0.69-0.79, HER=0.82-0.89, HIR=0.77-0.82, KFL=0.78-0.92, KEX=0.70-0.92, ADF=0.81-0.92, AIN=0.77-0.85, AEV=0.86-0.92.	Inter-session reliability is lower than intra-tester. Muscles with higher strength or mechanical advantage over the tester had lower reliability values. HHD testing of lower limb muscles is a reliable test. Mechanical advantage of the tester should be applied to improve the reliability.

(Kolber & Cleland, 2005)	NA (Literature review)	NA	NA	Yes (intra-tester for intra-session and inter-session)	NA	NA	Reliability (ICC): Upper extremity testing intra-tester reliability in testing healthy participants was 0.60 to 0.99 (r=0.83 to 0.96) for intra-session and 0.64 to 0.98 (r=0.88 to 0.99) for inter-session studies. Testing patients had intra-tester reliability of 0.89 to 0.99 (r=0.91 to 0.99) for intra-session and 0.70 to 0.99 (r=0.69 to 0.99) for inter-session studies. Lower extremity testing: testing healthy participants had intra-session reliability of 0.16 to 0.98 and inter-session reliability of 0.54 to 0.98 (r=0.44 to 0.96). Testing patients had higher intra-session reliability of 0.93 to 0.99 (r=0.84 to 0.99) and inter-session reliability of 0.75 to 0.97 (r=0.74 to 0.99).	Higher testing reliability of patients due to reduced variability in force production.
(Martin, et al., 2006a)	Model 01163, Lafayette Company.	KEX	Yes (Biodex)	No	9 males and 11 females (n=20).	61 to 81	Validity (Pearson r): KEX with isokinetic testing (r=0.91, p<0.001). MD=14.5 N·m, 95% CI [8.5, 20.6].	High correlation between the two devices. The HHD underscored KEX strength.
(Muff, et al., 2015)	MicroFET2	KFL and KEX.	Yes (Con-Trex MJ at 0°.sec^{-2}, 60°.sec^{-2} and 180°.sec^{-2}.	Yes (intra-tester)	20 males and 10 females (n=30).	32.8±11.5	Validity (Pearson r): KEX with isokinetic testing (r=0.84-0.87) and with isometric testing (r=0.85-0.87). KFL with isokinetic testing (r=0.75-0.85) and with isometric testing (r=0.72-0.75). Reliability (MD, 95% LoA, and CV): KEX DS=0.1, [-7.3, 7.6], 3.2%. KEX NS=0.4, [-7.5, 7.5], 3.3%. KFL DS=0.4, [-4.6, 5.4], 4.2%. KFL NS=0.3, [-3.9, 4.6], 3.7%.	KEX correlated more with isometric testing. KFL correlated more with isokinetic testing. Minimal differences in retesting KEX and KFL strength using a HHD.

(Thorborg, et al., 2010a)	NA	HAB, HAD, HFL, HEX, and HIR.	No	Yes (intra-tester within a week)	5 males and 4 females (*n*=9).	Males: 27±5; females: 25±4	(Best of 4 reps; ICC). HAB-SLP=0.74, HAB-SUP=0.97, HAD-SLP=0.75, HAD-SUP=0.93, HEX-PP-LL=0.86, HEX-PP-SL=0.78, HER-PP=0.99, HER-SIP=0.92, HEX-SIP=0.91, HFL-SUP=0.94, HIR-PP=0.89, HIR-SIP=0.80.	Testing hip muscle strength using a HHD has a high reliability. Strength changes above 10% are considered real changes.

M, mean; SD, standard deviation; NA, not available; HHD, hand-held dynamometer; SER, shoulder external rotation; EFL, elbow flexion; WEX, wrist extension; HAB, hip abductor; HAD, hip adductor; HER, hip external rotator; HIR, hip internal rotation; HEX, hip extensor; HFL, hip flexion; KEX, knee extensor; KFL, knee flexor; APF, ankle planter flexor; ADF, ankle dorsi flexor; AEV, ankle eversion; AIN, ankle inversion; STP, in a standing position; SLP, in a side-lying position; SUP, in a supine position; PP-LL, in a prone position long lever arm; PP-SL, in a prone position short lever arm; PP, in a prone position; SIP, in sitting position; Rt, right; Lt, left; ICC, intra-class correlation; CI, confidence interval; SEM, standard error of measurement; N.m, newton metre; MD, mean of difference; LoA, limits of agreement; CV, coefficient of variation.

Table 2.12 - Overview of the studies investigating the differences of lever arm length in HHD testing.

Study (first author, year)	Study Details			Sample Details		Results	Main Findings and Comments
	Muscles tested	Validity assessment	Reliability assessment	Number and sex	Age in (M±SD) or range in years		
(Krause, et al., 2007)	HAB and HAD.	No	Yes (intra-tester and inter-tester)	12 males and 9 females (*n*=21).	Males: 24±3; females: 25±3.	Intra-tester reliability for short and long lever arms consecutively (ICC): HAB=0.91 and 0.93, HAD=0.89 and 0.79, HAD with bench=0.83 and 0.89. Inter-tester reliability for short and long lever arms consecutively (ICC): HAB=0.68 and 0.73, HAD=0.74 and 0.64, HAD with bench=0.62 and 0.82.	Using a long lever arm is more reliable in hip abduction strength testing. Using a longer lever arm with bench stabilisation of the opposite side is more reliable in hip adduction strength testing.
(Krause, et al., 2014)	HEX, HAB and HER.	No	Yes (intra-tester and inter-tester)	11 males and 19 females (*n*=30).	24±1.4	[Reliability of tester 1, 2 and 3 consecutively] Intra-tester reliability for short lever arm (ICC): HEX=0.92, 0.93, and 0.91; HAB=0.91, 0.86, and 0.92; HER=0.96, 0.95 and 0.96. Intra-tester reliability for long lever arm (ICC): HEX=0.94, 0.93 and 0.94; HAB=0.82, 0.94 and 0.92; HER=0.96, 0.93 and 0.97. Inter-tester reliability for short lever arm (ICC): HEX=0.88; HAB=0.81 and HER=0.95. Inter-tester reliability for long lever arm (ICC): HEX=0.90, HAB=0.98 and HER=0.97.	Testing the strength of these muscles using a long lever arm is more reliable than using a short lever arm.
(Thorborg, et al., 2010a)	HEX	No	Yes (intra-tester within a week)	5 males and 4 females (*n*=9).	Males: 27±5; females: 25±4.	(Best of 4 reps [ICC]) HEX prone position long lever=0.86 and HEX prone position short lever=0.78.	Using a long lever arm is more reliable in hip extension strength testing.

Stab., stabilisation; M, mean; SD, standard deviation; HAB, hip abductor; HAD, hip adductor; HER, hip external rotator; HEX, hip extension; ICC, intra-class correlation.

2.7. *Stabilisation and Strength Assessment Platforms (SAP)*

Hand-held dynamometry testing can be affected by several issues such as the strength of the examiner and the location of placement on the limb. Therefore, some authors proposed solutions to overcome these problems. When the examiner is weaker than the patient, placing the dynamometer on a distal site will provide the examiner with a mechanical advantage. Thus, the strength difference is reduced (Krause, et al., 2007). Sufficient stabilisation is required to improve the validity and reliability of HHD testing. It can be achieved with proper positioning of the therapist and strapping the patient's pelvis and lower extremities with belts (Kolber & Cleland, 2005).

A study that included twenty elderly and twenty young female patients revealed that the hip abduction torque values were higher in belt-resisted testing compared to the examiner-resisted testing (Table 2.13). The belt-resisted testing was also more reliable in testing the young (Rt, ICC =0.92, 95% CI [0.81, 0.97]; Lt, ICC =0.92, 95% CI [0.80, 0.97]) and elderly patients (Rt, ICC =0.98, 95% CI [0.96, 0.99]; Lt, ICC =0.98, 95% CI [0.95, 0.99]). On the other hand, the examiner-resisted testing was less reliable in testing the young (Rt, ICC =0.84, 95% CI [0.64, 0.94]; Lt, ICC =0.93, 95% CI [0.81, 0.97]) and elderly patients (Rt, ICC =0.96, 95% CI [0.91, 0.99]; Lt, ICC =0.97, 95% CI [0.92, 0.99]) (Table 2.15) (Kramer et al., 1991). The study results indicate that stabilising the HHD with a belt is superior to simply holding it by hand while resisting the muscular contractions. Similar results were reported in another study showing high intra-tester (ICC= 0.880-0.958) and inter-tester (ICC= 0.899-0.948) reliability in using stabilised HHDs to measure hip abduction strength (Table 2.15) (Fenter, et al., 2003b).

Table 2.13 - Hip abductor torque (N·m) for the elderly and young patients in the first session (Kramer, et al., 1991).

		Right Hip		Left Hip	
Group	Test Method	Occasion 1	Occasion 2	Occasion 1	Occasion 2
Elderly (n = 20)	Belt	49 (22)	50 (23)	49 (21)	49 (21)
	Examiner	29 (13)	30 (13)	31 (12)	30 (13)
Young (n = 20)	Belt	77 (17)	77 (18)	78 (16)	78 (18)
	Examiner	54 (17)	50 (15)	52 (17)	50 (16)

A more recent study showed excellent reliability (ICC range 0.97 to 0.99) in belt-stabilised testing compared to the manual stabilisation method in hip muscle strength testing (ICC range 0.41 to 0.83) (Table 2.14) (Katoh & Yamasaki, 2009). Lower ICC values were reported in hip abduction and extension, knee extension and ankle planter flexion due to the strong nature of these muscle groups. It is most likely that the participants were stronger than the tester. Therefore, the HHD could not have been stabilised properly.

Another study investigated the validity of belt-resisted testing of the hip and knee muscles against the Cybex NORM, which is an IKD. The strength values obtained by the Cybex were significantly higher than those achieved with HHD testing with varying correlation levels between the two methods (exact *p* values were not provided). The hip external rotation and extension testing using the HHD had the strongest correlation (r =0.86 and r =0.84 respectively) with the Cybex, while hip abduction showed low to moderate correlation (r =0.34). However, when the highest three values were excluded, the correlation with the Cybex in abduction testing increased (r =0.65). Knee flexion and knee extension showed higher correlation values (r =0.88 and r =0.75 respectively) (Table 2.15) (Katoh, et al., 2011). The higher correlation values reported in knee muscle testing compared to the hip are probably due to the ease of testing and positioning. The participants were positioned in a seated position with maximal stabilisation of the thorax and thighs during knee muscle testing compared to the hip abduction and adduction testing, which were performed in a supine position. It may have been harder to stabilise the participants during hip testing.

Table 2.14 – Intra-class correlation coefficient of isometric muscle strength measurements using a HHD with and without a belt (Katoh & Yamasaki, 2009).

Exertion Task	Without Belt			With Belt		
	ICC (2.1)	95%CI		ICC (2.1)	95%CI	
		Lower	Upper		Lower	Upper
Hip flexion	0.66	0.11	0.86	0.98	0.96	0.99
Hip extension	0.61	–0.06	0.85	0.98	0.96	0.99
Hip abduction	0.41	–0.10	0.41	0.99	0.97	0.99
Hip adduction	0.83	0.58	0.93	0.98	0.97	0.99
Hip external rotation	0.82	0.67	0.90	0.97	0.95	0.98
Hip internal rotation	0.74	0.45	0.87	0.99	0.97	0.99
Knee flexion	0.79	0.40	0.91	0.98	0.96	0.99
Knee extension	0.04	–0.04	0.16	0.98	0.97	0.99
Ankle plantar flexion	0.33	–0.10	0.68	0.99	0.98	0.99
Ankle dorsiflexion	0.74	0.38	0.88	0.98	0.96	0.99

Both methods of stabilisation were compared in the study by Kim et al. (2014). The correlation between the fixed HHD and the Biodex together with the inter-tester reliability were the strongest in the supine position (Rt, r =0.806, ICC= 0.984; Lt, r =0.524, ICC= 0.952) compared to the seated position (Rt, r =0.473, ICC= 0.984; Lt, r =0.294, ICC= 0.983). While the lowest correlations were reported between the non-fixed HHD and the Biodex in the supine (Rt, r =0.185, ICC= 0.963; Lt, r =0.384, ICC= 0.95) and the seated positions (Rt, r =0.001, ICC= 0.94; Lt, r =0.016, ICC= 0.962). The reported torque values were the closest between the fixed HHD and the Biodex (Table 2.16) (Kim, et al., 2014). The findings suggest that the fixed method strongly correlated with the Biodex with better consistency in testing. Even though, the high ICC values reported with manual stabilisation do not necessarily mean that it should not be used. It is recommended that a test of agreement between scores of the HHD and the Biodex should have been performed instead of Pearson correlation.

The fixed stabilisation method improves the consistency of testing showing a stronger correlation with the isokinetic devices. Various authors introduced platforms and stabilising frames acting as anchoring stations to test the hip and thigh muscles' strength (Nadler, et al., 2000; Scott, et al., 2004; Tourville, et al., 2013; Traina, et al., 2010; Widler, et al., 2009).

The stabilising platform by Nadler et al. *(2000)*

The first anchoring station was developed more than fifteen years ago. The station, or what will be referred to as a platform in this research, was designed to test hip extensor and abductor muscle strength (Figure 2.13). To examine the reliability of this apparatus, ten participants had their muscle strength tested in two sessions.

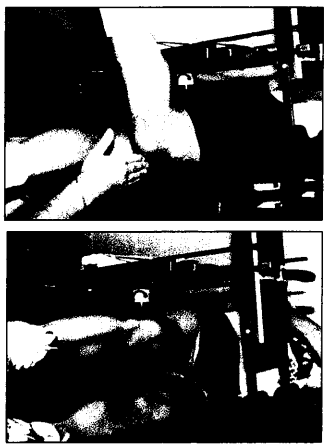


Figure 2.13 - Strength measurement of the right extensors (Top) and the hip abductors (bottom) (Nadler et al., 2000)

The results indicated good test-retest reliability for the abductor (ICC =0.95) and extensor (ICC =0.94) muscle groups. The average coefficient of variation (CV) was 4.77% (range: 0.7%-13.98%) for the hip abductor and 8.06% (range: 1.84%-27.7%) for the hip extensor muscles (Table 2.16) (Nadler, et al., 2000). The results support the use of this anchoring station in testing hip muscle strength with less consistency and increased variation in extensor muscles testing. However, the platform is large and may not fit all treatment beds.

The stabilising platform by Widler et al. *(2009)*

A few years later, another platform was developed to assess hip abductor muscle strength in side-lying, supine and standing positions. It was securely anchored to a physiotherapy treatment table using screw clamps. Muscle strength was measured using a HHD (Model 01163, Lafayette Instrument, IN, USA) which was attached to the frame (Figure 2.14) (Widler, et al., 2009).

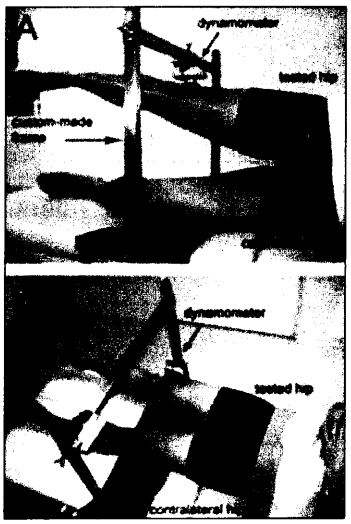


Figure 2.14 – A participant positioned for abductor testing (Widler, et al., 2009).

In the study, the dynamometer was positioned 5 cm proximal to the lateral femoral condyle (just above the knee joint) with the hip positioned in 10° of abduction, for

consistent results (Figure 2.14). The highest recorded value was taken into consideration after being normalised to body weight. Moreover, electromyography (EMG) of the gluteus medius muscle was recorded using a single-differential bipolar surface electrode (model: DE-2.1, Delsys, MA, USA).

The testing of sixteen participants revealed that the highest maximum voluntary contraction was recorded in the side lying position, which was significantly higher by 30% than the supine position (p <0.001) and 10% higher than the standing position (p <0.05) (Figure 2.15). The EMG activity ratio of the contralateral-to-tested side was the highest in the standing and supine positions with significant difference between the three positions (p <0.001, Figure 2.15). The side-lying position showed the least variation in results (CV =3.67%, limit of agreement (LoA) = ±6.87%) compared to the supine (CV =6.11%, LoA = ±8.37%) and the standing positions (CV =4.22%, LoA = ±7.48%). In terms of reliability, the side-lying position (ICC =0.90, 95% CI [0.74, 0.97]) was also more reliable than supine (ICC =0.83, 95% CI [0.57, 0.94]) and standing positions (ICC =0.88, 95% CI [0.69, 0.96]) (Table 2.16) (Widler, et al., 2009).

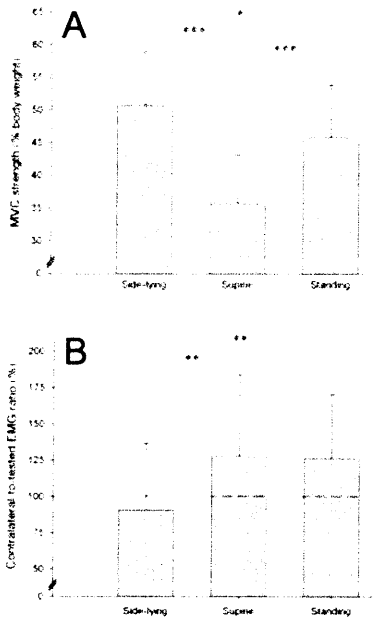


Figure 2.15 - Hip abductor strength and gluteus medius EMG (mean and standard deviation (A) Maximal voluntary contraction (MVC) strength (normalized to body weight; *, p < 0.05; ***p <0.001) (B) Contralateral-to-tested EMG ratio during maximal isometric hip testing. (**p < 0.01) (Widler, et al., 2009).

The results indicate that a side-lying position is preferable over the two other positions in testing the hip abductors. However, the researchers did not compare the obtained muscle strength values using this platform to an isokinetic device. Also, EMG analysis may not have been the best choice to assess the validity of the platform. It has been reported that a disconnection between EMG muscle activity and force production exists in hip muscle testing (Widler, et al., 2009).

The stabilising platform by Traina et al. *(2010)*

A group of researchers developed a reproducible and inexpensive method to measure hip abductor strength. The platform consisted of a built-in force measurement tool to assess hip abductor muscles in the supine position. The measurement tool consisted of a load cell that was attached to a rubber-padded plate positioned 2 cm proximal to the patella. Two adjustable pads were fixed to the table on the opposite side to counterbalance the abduction force (Figure 2.16).

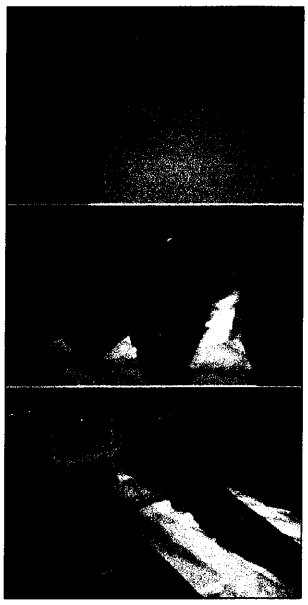


Figure 2.16 - The stabilising platform (frame) by Traina et al. (2010).

To test the reliability of this platform, ten healthy participants (age range: 25-35 years) had their hip abduction muscle strength tested over two separate sessions by two examiners. Testing in the second session was slightly more consistent (ICC_{right} =0.94, ICC_{left} =0.88) compared to the first session (ICC_{right} =0.84, ICC_{left} =0.81). There were no significant differences in scoring between the first (CV =4.93%; range: 1.95%-8.10%) and second examiners (CV =3.88%; range: 1.50%-6.55%) with low variability in the combined scores of both examiners (CV =4.23%; range: 1.46%-7.22%; p >0.05) (Table 2.16) (Traina, et al., 2010). The platform has a high reliability due to the contralateral stabilisation provided during abduction testing and the proximity between the two testing sessions (two days).

The stabilising platform by Tourville et al. *(2013)*

Tourville et al. (2013) designed and proposed a new stabilisation platform to test multiple hip muscle groups. The modifiable arm of the load cell allowed testing the hip joint in 0° to 45° in both flexion and extension (Figure 2.17).

The lever arm was measured from the point of application to the anterior superior iliac spine (ASIS) and was used for torque calculation. Nineteen healthy professional ice-hockey players volunteered to have their hip muscle strength tested using the frame. In hip abduction and adduction testing, the participants were positioned in a supine position with both hips in 15° of abduction from the centre line. The hip flexion and extension testing were performed in a supine position with the hips positioned at 10° of abduction and 10° of flexion. Testing was done at three-time points (baseline, one week, and three weeks) by two trained examiners.

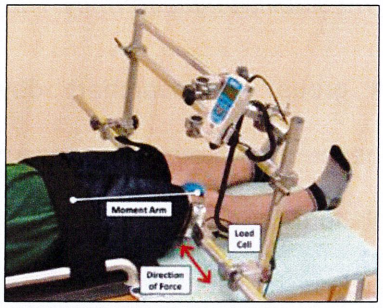


Figure 2.17 – A participant positioned for hip abduction/adduction strength testing (Tourville, et al., 2013).

The highest variability was between-subjects (range 74%-91%) followed by between-sessions (range 8%-21%). Intra-tester reliability of torque measurements was the highest for adduction (ICC =0.92) followed by flexion (ICC =0.91), abduction (ICC =0.91) and extension (ICC =0.78). These findings indicate that the frame has moderate to high reliability in testing the hip muscles. However, the increased variability between the sessions should be considered, which might have been caused by an increase in strength or variations in the testing position. Also, fixing the platform to a treatment bed and aligning the joints require additional time that is added to the overall clinical testing time.

The stabilising device by Jackson et al. *(2016)*

A simpler stabilisation device to test hip muscle strength was developed by another team of researchers. The device consists of a Polyvinyl chloride (PVC) pipe with a wide contact base and another end to accommodate the MicroFET2 HHD (Hoggan Health Industries, UT, USA).

Fifteen healthy active runners were recruited. They were positioned in a supine position for the hip abduction and adduction followed by a seated position for the hip external rotation, knee extension and ankle planter flexion strength testing (Figure 2.18). All tests were repeated three times (6 seconds each) in two sessions with a 15 minutes rest period in between.



Figure 2.18 - Hip abduction testing using the PVC pipe stabilisation device (Jackson, et al., 2016).

Intra-session reliability was the highest for the hip external rotators (ICC =0.98, 95% CI [0.94, 0.99]) and ankle planter flexors (ICC =0.98, 95% CI [0.93, 0.99]) followed by the hip adductors (ICC =0.97, 95% CI [0.90, 0.99]) and hip abductors (ICC =0.96, 95% CI [0.89, 0.99]). The knee extensors had the lowest reliability values using the stabilisation device (ICC =0.93, 95% CI [0.82, 0.98]). In the second session, the hip adductors, external rotators, and ankle planter flexors showed a reduction in the reported force in comparison to the other two muscle groups (Table 2.16) (Jackson, et al., 2016).

Overall, using the PVC pipe as a stabilisation device showed high reliability in testing these lower limb muscles. It is worth mentioning that all muscle groups were tested with a strap stabilising the pelvis or the thighs, which may have resulted in the increased repeatability of the results. A timeline representing the platforms from oldest to the most recent can be found in Figure 2.19. A summary of the reviewed platforms in addition to other can be found in Table 2.16.

Summary

Testing the hip and thigh muscle strength using a hand-held dynamometer could be affected by many factors. Out of those factors, the location of positioning and the stabilisation method are of importance. The findings of several research projects

revealed that placing the dynamometer at a distal location and applying a long lever arm is superior to the proximal location with a short lever arm.

Stabilising the dynamometer using a belt or a platform provides results that are more consistent. There is still a need for a platform that is smaller, easy to carry and assemble. In Chapter 4, the designing process of two platforms, which were designed by the researcher, will be discussed and critiqued.



Figure 2.19 - Sequence of the stabilising platforms listed according to the authors and the muscle groups of interest (HFL, hip flexion; HEX, hip extension; HAB, hip abduction; HAD, hip adduction; HER, hip external rotation; KEX, knee extension; APF, ankle planter flexion).

2.8. *Literature Review Key Points*

- Intra-articular hip pain can be caused by multiple conditions. Hip arthroscopy is an effective treatment to relieve non-arthritic pain.
- Hip arthroscopy results in muscle strength, range of motion, function and pain improvement. However, the outcome varies based on the treated condition and sample of patients.
- Several factors such as age, sex, BMI and the presence of chondropathy can affect the outcome. However, the effect of such factors was dependent on the studied sample of patients. An inclusive and heterogeneous sample is required to study these factors further.

- The follow-up with hip arthroscopy patients was limited to few weeks after the procedure. The short to medium-term postoperative changes should be investigated. Thus, a study that includes postoperative improvements in muscle strength, range of motion, function and pain in a more comprehensive approach is required.
- The hand-held dynamometer is considered a valid and reliable tool in assessing muscle strength. Still, the measurements can be affected by the method of stabilisation and location of placement. Such factors need to be further investigated.
- The use of strength assessment platforms in muscle strength testing can overcome the strength differences between the patient and the examiner together with the issues related to the stability of HHD during testing. Nonetheless, these platforms are limited in their use because of their size and configuration. Moreover, the validity of some of these platforms was not assessed. Furthermore, there were no follow-up studies to test the validity and reliability of the platforms on patients, nor there were any improvements made to the design.
- There is still a need for a versatile, quick and easy tool to improve the objectivity of HHD testing.

Table 2.15 - Overview of the studies investigating the types of stabilisation in HHD testing of the lower extremities muscles.

Study (first author, year)	Study Details				Sample Details		Results	Main Findings and Comments
	Type of stab.	Muscles tested	Validity assessment	Reliability assessment	Number and sex	Age in (M±SD) or range in years		
(Bohannon, et al., 2011)	Belt-fixed HHD.	KEX	Yes (Biodex)	No	71 males and 113 females (n=184).	46.3±22.7	Correlation (Pearson r): Rt=0.884 and Lt=0.864 (both p<0.001). MD Rt=33.7 N.m and MD Lt=35.6 N.m.	KEX torque measurement using a belt-stabilised HHD have high correlation with the Biodex. The HHD underscored the participants.
(Fenter, et al., 2003a)	3 HHDs stabilised on a portable treatment bed (HHD1=MicroFET, HHD2= Dial Push Pull, HHD3=Jammar).	HAB	No	Yes (intra- and inter-tester)	Females (n=10).	23.5±1.91	Inter-tester reliability (ICC): HHD1=0.948, HHD2=0.899, HHD3=0.914. Intra-tester reliability (ICC) (Rater A then B): HHD1=0.880 and 0.942, HHD2=0.898 and 0.958, HHD3=0.895 and 0.932.	The reliability of three stabilised HHD models is high. The measurement values should have been compared with a non-fixed method.
(Katoh, et al., 2011)	Belt-fixed HHD.	HEX, HFL, HAB, HAD, HIR, HER, KEX and KFL.	Yes (Cybex at varying positions)	No	12 males and 12 females (n=24).	20.4±2.2	Correlation (Pearson r) HFL=0.52, HEX=0.84, HAB=0.65 (when n=3 strongest subjects were excluded from analysis), HER=0.86, HIR=0.77, KFL=0.88, KEX=0.75 (all p<0.05).	HHD-belt underscored muscle strength. Moderate to high correlation between belt stabilised HHD and the Cybex. Limit of agreement studies should have been used.
(Kramer, et al., 1991)	Manual and belt-fixed HHDs.	HAB	No	Yes (intra-tester)	Elderly (n=20) and young females (n=20).	Elderly: 68.4±5.2; young: 23.5±3.4.	Intra-tester reliability (ICC): Belt-resisted testing in young (Rt=0.92, Lt=0.92) and elderly patients (Rt=0.98, Lt=0.98). Manual-resisted testing in young (Rt=0.84, Lt=0.93) and elderly patients (Rt=0.96, Lt=0.97).	Belt-resisted testing was more reliable than manual-resisted. No validity tests were performed.

(Toonstra & Mattacola, 2013)	Belt-fixed HHD and an isokinetic device.	KEX and KFL	Yes (Cybex)	Yes (inter-session within a day)	Adults (n=16).	29.3±7.2	Inter-session reliability (ICC): HHD: KEX=0.76 and KFL=0.49. Belt-fixed HHD: KEX=0.92 and KFL=0.96. Isokinetic: KEX=0.93 and KFL=0.89.	High reliability for testing KEX and KFL using a belt-stabilised HHD and an isokinetic device compared to a manually stabilised HHD.
(Whiteley, et al., 2012)	Belt-fixed HHD and an isokinetic device.	KEX and KFL	Yes (Biodex)	Yes (inter-tester in one session)	Football male players (n=10).	-	Validity (Pearson r): Correlation between isometric HHD KEX and concentric testing at 60°.sec^{-2} (r=0.540-0.617, p<0.001) and at 300°.sec^{-2} (r=0.447-0.472, p<0.001). Correlation between isometric HHD KFL and concentric testing at 60°.sec^{-2} (r=0.525-0.548, p<0.001) and at 300°.sec^{-2} (r=0.334-0.379, p<0.001). Correlation between eccentric HHD KFL and concentric testing at 60°.sec^{-2} (r=0.519-0.559, p<0.001) and at 300°.sec^{-2} (r=0.322-0.368, p<0.001). Inter-rater reliability (ICC): Isometric KEX=0.96 and KFL=0.91. Eccentric KFL=0.90.	Excellent reliability of using a belt-stabilised HHD in KEX and KFL strength testing. Isometric testing with HHD have higher correlation with concentric testing at 60°.sec^{-2}. Limit of agreement studies should have been performed instead of Pearson r.

Stab., stabilisation; M, mean; SD, standard deviation; HHD, hand-held dynamometer; HAB, hip abductor; HAD, hip adductor; HER, hip external rotator; HIR, hip internal rotation; HEX, hip extensor; HFL, hip flexion; KEX, knee extensor; KFL, knee flexor; Rt, right; Lt, left; ICC, intra-class correlation.

Table 2.16 - Overview of the studies that were performed using stabilising platforms or frames in lower extremities muscle strength testing.

Study (first author, year)	Study Details				Sample Details		Results	Main Findings and Comments
	Platform/frame info.	Muscles tested	Validity assessment	Reliability assessment	Number and sex	Age in (M±SD) or range in years		
(Beshay, et al., 2011)	Two chairs with stabilising rods (ORI-SST).	SIR, SER, SAB, and SAD.	No	Yes (intra- and inter-tester)	Healthy participants (sex NA, *n*=6).	23 to 42	ORI-SST: Intra-tester (ICC) for SIR=0.98, SER=0.94, SAB=0.98 and SAD=0.97. Inter-tester (ICC) for SIR=0.98, SER=0.93, SAB=0.96 and SAD=0.94. HHD: Intra-tester (ICC) for SIR=0.98, SER=0.95, SAB=0.93 and SAD=0.98. Inter-tester (ICC) for SIR=0.89, SER=0.89, SAB=0.83 and SAD=0.92.	No validity test was performed. Higher reliability in using the ORI-SST as a stabilisation method to test the shoulder muscles.
(Jackson, et al., 2016)	PVC pipe with a HHD.	HAB, HAD, HER, KEX and APF.	No	Yes (inter-session); two sessions	8 males and 7 females (*n*=15).	34.3±3.5	Reliability (ICC, 95% CI [lower, higher]): HAB= 0.96 [0.89, 0.99), HAD=0.97 [0.90, 0.99], HER=0.98 [0.94, 0.99], KEX=0.93 [0.82, 0.98], APF=0.98 [0.93, 0.99].	No validity test was performed. High reliability in using the PVC pipe to stabilise the HHD.
(Kim, et al., 2014)	Wooden board with a HHD.	KEX	Yes (Biodex, isometric at 35°)	Yes (intra- and inter-tester)	Females (*n*=7).	23.9±2.9	Fixed HHD: inter-tester reliability in a supine position (Rt, *r*=0.806, ICC= 0.984; Lt, *r*=0.524, ICC=0.952) and a seated position (Rt, *r*=0.473, ICC=0.984; Lt, *r*=0.294, ICC=0.983), all $p<0.05$. Non-fixed HHD: in a supine position (Rt, *r*=0.185, ICC=0.963; Lt, *r*=0.384, ICC=0.95, $p<0.05$ Lt only) and in a seated positions (Rt, *r*=0.001, ICC=0.94; Lt, *r*=0.016, ICC=0.962). Remaining $p>0.05$.	Fixed method is more correlated with the Biodex in KEX testing. Limits of agreement test should have been used instead of Pearson *r*.

(Kollock Jr, et al., 2013)	Belt and strain gauge.	HEX, HFL, HAB, HAD, HIR and HER.	Yes (isokinetic at 30°.sec^{-1} and 60°.sec^{-1})	Yes (intra-session)	9 males and 9 females (n=18).	22±3	Validity (Pearson r): HEX=0.77, HFL=0.80, HAD=0.82, HAB=0.69, HIR=0.60, HER=0.87 Reliability (ICC): HEX=0.92, HFL=0.66, HAB=0.90, HAD=0.90, HER=0.88, HIR=0.78	Isometric values were compared against isokinetic values. Highest correlation was seen in HAD and HER testing. The highest reliability was noted with HAB and HAD testing.
(Nadler, et al., 2000)	Wooden frame with a HDD.	HEX and HAB	No	Yes (intra- and inter-tester in two sessions)	7 males and 3 females (n=10).	25 to 35	Reliability (ICC): Maximum values, HAB=0.95 and HEX=0.94; Average values, HAB=0.98 and HEX=0.98.	Bulky stabilising station. No validity test was performed.
(Scott, et al., 2004)	Wooden frame with a HDD - Dynamometer Anchoring Station (DAS).	HFL, HAB and HEX.	No	Yes (intra- and inter-tester in two sessions)	5 males and 10 females (n=15).	23 to 44	DAS Intra-tester reliability (ICC) for tester 1: HFL=0.87, HAB=0.84 and HEX=0.59 and for tester 2: HFL=0.85, HAB=0.76 and HEX=0.72. HHD Intra-tester reliability (ICC) for tester 1 only: HFL=0.81, HAB=0.67 and HEX=0.74. DAS Inter-tester reliability (ICC): HFL=0.84, HAB=0.75 and HEX=0.56.	DAS has good intra- and inter-tester reliability. It also came superior to the HHD in testing the strength of HFL and HAB.
(Tourville, et al., 2013)	Frame with a load cell.	HAB, HAD, HFL and HEX.	No	Yes (intra- and inter-tester in three sessions)	10 males and 9 females (n=19).	Males: 23.1±3.1; females: 20.0±0.7	Intra-tester reliability (ICC): HFL=0.91, HEX=0.78, HAB=0.91 and HAD=0.92. Inter-tester reliability (ICC): HFL=0.91, HEX=0.74, HAB=0.91 and HAD=0.92.	The participants were stabilised in the pelvic region which may have increased the measured values. Positioning the participants and setting up the frame is a complicated process.

(Traina, et al., 2010)	Frame with a load cell.	HAB	No	Yes (intra- and inter-tester in two sessions)	6 males and 4 females (n=10).	25 to 35	Intra-tester reliability (ICC) for examiner 1: Rt=0.90, Lt=0.85; examiner 2, Rt=0.96, Lt=0.98. Inter-tester reliability in session 1, Rt=0.84, Lt=0.81 and session 2, Rt=0.94, Lt=0.88. CV% for examiner 1 was 4.93% and for examiner 2 it was 3.88%. Inter-rater CV% was 4.23%.	No validity test was performed. High ICC values and low CV% for HAB testing.
(Widler, et al., 2009)	Frame with HHD.	HAB	Yes (EMG activity)	Yes (inter-session)	8 males and 8 females (n=16).	31±6	Validity: Lower side-to-side HAB muscles activity in side-lying position (<100%, p<0.01) Reliability (ICC, 95% CI [lower, higher]): Side-lying=0.90 [0.74,0.97], supine=0.83 [0.57,0.94], standing=0.88 [0.69,0.96]. CV% was 3.67% for side-lying position, 6.11% for supine, and 4.22% for standing positions.	Multi-use simple frame. Testing the hip HAB using the frame in side-lying position is more reliable than standing and supine positions.
(Wollin, et al., 2016)	Platform with a load cell.	KFL	No	Yes (inter-tester in one day and intra-tester in one week)	Male Football players (n=21).	16.81±0.54	Intra-tester reliability (ICC): Rt=0.91, Lt=0.80 and combined=0.86. Inter-tester reliability: Rt=0.91, Lt=0.83 and combined=0.87.	Excellent reliability for using the platform in testing hamstrings muscles strength in young football players. No validity tests were performed.

Infor., information; M, mean; SD, standard deviation; PVC, ployvinyl chloride; ORI-SST, Orthopaedic Research Institute Shoulder Strength Tester; HHD, hand-held dynamometer; SAD, shoulder adduction; SAB, shoulder abduction; SIR, shoulder internal rotation; SER, shoulder external rotation; HAB, hip abductor; HAD, hip adductor; HER, hip external rotator; HIR, hip internal rotation; HEX, hip extensor; HFL, hip flexion; KEX, knee extensor; KFL, knee flexor; APF, ankle planter flexion; Rt, right; Lt, left; ICC, intra-class correlation; CI, confidence interval; CV, coefficient of variation.

Chapter 3 Study One: Hip Arthroscopy Data Secondary Analysis

3.1. *Chapter Overview*

This study was performed to analyse the secondary data of patients that have undergone hip arthroscopy surgery to treat intra-articular pathologies. The data was analysed to evaluate the changes in hip muscle strength, the range of motion and function over 24-weeks post-arthroscopy. Also, factors such as the sex, age and duration of symptoms were analysed to examine their effects on the pre- and post-operative measurements and surgery outcome.

3.2. *Description of the Data*

Anonymised secondary data was provided by Professor Ernest Schilders*, an orthopaedic consultant in hip and groin surgeries. It was collected in a standardised method by Prof. Schilders and his physiotherapists at the Fortius Clinic (London), the Wellington Hospital (London), and the Yorkshire Clinic (Bingley) in the period from February 2010 to December 2013 (Appendix B. 1).

The dataset consists of repeated measurements of hip muscle strength, range of motion (ROM), function, and pain levels of hip arthroscopy patients (n=316). Descriptive information including sex, date of birth, operation and dominant sides, duration of symptoms, main sport, and level of athletic activity was collected from the patients.

The muscle strength, ROM, function and pain measurements were obtained preoperatively and at 2 weeks, 8 weeks (2 months), 24 weeks (6 months) and 52 weeks (12 months) postoperatively. The strength of the hip abductor (HAB) and adductor (HAD) muscles was measured at these intervals using a hand-held dynamometer

* Consultant at the Fortius Clinic (London), the Wellington Hospital (London), Yorkshire Clinic (Bingley), and a Visiting Professor in Orthopaedic Sports Medicine at the Carnegie School of Sport, Leeds Beckett University.

(HHD). Hip internal rotation (HIR), external rotation (HER), and overall hip joint ROM were assessed using a 360° plastic goniometer and the Flexion-Abduction-External Rotation (FABER) test distance. The measurements were followed by self-completion of the Modified Harris Hip Score (MHHS) form, which was used to monitor pain and function changes over time (Appendix A. 3). The form was filled at all intervals except 2 weeks postoperatively. Lastly, the Visual Analogue Scale (VAS) was used to assess the changes in pain. All variables were analysed using the Statistical Package for the Social Sciences (SPSS) version 22 (IBM, Chicago, IL). Overall, the dataset contained the information on 316 patients and 67 variables (Table 3.1).

Table 3.1 - Variables in the hip arthroscopy dataset (Count = 67)

Sex	Abduction (Rt) at 24 Weeks	External Rotation (Rt) at 52 Weeks
Level of Activity	Abduction (Lt) at 24 Weeks	External Rotation (Lt) at 52 Weeks
Age (years)	Abduction (Rt) at 52 Weeks	Faber Test (Rt) Preoperative
Duration of Symptoms	Abduction (Lt) at 52 Weeks	Faber Test (Lt) Preoperative
Operation Side	Internal Rotation (Rt) Preoperative	Faber Test (Rt) at 2 Weeks
Dominant Side	Internal Rotation (Lt) Preoperative	Faber Test (Lt) at 2 Weeks
Adduction (Rt) Preoperative	Internal Rotation (Rt) at 2 Weeks	Faber Test (Rt) at 8 Weeks
Adduction (Lt) Preoperative	Internal Rotation (Lt) at 2 Weeks	Faber Test (Lt) at 8 Weeks
Adduction (Rt) at 2 Weeks	Internal Rotation (Rt) at 8 Weeks	Faber Test (Rt) at 24 Weeks
Adduction (Lt) at 2 Weeks	Internal Rotation (Lt) at 8 Weeks	Faber Test (Lt) at 24 Weeks
Adduction (Rt) at 8 Weeks	Internal Rotation (Rt) at 24 Weeks	Faber Test (Rt) at 52 Weeks
Adduction (Lt) at 8 Weeks	Internal Rotation (Lt) at 24 Weeks	Faber Test (Lt) at 52 Weeks
Adduction (Rt) at 24 Weeks	Internal Rotation (Rt) at 52 Weeks	VAS Preoperative
Adduction (Lt) at 24 Weeks	Internal Rotation (Lt) at 52 Weeks	VAS at 2 Weeks
Adduction (Rt) at 52 Weeks	External Rotation (Rt) Preoperative	VAS at 8 Weeks
Adduction (Lt) at 52 Weeks	External Rotation (Lt) Preoperative	VAS at 24 Weeks
Abduction (Rt) Preoperative	External Rotation (Rt) at 2 Weeks	VAS at 52 Weeks
Abduction (Lt) Preoperative	External Rotation (Lt) at 2 Weeks	MHHS Preoperative
Abduction (Rt) at 2 Weeks	External Rotation (Rt) at 8 Weeks	MHHS at 8 Weeks
Abduction (Lt) at 2 Weeks	External Rotation (Lt) at 8 Weeks	MHHS at 24 Weeks
Abduction (Rt) at 8 Weeks	External Rotation (Rt) at 24 Weeks	MHHS at 52 Weeks
Abduction (Lt) at 8 Weeks	External Rotation (Lt) at 24 Weeks	Distraction Test

Rt, right; Lt, left; VAS, visual analogue scale; MHHS, modified Harris Hip Score

3.3. ***Research Significance, Contribution and Questions***

3.3.1. Research and Clinical Significance

Hip arthroscopy procedures result in favourable improvements in hip mechanics, muscle strength, quality of life and function for an extended period (Byrd & Jones, 2009; Casartelli, et al., 2014). Moreover, the pattern of clinical changes differs from a group of patients to another based on several factors (McCarthy, et al., 2011; McCarthy & McMillan, 2013). Thus, the evaluation of medical procedures such as hip arthroscopy (HA) or total hip arthroplasty (THA) is not limited to the success of the procedure itself. It extends to include the return to optimal and pain-free functional status with good prognosis. Therefore, longitudinal studies are essential to determine the degree of success and the potential factors that may influence clinical improvement. The results of the current longitudinal studies in the literature are limited regarding the sample size and characteristics, and outcome measures (Daivajna, et al., 2015; Kemp, et al., 2014b). The researcher believes that there is a demand for short- to medium-term longitudinal studies that include larger samples of patients and multiple outcome measures. Another point to consider is the lack of research that discusses between-patients variation in clinical progression based on their sex and level of athletic activity.

In this study, the postoperative changes in muscle strength, the range of motion, function and pain levels of hip arthroscopy patients were evaluated using various statistical tests to look at the pattern of these changes. Another aspect of this study investigated the effects of age, sex, the level of athletic activity and the duration of symptoms on clinical outcome and recovery. Lastly, the muscle strength, the ROM, and the MHHS values were summarised at set intervals.

The significance of this study is based on the design and statistical tests that were used. The patients were followed-up for 24 weeks to look at the pattern of clinical changes during this phase. Therefore, the clinician can estimate the time it will take for the muscles to gain strength, and for the range of motion, function and pain to improve. When the factors that affect the prognosis are considered, the clinician can have a better

understanding of why some patients perform better than other. Moreover, it will be beneficial to design slightly different rehabilitation protocols based on the patient's sex, level of athletic activity and age. Also, it will help in giving the patient insight into postoperative prognosis and eligibility for surgery if possible. Lastly, significant deviations from the reference values will alert the clinician to a potential underlying problem affecting progression and recovery.

3.3.2. Novel Contribution and Strengths

- The sample of patients in the present study is heterogeneous regarding sex, age and level of athletic activity. It is one of the few studies that considered various aspects of postoperative physical and functional changes in a group of mixed hip arthroscopy patients.
- It is one of the largest studies to report the clinical changes following hip arthroscopy. The sample consists of a more prominent cohort of patients (n=316), compared to the smaller sample sizes of the previous studies (section 2.5).
- The study explored the simultaneous postoperative clinical changes in hip muscle strength, the range of motion, function, and pain over 24 weeks.
- The multitude and diversity of the outcome measures in this study looking at the different aspects of clinical changes is a major advantage.
- The patients were followed-up for 24 weeks, which falls between short- to medium-term. The previously published studies followed-up with the patients after one or two years from the arthroscopy.
- The study results identified to which extent do age, sex, the level of athletic activity and duration of symptoms affect postoperative changes. The findings complement the research by Kemp et al. (2014a), Philippon et al. (2010) and McCormick et al. (2012).

3.3.3. *Research Questions (RQ)*

1. What are the values of hip muscle strength, functional and pain levels of hip arthroscopy patients preoperatively, and at 2, 8, and 24 weeks postoperatively?
2. What are the changes in hip muscle strength, the range of motion, function and pain levels of hip arthroscopy patients over the period of 24 weeks?
3. What are the variations in hip muscle strength, function, and pain level measurements between- and within-patient subgroups?
4. What are the effects of age, sex, physical activity level and the duration of symptoms on postoperative changes?

3.4. ***Ethical Consideration***

3.4.1. *Personal Data and Anonymity*

Standardised data collection was achieved through a questionnaire and physical assessment of muscle strength and hip ROM by experienced clinicians. The collected data were managed using the Standardised Orthopaedic Clinical Research and Treatment Evaluation Software (SOCRATES). The data contains no photographic or videotape recordings of the patients and are not available in the public domain. The anonymity of data was maintained by omitting the names of patients and any personal (identity-related) records before the researcher received the data. The data was secured in an encrypted USB stick with no hard copies. Raw anonymised data were stored securely for further research.

3.4.2. *Ethical Approval*

Initial ethical approval was obtained by Professor Schilders (Appendix B. 1). Final ethical approval to analyse the data was obtained on the 27th of April 2015. It was signed and approved by the Director of Studies (DoS) and the Local Research Ethics Coordinator (LREC) at Leeds Beckett University (Appendix D. 1).

3.5. *Methodology*

3.5.1. General Design

The secondary data were analysed in several steps. In the first phase of the analysis, the data was tested for normality and checked for outliers and missing values. In the second phase, a detailed analysis using several statistical procedures was performed. Throughout the analysis, results, and discussion, the values obtained from the patients were referred to as scores. The general design of the study is displayed in Figure 3.1.

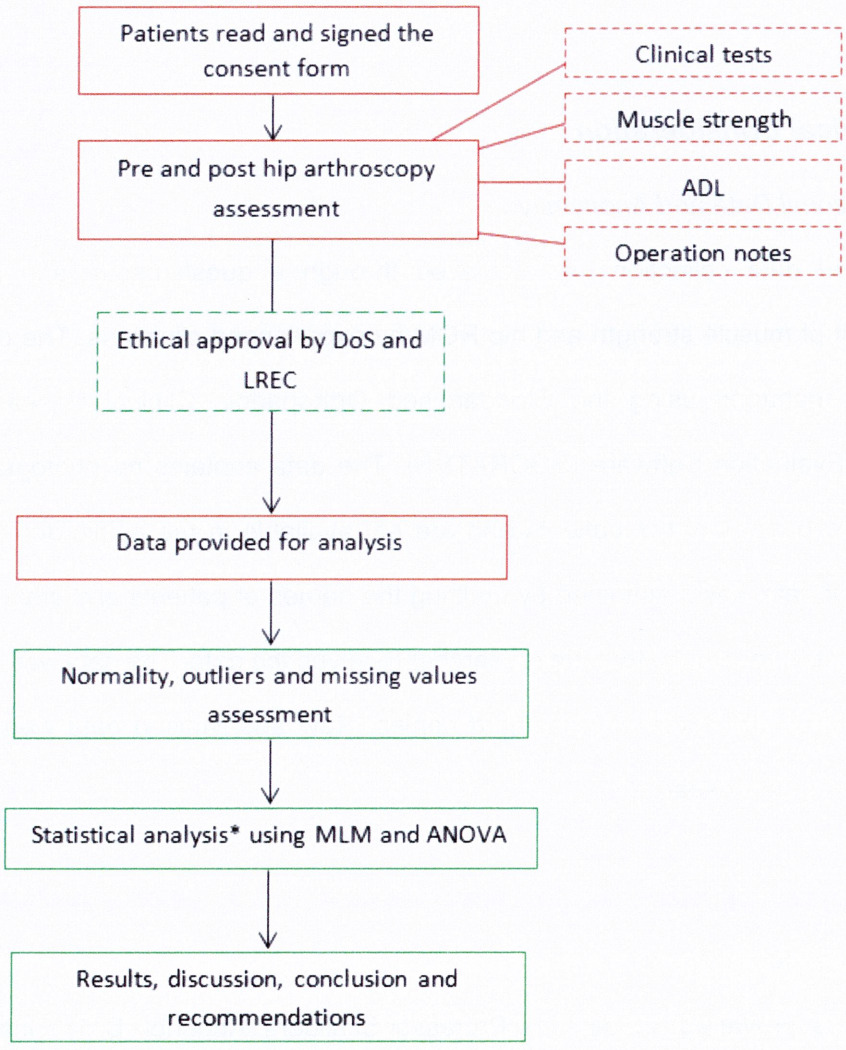


Figure 3.1 - General design of the study showing the major steps from data collection to conclusion (Red boxes refer to the steps performed by Professor Schilders and his staff with no further information; Green boxes refer to the steps performed by the researcher; ADL, activities of daily living; DoS, director of studies; LREC, local research ethics coordinator; *, will be explained in a separate flow chart).

3.5.2. Outcome Measures

The hip muscle strength was measured using the same hand-held dynamometer (Model: MicroFET2, Hoggan Health Industries, Utah, USA) in all clinics. The ROM was measured using a 360° plastic transparent goniometer calibrated with the International Standards of Measurement system. The FABER distance was measured from the lateral edge of the patella to the top surface of the bed with the hip positioned in flexion, abduction, and external rotation. Combined hip pain-function and activities of daily living were measured using the MHHS. The pain level was measured separately using the VAS, which consists of a 10-cm long line used by patients to express pain level.

3.5.3. Statistical Analysis: First Phase

3.5.3.1 Step 1: Normality Assessment

Many statistical procedures are based on the Central Limit Theory. Therefore, the Shapiro-Wilk test, histograms, quantile-quantile (*q-q*) plots, skewness, and kurtosis were used to assess the normality of distribution. Variables with $p > 0.05$ in the Shapiro-Wilk test were classified as normally distributed, while values of $p < 0.05$ were classified as non-normally distributed. Further analysis of the non-normally distributed variables was performed using visual inspection of the histogram, *q-q* plots, and *z*-scores (Equation 3.1). Based on the results of collective tests, most variables were normally distributed. The Detailed analysis and results of the normality tests can be found in Appendix B. 2 and B. 3.

$$z = \frac{Skew\ value}{Standard\ Error\ of\ Skewness}, \quad z = \frac{Excess\ Kurtosis\ value}{Standard\ Error\ of\ Kurtosis}$$

Equation 3.1 - Calculating z-scores from Skewness and Kurtosis values

3.5.3.2 Step 2: Outliers

Exploring outliers was performed visually and statistically using the 5% trimmed mean because they affect the mean (Field, 2013). The duration of symptoms was the only variable that required outliers trimming. The detailed analysis and results of the outliers' assessment can be found in Appendix B. 4 and B.5.

3.5.3.3 Step 3: Missing Values

The analysis showed that the data contains high percentages (>70%) of missing values in many variables (e.g. Adduction at 52 weeks) (Appendix B. 8). The measurements obtained at 52-weeks postoperatively were excluded from the analysis due to the very high percentage of missing values (Figure 3.2). The detailed analysis of the missing values can be found in Appendix B. 6 to B. 8. A comparison between the different methods of analysing data with missing values can be found in Appendix G.

3.5.3.4 Overview

The data exploration in the first phase showed that some variables contained a high percentage of missing values while others did not follow a normal distribution pattern (Appendix B. 9). Considering these findings, the tests used in the second phase of analysis should be robust against missing values and non-normality, or at least account for these issues (Appendix G.). Therefore, multilevel modelling (MLM), which is based on the analysis of variance (ANOVA) and regression models, was used to assess within- and between-patient variability in scores. Moreover, it was used to evaluate the effect of certain variables such as age and sex on the measurements. This decision was reached based on the unique and powerful features that MLM has over other methods. To fulfil the remaining aims of the study, repeated measures ANOVA was used to compare means of measurements between sides and weeks. A comprehensive review and critique of MLM and ANOVA tests can be found in Appendix G.

3.5.4. Statistical Analysis: Second Phase

3.5.4.1 Step 1: Dividing Data

The patients had an operation on the right, left or both sides (*n*=316). Those who had operations on both sides were excluded (*n*=7). The remaining dataset (*n*=309) was divided into right (*n*=169) or left (*n*=140) based on the operation sides (Figure 3.2).

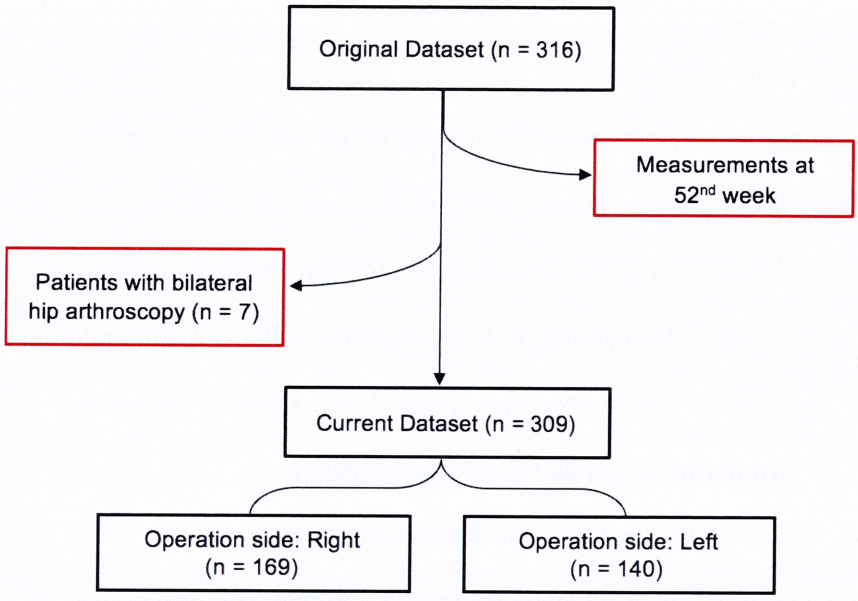


Figure 3.2 - Changes to the dataset after exclusion (red boxes) and dividing patients based on the operation side (black boxes)

3.5.4.2 Step 2: Calculating Strength

The muscle strength measurements of the patients were provided in pounds (lb). Therefore, all muscle strength measurements were converted into Newton (*N*) using the following equation:

$$Muscle\ strength\ in\ Newton\ (N) = Muscle\ strength\ in\ Pound\ (lb) \times 4.4$$

3.5.4.3 Step 3: Recoding and Grouping

- Age was calculated in years from the date of birth (as of 2015). The classification and selection of the age groups were based on the recommendations of the World Health Organization (WHO, 2016). Age was divided into:
    1. Group 1: Adolescent patients (≤19 years old).
    2. Group 2: Adult patients (20 to 59 years old).
    3. Group 3: Older adult patients (≥60 years old) → set as the reference group.
- The duration of symptoms was converted from weeks and years into months and was divided into three groups:
    1. Group 1: One year or less (≤12 months), reflecting acute pain.
    2. Group 2: Two years (13 to 24 months), reflecting subacute pain.
    3. Group 3: More than two years (≥25 months), reflecting very chronic symptoms → set as the reference group.
- The level of athletic activity was divided into three groups:
    1. Group 1: Non-athletes or inactive patients who had no history of playing sports.
    2. Group 2: Amateur athletes who play sports recreationally.
    3. Group 3: Professional athletes who play sports as a career or at elite levels → set as the reference group.

3.5.4.4 Step 4: Determining Factors Affecting Measurements

Based on the available data, information about the patients, and previous research (Keating & Matyas, 1996; Sisto & Dyson-Hudson, 2007), it was determined that age, sex, athletic activity, and the duration of symptoms were likely to affect strength, the range of motion, and function changes in hip arthroscopy patients. Moreover, they were the only available information about the patients. These factors were examined and analysed using MLM.

Step 5: Statistical Analysis Methods

- Repeated Measures Analysis of Variance (*ANOVA*) with post-hoc analysis using Bonferroni adjustment was used to examine the systematic differences between 2, 8, and 24 weeks postoperatively to preoperative baseline scores (RQs 1 and 2, Figure 3.3).

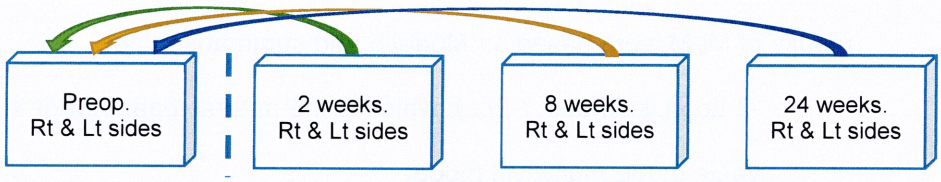


Figure 3.3 - Structure of the repeated measures ANOVA analysis model (Preop, preoperative; Rt, right; Lt, left).

The partial eta squared (ηp^2), an outcome of ANOVA tests, is an estimate of effect size, which measures the magnitude of the difference between scores and the practical significance regardless of the scale used (Lakens, 2013). The results were interpreted as 0 - 0.1, weak effect; 0.1 - 0.3, modest effect; 0.3 - 0.5, moderate effect and >0.5 was considered a strong effect (Levine & Hullett, 2002).

- Multilevel Modelling is a mixed model statistical approach that combines the theoretical bases of regression and ANOVA tests into one framework. It was used to explore the variability in strength, function, ROM, and pain measurement scores of patients with right or left hip arthroscopy. The comparison was made between the measurement scores preoperatively, and at 2, 8, and 24 weeks postoperatively (4-time points) at Level 1 to examine within-patients variability (ε_{oj}) and at Level 2 to examine between-patients variability (μ_{oj}) (RQ 3) (Figure 3.4). It was also used to measure the effects of age, sex, athletic activity level, duration of symptoms and dominance on the measurements (RQ 4). The analysis was performed twice, once for each operated side (Figure 3.4). Building

up the models and interpretation of the results was based on previous recommendations (Field, 2013; Heck, et al., 2014; Steele, 2008). Model 1 was considered a null model, in which no factors were considered. It was used to establish within- and between-patient variabilities. This model was followed by four models, each contained an additional factor to explain the within- and between-patients variabilities in scores (Table 3.2). In order of inclusion, the factors were age, sex, athletic activity level, and the duration of symptoms. The results of MLM were based on Model 5 and summarised as:

- -2 Log Likelihood* (*-2LL)*: which is the natural logarithm of the likelihood value of the multilevel model.
- χ^2 change or the log ratio: the difference in the -2LL of a model when a factor was added to the model.
- The fixed part of the model was presented as:
  - β_{oj} (intercept) or the grand mean of the regression model. It represented the mean of the measurement of interest.
  - The regression coefficient of the regression model and the standard error. The coefficient represented the estimate of reference group relative to the other groups.
- The random part of the model was presented as:
  - Within-participants variation in scores (ε_{oj})
  - Between-participants variation in scores (μ_{oj})

The conclusion was based on Model 5, which included all factors (Table 3.2). Interpretation of the results of Model 5 was dependent on:

- The degrees of freedom (*df*) and -2 Log Likelihood (*-2LL)* change from Model 1 to Model 5, which were compared against the Chi Square (χ^2)

* Likelihood for a model is "proportional to the probability of obtaining the observed data (values for individuals in the sample) if that model was true. The higher the likelihood, the better the fit of the model to the data" (Steele, 2008).

distribution table. After adding the factors, if the change in *df* and *-2LL* was more than the critical value (df_{change} =5, $\chi^2_{0.05}$ =11.1) it was considered a significant change at p =0.05. Therefore, this model was considered to have a statistically significant effect on explaining within- and between-participants' variability in scores. A copy of the Chi Square table can be found in Appendix B. 10.

- The estimate (regression coefficient) of a group or factor was considered to have a significant effect at p <0.05.
- The reference groups in the coefficients table were set as the third or last group of each variable as default by SPSS (v. 22). Further information can be found under each table.

Table 3.2 – The models used in multilevel modelling.

	Model 1	Model 2	Model 3	Model 4	Model 5
Factors	-	Age	Age	Age	Age
	-	-	Sex	Sex	Sex
	-	-	-	Athletic activity	Athletic activity
	-	-	-	-	Duration of symptoms
Variation	ε_j & μ_j	ε_j & μ_j	ε_j & μ_j	ε_j & μ_j	ε_j & μ_j

ε_j, within patients variability; μ_j, between patients variability; Note: Model 1 is a null model and does not contain any factors (Heck, et al., 2014).

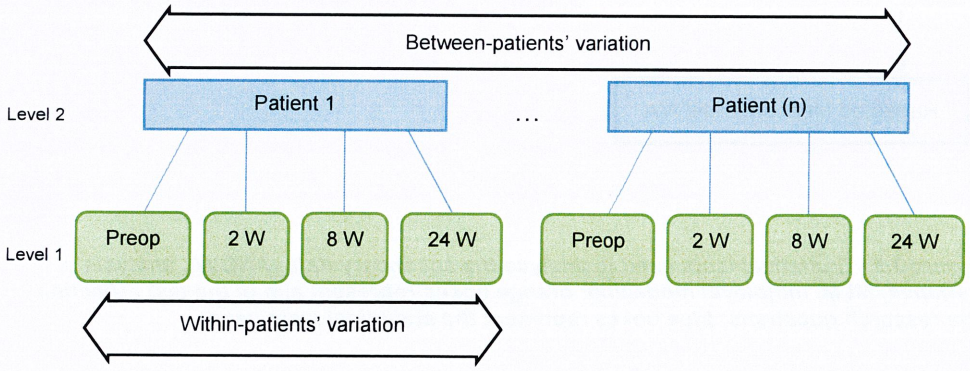


Figure 3.4 - Structure of the multilevel modelling analysis (Preop, preoperative; W, weeks; green boxes represent level 1 and the blue boxes represent level 2 of the model).

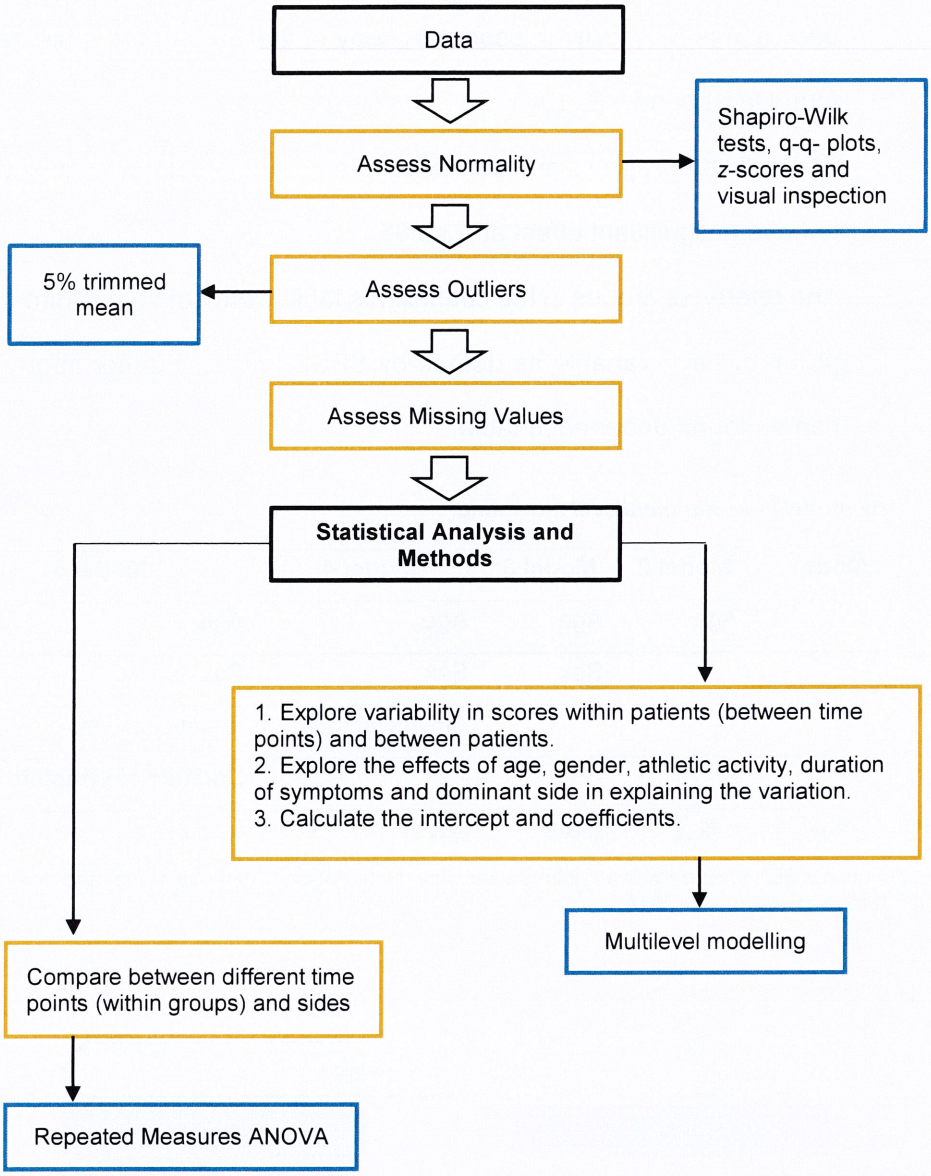


Figure 3.5 - Statistical tests used to analyse the secondary data (ANOVA, analysis of variance; MLM, multilevel modelling; orange boxes represent aim of the test reflecting the research questions; blue boxes represent the statistical tests used).

3.6. *Results*

3.6.1. Descriptive Results

The mean age of the patients was (*M* ± *SD*) 41.4 ± 13.9 years with the male patients (*n*=167) being younger than the females (*n*=142) (39.9 ± 13.2 years and 43.3 ± 14.4 years respectively). The adults made up the majority of patients and males were aged 38.9 ± 10.6 years and females 42.7 ± 9.8 years. The average duration of symptoms was 21.1 ± 23.2 months. About 56% of the patients had the symptoms for less than a year (7.1 ± 3.7 months, *n*=140) and 19% had the symptoms for more than two years (60.3 ± 25.3 months, *n*=48). Just over half of the patients were inactive (51%, *n*=158), and most of them were females (*n*=88). The professional athletes made up just 12% of the sample (*n*=37), and 81% (*n*=30) of them were males (Table 3.3).

Based on the available information, 84% of the patients were right dominant (*n*=161 of 191) compared to 16% as left dominant (*n*=30 of 191). Concerning the operation, the right hip (*n*=169) was operated on more than the left hip (*n*=140) (Table 3.4). Further descriptive findings can be found in Table 3.3 and Table 3.4.

Table 3.3 - Characteristics of the hip arthroscopy patients (1).

Parameters	Males	Females	Overall
Age [in years]	39.9 ± 13.2 (165)	43.3 ± 14.4 (141)	41.4 ± 13.9 (309)
Adolescents (≤19 years old) (*n*)	18.0 ± 1.4 (8)	17.3 ± 1.2 (13)	17.6 ± 1.3 (21)
Adults (20 to 59 years old) (*n*)	38.9 ± 10.6 (145)	42.7 ± 9.8 (111)	40.6 ± 10.4 (256)
Older Adults (≥60 years old) (*n*)	66.1 ± 5.1 (14)	66.7 ± 4.4 (18)	66.4 ± 4.6 (32)
Duration of Symptoms [in months]	21.3 ± 22.9 (131)	20.8 ± 23.7 (118)	21.1 ± 23.2 (249)
Less than a year (≤12 months) (*n*)	6.8 ± 3.8 (69)	7.3 ± 3.7 (71)	7.1 ± 3.7 (140)
One to two years (13 to 24 months) (*n*)	22.4 ± 22.3 (36)	22.3 ± 3.2 (25)	22.4 ± 3.2 (61)
More than two years (≥25 months) (*n*)	58.2 ± 25.8 (26)	62.6 ± 25.1 (22)	60.3 ± 25.3 (48)
Level of Athletic Activity			
Inactive patients (*n*)	70	88	158
Amateur players (*n*)	67	47	114
Professional players (*n*)	30	7	37
Total (*n*)	167	142	309

Notes: *(n)* is the number of valid cases only; values are presented as mean ± standard deviation (*M* ± *SD*) where applicable; NA, not available.

Table 3.4 - Characteristics of the hip arthroscopy patients (2).

Parameters		Males	Females	Total
Dominant Side (*n*)				
	Right (*n*)	84	77	161
	Left (*n*)	15	15	30
Operation Side (*n*)				
	Right (*n*)	81	88	169
	Left (*n*)	86	54	140
Total (*n*)		167	142	309

Notes: *(n)* is the number of valid cases only.

3.6.2. RM-ANOVA Results (RQs 1 and 2)

3.6.2.1 Changes in Strength

- Performing the repeated measures ANOVA reduced the sample size due to missing values*. The hip abductor (HAB) and adductor (HAD) muscle strength of the patients were reduced at the second week postoperatively compared to preoperative baseline measurements. In HAB (n=54), there were statistically significant differences between all measurements with modest effect size between preoperative and 2 weeks, 8 weeks and 24-weeks postoperative values (F [2.37, 125.8] =29.41, ηp^2 =0.357, p <0.001). The highest improvement was at 24 weeks (F [1, 53] =25.59, ηp^2 =0.326, MD= 16.57, 95% CI [7.59, 25.55], p <0.001, Table 3.5). The gain in strength was equivalent to 21.4% and 17.5% for the right and the left sides respectively (Figure 3.6).
- In HAD (n=53), there were statistically significant changes between all HAD strength measurements with moderate effect size between preoperative, 8 and 24 weeks (F [2.55, 132.8] =33.36, ηp^2 =0.391, p <0.001). The highest improvement of 27.0% and 18.2% for the right and left sides respectively was observed at 24 weeks (F [1, 52] =38.19, ηp^2 =0.423, MD= 18.29, 95% CI [10.17, 26.40], p <0.001) (Table 3.5, Figure 3.7). There were no significant differences

* The ANOVA tests follow a listwise deletion approach. Patients with any missing values were excluded from the analysis.

between the right and left operated sides in HAB (p =0.215) and HAD strength measurements (p =0.954).

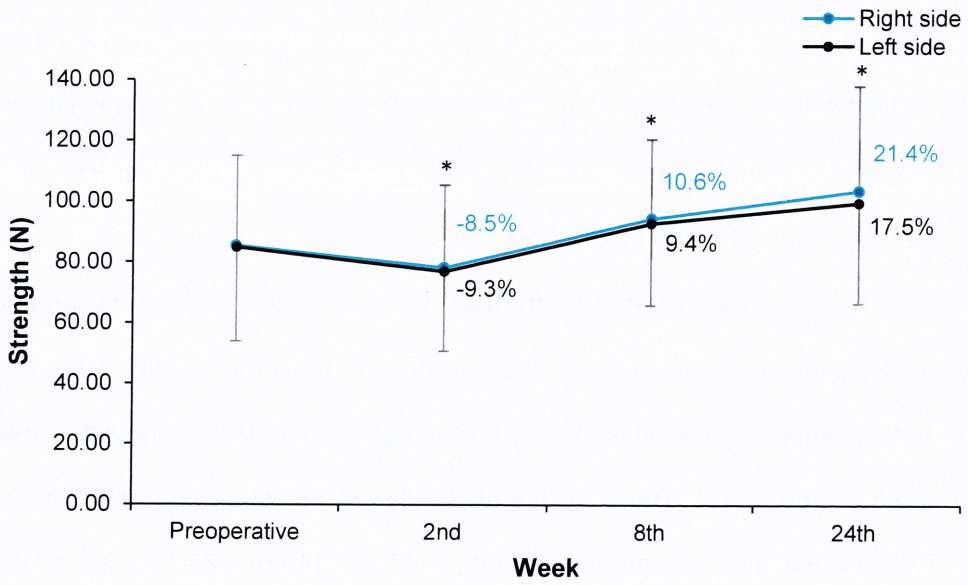


Figure 3.6 - Hip abductor muscle strength changes pre- and post-operatively based on the RM-ANOVA results and operation side (n=54) (With standard deviation bars and percentage of change from pre-operative measurements; *, statistically significant difference at 0.05 from pre-operative measurements).

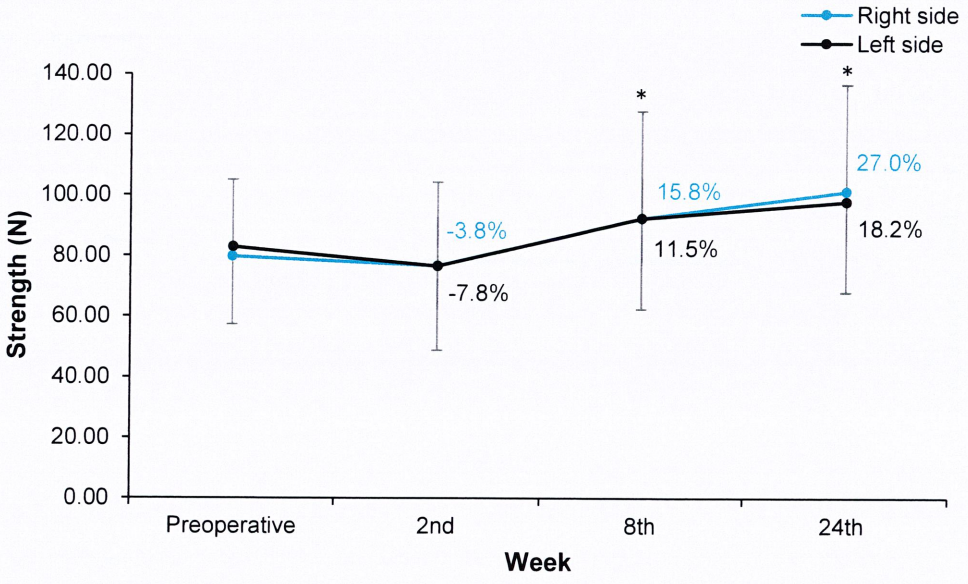


Figure 3.7 - Hip adductor muscles strength changes pre- and post-operatively based on the RM-ANOVA results and operation side (n=53) (With standard deviation bars and percentage of change from pre-operative measurements; *, statistically significant difference at 0.05 from pre-operative measurements).

3.6.2.2 Changes in ROM and Flexibility

- The hip external (HER) and internal rotation (HIR) ROM did not show statistically significant changes postoperatively, with the highest ROM improvement observed at the 24th week (p =0.339, Table 3.5). HER ROM (n=51) increased by 9.5% and 5.7%, while the HIR ROM (n=51) rose by 6.7% and 9.6% for the right and left sides respectively at the 24th week. There were no significant differences between the right and left operated sides in HER (p =0.824) and HIR ROM (p =0.627, Figure 3.8 and Figure 3.9).
- The FABER test distance increased for both sides at the second week postoperatively followed by a steady reduction towards the 24th week (n=60). The distance decreased by 18.2% on the right and 13.0% on the left at the 24th week (Figure 3.10). There were statistically significant differences between all FABER test distance measurements (F [2.54, 149.6] =25.14, ηp^2 =0.299, p <0.001). The biggest improvement was seen at 24 weeks postoperatively (F [1, 59] =28.20, ηp^2 =0.323, MD= -3.53, 95% CI [-5.34, -1.71], p <0.001, Table 3.6). The right side showed more gain in joint flexibility compared to the left side at the 24th week, but there was no statistical difference between sides (p =0.875).

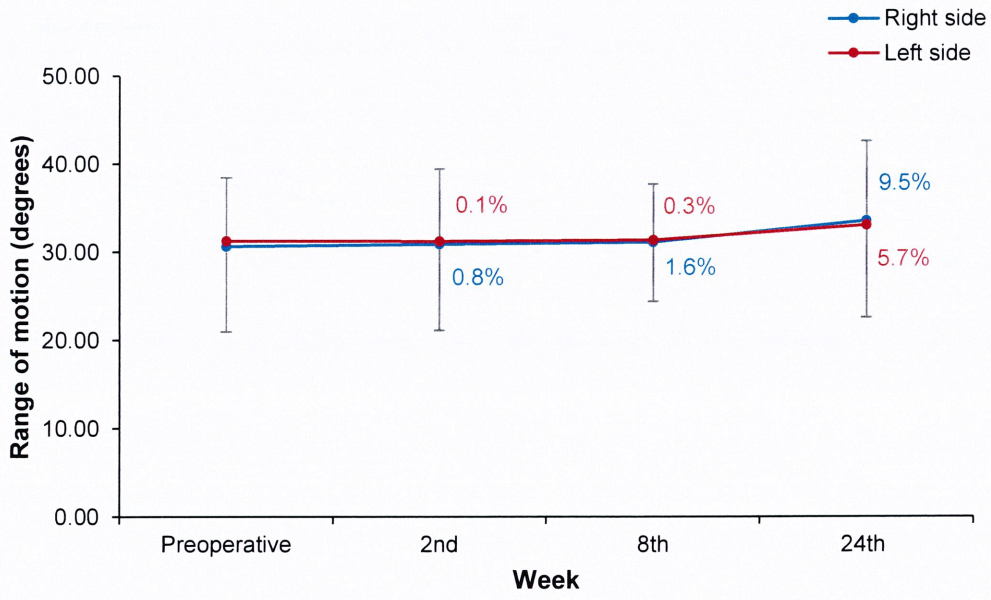


Figure 3.8 - Hip external rotation range of motion changes pre- and post-operatively based on the RM-ANOVA results and operation side (n=51) (With standard deviation bars and percentage of change from pre-operative measurements).

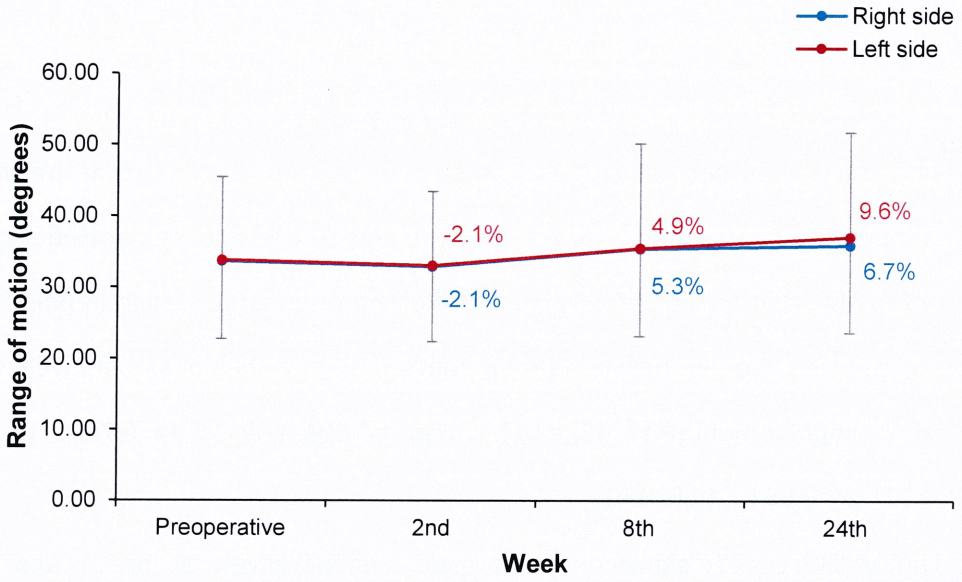


Figure 3.9 - Hip internal rotation range of motion changes pre- and post-operatively based on the RM-ANOVA results and operation side (n=51) (With standard deviation bars and percentage of change from pre-operative measurements).

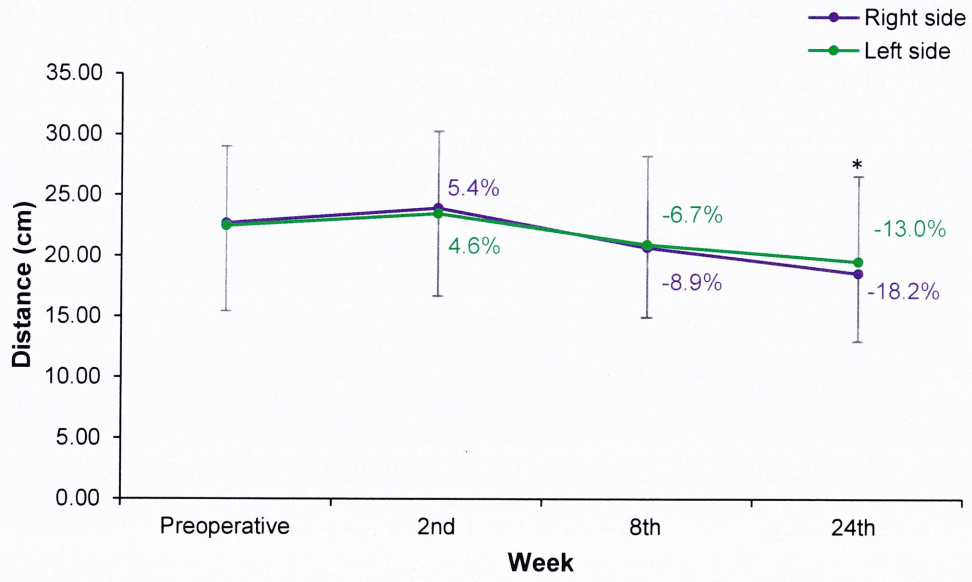


Figure 3.10 - The Flexion-Abduction-External Rotation test distance changes pre- and post-operatively based on the RM-ANOVA results and operation side (n=60) (With standard deviation bars and percentage of change from pre-operative measurements; *, statistically significant difference at 0.05 from pre-operative measurements).

3.6.2.3 Changes in Pain and Function

- Performing the hip arthroscopy procedure resulted in a marked reduction in the pain VAS scores (n=49) (Figure 3.11). There were statistically significant differences between all VAS scores with a strong effect size between preoperative scores, and 2 weeks, 8 weeks and 24 weeks postoperative scores (F [3, 144] =82.98, ηp^2 =0.634, p <0.001). The greatest reduction in pain VAS scores was observed between preoperative and 24 weeks postoperatively with 76.1% improvement (F [1, 48] =197.47, ηp^2 =0.804, MD= -5.44, 95% CI [-6.50, -4.37], p <0.001, Table 3.6).
- The MHHS (n=27) significantly improved postoperatively at the 8th and 24th weeks (F [2, 52] =22.21, ηp^2 =0.461, p <0.001). The highest improvement in scores was seen at the 24th week with a 29.6% increase (F [1, 26] =39.99, ηp^2 =0.606, MD= 19.37, 95% CI [11.53, 27.21], p <0.001, Table 3.6, Figure 3.12). For all reference values, the reader is referred to Appendix B. 11 and Appendix B. 12.
- The positive means of differences in HAB, HAD, IR ROM, ER ROM, and MHHS indicated the increase of strength, the range of motion and function. On the other hand, the negative means of differences in FABER distance and the VAS indicated improved joint flexibility and reduced pain.

Key Results

- Statistically significant differences between measurements over time were found in HAB strength, HAD strength, FABER, VAS and MHHS scores.
- No statistical differences were found in HER and HEX ROM scores.
- The highest differences and strongest effect sizes were between preoperative and 24 weeks postoperative measurements.

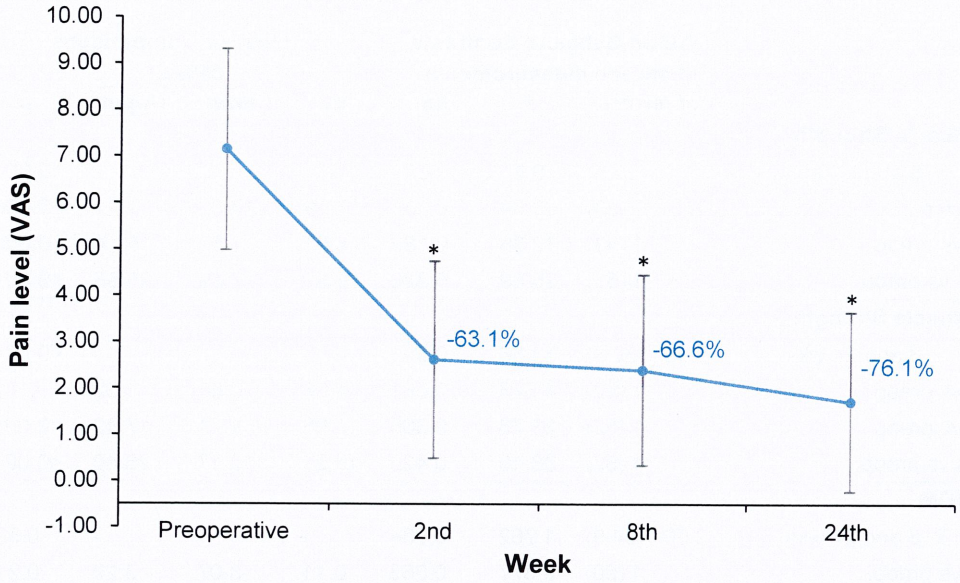


Figure 3.11 - The Visual Analogue Scale changes pre- and post-operatively based on the RM-ANOVA results (n=49) (With standard deviation bars and percentage of change from pre-operative measurements; VAS, visual analogue scale*, statistically significant difference at 0.05 from pre-operative measurements).

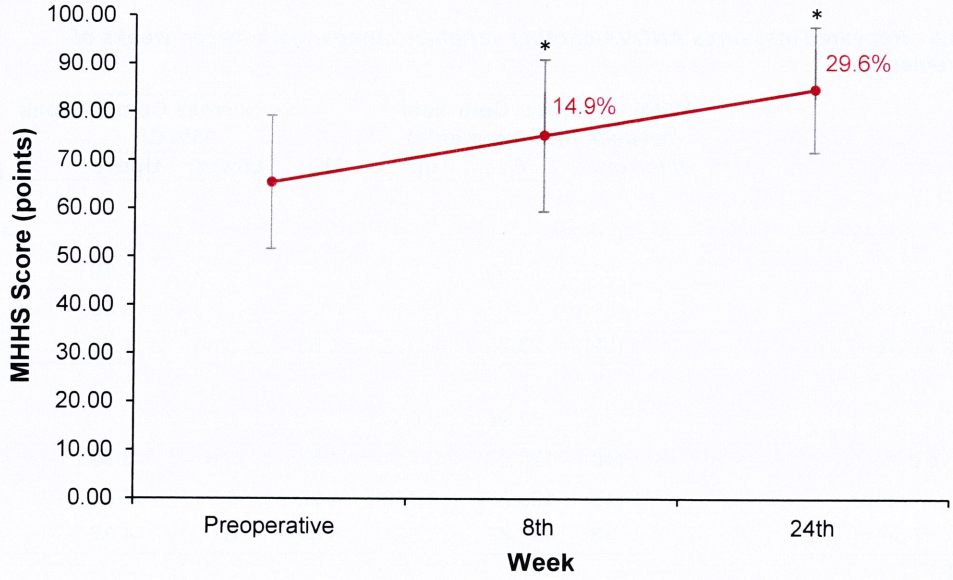


Figure 3.12 - The Modified Harris Hip score changes pre- and post-operatively based on the RM-ANOVA results (n=27) (With standard deviation bars and percentage of change from pre-operative measurements; *, statistically significant difference at 0.05 from pre-operative measurements).

Table 3.5 - Repeated measures ANOVA of strength and range of motion comparing between weeks of measurement.

	Within-Subjects Contrasts (between measurements)			Pairwise Comparisons 95% CI			
	df (error)	*F*	ηp^2	MD	Lower	Upper	*p*
HAB Muscle Strength							
Preop., 2, 8 and 24 wks	2.37 (125.8)	29.41	0.357	-	-	-	<0.001*
2 wks vs preop.	1 (53)	10.64	0.167	-7.59	-13.96	-1.21	0.006*
8 wks vs preop.	1 (53)	12.36	0.189	8.51	1.87	15.14	0.003*
24 wks vs preop.	1 (53)	25.59	0.326	16.57	7.59	25.55	<0.001*
HAD Muscle Strength							
Preop., 2, 8 and 24 wks	2.55 (132.8)	33.36	0.391	-	-	-	<0.001*
2 wks vs preop.	1 (52)	4.578	0.081	-4.67	-10.66	1.32	0.111
8 wks vs preop.	1 (52)	22.78	0.305	11.06	4.71	17.42	<0.001*
24 wks vs preop.	1 (52)	38.19	0.423	18.29	10.17	26.40	<0.001*
HER ROM							
Preop., 2, 8 and 24 wks	2.50 (124.9)	1.762	0.034	-	-	-	0.501
2 wks vs preop.	1 (50)	3.377	0.063	0.11	-3.07	3.29	0.216
8 wks vs preop.	1 (50)	2.391	0.046	0.29	-2.26	2.85	0.384
24 wks vs preop.	1 (50)	2.607	0.050	2.35	-1.17	5.87	0.339
HIR ROM							
Preop., 2, 8 and 24 wks	3 (150)	3.415	0.064	-	-	-	0.057
2 wks vs preop.	1 (50)	0.328	0.007	-0.72	-4.15	2.72	0.569
8 wks vs preop.	1 (50)	2.133	0.041	1.73	-1.52	4.97	0.450
24 wks vs preop.	1 (50)	3.849	0.071	2.76	-1.10	6.61	0.165

HAB, hip abduction; HAD, hip adduction; HER, hip external rotation; HIR, hip internal rotation; ROM, range of motion; Preop. preoperative; wks, weeks; MD, *m*ean of difference; df, degrees of freedom; ηp^2, partial eta squared; *, significant difference at 0.05.

Table 3.6 - Repeated measures ANOVA of other variables comparing between weeks of measurement.

	Within-Subjects Contrasts (between measurements)			Pairwise Comparisons 95% CI			
	df (error)	*F*	ηp^2	MD	Lower	Upper	*p*
FABER D.							
Preop., 2, 8 and 24 wks	2.54 (149.6)	25.14	0.299	-	-	-	<0.001*
2 wks vs preop.	1 (59)	3.399	0.054	1.13	-0.54	2.79	0.210
8 wks vs preop.	1 (59)	7.292	0.110	-1.77	-3.55	0.02	0.027*
24 wks vs preop.	1 (59)	28.20	0.323	-3.53	-5.34	-1.71	<0.001*
VAS							
Preop., 2, 8 and 24 wks	3 (144)	82.98	0.634	-	-	-	<0.001*
2 wks vs preop.	1 (48)	120.03	0.714	-4.53	-5.67	-3.39	<0.001*
8 wks vs preop.	1 (48)	129.38	0.729	-4.76	-5.91	-3.61	<0.001*
24 wks vs preop.	1 (48)	197.47	0.804	-5.44	-6.50	-4.37	<0.001*
MHHS							
Preop., 8 and 24 wks	2 (52)	22.21	0.461	-	-	-	<0.001*
8 wks vs preop.	1 (26)	10.94	0.296	9.74	2.20	17.28	0.006*
24 wks vs preop.	1 (26)	39.99	0.606	19.37	11.53	27.21	<0.001*

Preop., preoperative; wks, weeks; FABER D, distance between the knee and bed; VAS, visual analogue scale; MHHS, modified Harris Hip Score; ηp^2, partial eta squared; MD, *m*ean of difference; df, degrees of freedom; *, significant difference at 0.05.

3.6.3. MLM Results (RQs 3 and 4)

Multilevel modelling was performed to explore within-patients (between different time points) and between-patients (between individuals) variation in scores. The results were based on model 5, which showed statistically significant variations in scores within- and between-patients in all measured variables. Within-patients variation was higher than between-patients in all measurements except HIR ROM of the right side. All values can be found in Table 3.8 and Table 3.9.

Age, sex, athletic group and the duration of symptoms were examined as factors affecting the postoperative outcome. The grand mean for the measurement was represented by the intercept. A positive regression coefficient in MLM indicates that the reference group scored higher than the other group(s) and the vice versa.

Age was the first factor to be considered. The patients aged ≥60 years old (older adults) were considered as the reference group (Table 3.8 and Table 3.9). Age was found to affect some postoperative measurements in which:

- The older adults had lower HER ROM of -5.70° (SE= 1.86°) on the left side only compared to the other groups (p =0.003).
- The older adults had lower HIR ROM of -4.20° (SE= 1.51°) on the right side only compared to the other groups (p =0.006).
- The older adults had a higher FABER distance of 3.43 cm (SE= 1.46 cm) on the left side only (p =0.02).

When sex was examined, the female patients were considered the reference group (Table 3.8 and Table 3.9). Sex was found to affect some postoperative measurements in which:

- The females had lower HAB strength of -32.77 *N* (SE= 3.37 *N*) on the right, and -33.0 *N* (SE= 4.56 *N*) on the left sides compared to males (both p <0.001).
- The females had lower HAD strength of -38.7 *N* (SE= 3.29 *N*) on the right, and -30.82 *N* (SE= 4.30 *N*) on the left sides compared to males (both p <0.001).
- The females had lower HER ROM of -2.82° (SE= 1.36°) on the left side only compared to males (both p =0.041).
- The females had higher HIR ROM of 5.71° (SE= 1.38°) on the right (p <0.001), and 5.49° (SE= 1.69°) on the left sides compared to males (p =0.001).
- The females had lower MHHS scores of -7.21 points (SE= 2.47 points) compared to males (p =0.004).
- The females had higher VAS pain scores of 0.53 (SE= 0.24) compared to males (p =0.029).

Regarding the level of athletic activity, the professional athletes (third group) were considered as the reference group (Table 3.8 and Table 3.9). The level of athletic activity was found to affect postoperative measurements in which:

- The professional athletes had higher HAB strength of 8.61 *N* (SE= 2.79 *N*) on the right (p =0.002), and 13.46 *N* (SE= 3.37 *N*) on the left sides (p <0.001) compared to the other groups.
- The professional athletes had higher HAD strength of 9.89 *N* (SE= 2.71 *N*) on the right (p <0.001), and 14.72 *N* (SE= 3.18 *N*) on the left sides (p <0.001) compared to the other groups.
- The professional athletes had lower MHHS score of -4.54 points (SE= 1.87 points) compared to the other groups (p =0.016).

The duration of symptoms and the dominant side did not have a statistically significant effect on postoperative changes (Table 3.8 and Table 3.9). A summary of the findings can be found in Table 3.7, and detailed results can be found in Table 3.8 and Table 3.9.

Table 3.7 - Summary of the effects of factors on strength, range of motion, function and pain postoperatively (based on Model 5).

	HAB strength	HAD strength	HER ROM	HIR ROM	FABER	MHHS	VAS
Age [≥60 years]	-	-	↓	↓	↑	-	-
Sex [Female]	↓	↓	↓	↑	-	↓	↑
Athletic Group [Professionals]	↑	↑	-	-	-	↓	-

HAB, hip abduction; HAD, hip adduction; HEX, hip external rotation; HIR, hip internal rotation; ROM, range of motion; FABER, flexion abduction external rotation distance; MHHS, modified Harris Hip Score; VAS, visual analogue scale.

Table 3.8 - Results (1) of multilevel modelling based on Model 5 showing the regression coefficients (and standard error) to estimate trends in means of measurement scores.

	Abduction (*N*)		Adduction (*N*)		External Rotation (°)		Internal Rotation (°)	
	OS Rt (SE)	**OS Lt (SE)**	**OS Rt (SE)**	**OS Lt (SE)**	**OS Rt (SE)**	**OS Lt (SE)**	**OS Rt (SE)**	**OS Lt (SE)**
Model 5 Parameters								
df (-2LL) of Model 5	10 (3590.2)	10 (2670.6)	10 (3550.0)	10 (2716.0)	10 (2741.1)	10 (2227.8)	10 (2726.4)	10 (2280.5)
df and χ^2 Change from Model 1	5 (99.8)*	5 (71.8)*	5 (86.8)*	5 (122.3)*	5 (4.1)	5 (17.0)*	5 (26.7)*	5 (15.5)*
Fixed Part (Regression model)								
β_{oj} (intercept)	135.4 (13.36)***	106.4 (18.67)***	135.6 (13.03)***	87.10 (17.62)***	28.72 (4.68)***	45.96 (5.46)***	33.13 (5.32)***	29.66 (6.75)***
Age [≥60 years]	-7.06 (3.64)	2.90 (6.37)	-5.14 (3.55)	9.53 (4.30)	-0.48 (1.33)	-5.70 (1.86)**	-4.20 (1.51)**	-1.45 (2.31)
Sex [Female]	-32.77 (3.37)***	-33.00 (4.56)***	-38.7 (3.29)***	-30.82 (4.30)***	0.58 (1.20)	-2.82 (1.36)*	5.71 (1.38)***	5.49 (1.69)**
Athletic Group [Professionals]	8.61 (2.79)**	13.46 (3.37)***	9.89 (2.71)***	14.72 (3.18)***	1.33 (0.97)	0.16 (0.99)	0.46 (1.13)	-0.20 (1.22)
Duration of Symptoms [≥2 years]	-1.51 (1.58)	-3.40 (1.89)	-0.99 (1.54)	-2.58 (1.78)	-0.41 (0.54)	-0.71 (0.55)	-0.98 (0.63)	-0.95 (0.68)
Random Part (ANOVA model)								
Within-Var. (ε_{oj})	470.5 (43.07)***	417.7 (46.68)***	426.2 (39.21)***	400.8 (43.34)***	48.24 (4.48)***	64.01 (6.64)***	36.57 (3.37)***	61.43 (6.27)***
Between-Var. (μ_{oj})	203.8 (48.72)***	335.7 (74.28)***	203.3 (46.6)***	290.7 (63.90)***	28.65 (6.10)***	21.28 (6.75)**	50.20 (7.76)***	48.31 (9.50)***

OS, operation side; Rt, right; Lt, left; SE, standard error; *N,* Newton; °, degrees; β_{oj}, intercept representing grand mean; Within-Var. (ε_{oj}) = within participants variance; Between-Var. (μ_{oj}) = between participants variance; significant increase or decrease or significant variation in scores at p <0.05, *; p <0.01, **; p <0.001, ***; Note 1: Intercept, coefficients and variation estimates are based on Model 5 results; Note 2: Reference group in categorical variables is the group between the square brackets [..].

Table 3.9 - Results (2) of multilevel modelling based on Model 5 showing the regression coefficients (and standard error) to estimate trends in means of measurement scores.

	FABER D. (cm)		MHHS (points)	VAS (cm)
	OS Rt (SE)	OS Lt (SE)		
Model 5 Parameters				
df (-2LL) of Model 5	10 (2558.7)	10 (2033.1)	9 (2305.9)	10 (3382.2)
df and χ^2 Change from Model 1	5 (9.2)	5 (15.2)*	5 (10.1)	5 (8.8)
Fixed Part (Regression model)				
β_{oj} (intercept)	24.89 (3.57)***	22.11 (4.31)***	91.41 (9.96)***	1.93 (0.04)*
Age [≥60 years]	1.13 (1.02)	3.43 (1.46)*	-0.83 (3.21)	0.42 (0.28)
Sex [Female]	-0.94 (0.92)	-1.46 (1.07)	-7.21 (2.47)**	0.53 (0.24)*
Athletic Group [Professionals]	-1.60 (0.75)*	-1.43 (0.78)	-4.55 (1.87)*	0.11 (0.18)
Duration of Symptoms [≥2 years]	0.23 (0.42)	0.28 (0.44)	0.92 (1.02)	0.17 (0.10)
Random Part (ANOVA model)				
Within-Var. (ε_{oj})	25.82 (2.33)***	25.44 (2.61)***	335.8 (43.96)***	8.65 (0.47)*
Between-Var. (μ_{oj})	18.48 (3.55)***	18.65 (3.96)***	4.89 (33.40)	NA

OS, operation side; Rt, right; Lt, left; cm, centimetres; FABER, flexion abduction external rotation; MHHS, modified Harris Hip Score; VAS, visual analogue scale; β_{oj}, intercept representing grand mean; Within-Var. (ε_{oj}) = within participants variance; Between-Var. (μ_{oj}) = between participants variance; NA, not available for statistical analysis because the variance cannot be calculated; significant increase or decrease or significant variation in scores at p <0.05, *; p <0.01, **; p <0.001, ***; Note 1: Intercept, coefficients and variation estimates are based on Model 5 results; Note 2: Reference group in categorical variables is the group between the square brackets [..].

3.7. *Discussion*

The current study was performed to analyse the data of 316 hip arthroscopy patients. Hip abduction (HAB) and adduction (HAD) muscle strength showed remarkable improvements with the biggest changes seen in the 24th week postoperatively. Similar changes were observed in pain and function. Also, age, sex and the level of athletic activity were found to affect strength, the range of motion (ROM), function and pain at varying levels. Details of the diagnosis and the intra-operative procedures of each patient were not provided as part of the dataset. Therefore, the analysis was based on the results of physical clinical examinations and patient reported outcomes.

The presence of missing values constrained the full potentials of using such valuable information. The dataset was not complete because the medical insurance policies don't cover follow-up appointments after 24 weeks (6 months) of the arthroscopy. Moreover, the professional athletes usually don't return for follow-up appointments after 8 weeks (2 months) from the start of their rehabilitation programs. These explanations were based on clinical observations, which could explain the reduction in the sample size to <60 once the ANOVA tests were applied.

Femoroacetabular impingement (FAI) is one of the leading indications to perform hip arthroscopy. Cam impingement occurs mostly in males with idiopathic causes, while pincer impingement is genetic and occurs mainly in females (Samora, et al., 2011). FAI can be found in about 30% of the symptom-free population, and in 17% of groin pain patients (Röling, et al., 2016). Furthermore, whites are more susceptible to developing FAI at a young age compared to the Chinese (Dudda, et al., 2011; Houck, et al., 2004). More importantly, a genetic component to the development of FAI has recently been described. The siblings of patients with a cam or pincer impingement are at a higher risk of developing a similar deformity (Pollard, et al., 2010). The patients included in this study could have also had hip dysplasia, who usually don't have reduced ROM compared to the patients with FAI. On the contrary, they are characterised with

hypermobility (Sturridge & Bankes, 2015). Combining the patients affected with these conditions may have affected the results and conclusions drawn from the analysis.

Almost 49% of the patients were active at the time of surgery. Out of these patients, about 75% of them are physically active at a recreational or an amateur level. However, the remaining 51%, who were classified as inactive, may have participated in physical activity or sport before having the surgery. Another point to consider is that 48% of the patients had the symptoms for more than one year and some of them for up to 7 years. The presence of intra-articular pain may have ceased any moderate to high-intensity activities such as running or playing football. A previous study has shown that FAI is more prevalent in the active population. Moreover, active individuals playing high-impact sports such as basketball and hockey are at a significantly higher risk of developing cam impingements with odds ratio between 1.9 and 8.0 (Kuhns, et al., 2015). However, the full medical and physical activity history of the patients was not available for further analysis.

Key Points

- Combining patients with various diagnoses could have affected the results.
- Almost half of the participants were physically active before the arthroscopy, which puts them at a higher risk of developing FAI.

3.7.1. Muscle Strength

The preoperative hip muscle strength values of the current study are 35% less than previously reported normative values of healthy individuals (Appendix B. 11). A previous study based on healthy participants (n=231) showed that the average HAB muscle strength was 240.6 *N* (Bohannon, 1997c). Another study reported HAB strength values as high as 303.4 ± 64.6 *N* for males (n=25) and 202.5 ± 40.6 *N* for females (n=25) aged 50 to 59 years old (Andrews, et al., 1996). Long-standing joint pain and reduced intensity

of physical activity may have caused muscular weakness over time. Moreover, combining patients of varying age groups and activity levels may have also lowered the average muscle strength scores and increased the standard deviation.

The pre-operative HAB muscle strength of the patients was lower than the HAD strength. Patients with intra-articular hip pathology are often presented with lateral hip pain, which inhibits muscular activity (Safran, 2005). A study based on EMG revealed that patients with symptomatic FAI have weaker HAB, HAD, and hip flexion (HFL) muscles compared to healthy individuals (Casartelli, et al., 2011). From a clinical perspective, HA patients often experience HFL muscle insufficiency, which results in over-activity of the rectus femoris and TFL muscles as a compensatory mechanism. This process predisposes the patients to develop tendonitis (Casartelli, et al., 2011; Edelstein, et al., 2012; Enseki, et al., 2010). The gluteus maximus (GMax), can also be affected because it inserts into the gluteal tuberosity and TFL. A study based on TMG analysis of this muscle indicated a statistically significant increase in the contraction time on the impingement side (37.0 ± 9.5 ms) compared to the healthy side (32.9 ± 7.2 ms, p =0.01), which suggests a reduced ability of the muscle to rapidly generate tone upon contraction (Seijas, et al., 2016). Thus, the functional abilities of this muscle could be compromised.

Intra-articular hip pathology can be misdiagnosed as a groin injury because the patients commonly experience radiating pain to the groin area. The squeeze test, which recruits the hip adductor muscles, is a commonly used clinical test to detect adductor related groin pain (Holmich, et al., 2004). The HAD muscle strength is measured by squeezing a HHD or a pressure cuff placed between the knees (Light & Thorborg, 2016). When a pressure cuff was used in testing, the HAD strength of the groin pain patients (172.3 ± 28.2 mmHg) was significantly less than the strength of the healthy individuals in the squeeze test at 0° of hip flexion (210.8 ± 39.3 mmHg, p =0.01) (Malliaras, et al., 2009). Weakness, atrophy or reduced activation of the hip muscles can result from intra-articular pathology, a groin injury or fear of pain (Casartelli, et al., 2011).

The postoperative changes in HAB and HAD muscle strength were expected. In the current study, the drop of 3.8% to 9.3% in muscle strength two weeks after the surgery can be due to muscle deactivation and joint inflammation, which have been reported in hip surgeries and knee ligamentous pathologies (Holm, et al., 2013; Hart, et al., 2010). Although the procedure was performed on one side only, bilateral neuromuscular inhibition could have developed. In this type of a physical response, the body reacts to soft tissue disruption by inhibiting muscular activity on both sides. Such an event has been noticed in patients with knee osteoarthritis, chronic lateral ankle sprain and anterior cruciate ligament insufficiency (Charlton, et al., 2016). However, muscular weakness may have been present before the clinical symptoms became apparent (Casartelli, et al., 2011). Nonetheless, a direct causal relationship between the two cannot be established with the available information.

The findings of the current study indicated that it may take eight weeks for the HAB and HAD muscles to show statistically significant improvement in strength. The postoperative strength improvement of the hip arthroscopy patients included in the current study is accompanied by a significant drop of 63.1% in the VAS scores as seen in Figure 3.11. In the following weeks, muscle strength steadily improved to reach the highest level in the 24th week. On that week, HAB muscle strength had improved by 17.5% to 21.4%, while HAD strength improved by 18.2% to 27.0% in relation to the preoperative measurements as illustrated in Figure 3.6 and Figure 3.7. Such findings indicate that muscle strength development could take few weeks and the rehabilitation programs should be designed to accommodate this gradual increase in strength.

Other researchers have also investigated the changes in muscle strength. Holm et al. (2013) reported 41% to 58% (p <0.001) reduction in leg press strength on the second day following surgery in THA patients compared to 3.8% to 9.3% after two weeks in the current study. Philippon et al. (2009) monitored an American Football player (25 years), who was treated arthroscopically for posterior hip subluxation, for nine weeks. He noticed marked increase in the strength of the hip abductor (83%), adductor (80%),

extensor (23%), flexor (65%), external rotator (89%), and internal rotator (57%) muscle groups from preoperative baseline scores. Casartelli et al. (2014) also found noticeable gains in the hip abductor (9%), adductor (34%), internal rotation (59%), external rotation (37%), flexor (23%), and extensor (43%) muscle groups after 2.5 years from the arthroscopy. The results of the current study together with the findings of Philippon et al. (2009) and Casartelli et al. (2014) indicate that postoperative muscle strength requires few weeks to improve, and continue to develop for more than two years compared to preoperative measurements.

The results of the repeated measures ANOVA of the hip arthroscopy data showed a similar pattern of strength improvement in both muscle groups. However, the strength gain of the HAB muscle was slightly lower (17.5% to 21.4%) than HAD muscle (18.2% to 27.0%). The surgical technique used to treat intra-articular pathologies of the central compartment could be the reason. The surgery involves inserting the anterolateral port through the gluteus medius, tensor fascia latae, and the joint capsule. Similarly, the posterolateral portal penetrates the gluteus maximus, medius, and minimus muscles (Dienst, 2005). These minor injuries to the muscle fibres can also play a role in prolonging pain and recovery time. Nonetheless, the technique could differ from one surgeon to another depending on the training, experience, and pathology.

The outcome of multilevel modelling showed that sex and level of athletic activity influence strength (Table 3.7). In terms of sex, the analysis revealed that the strength of the female patients was reduced compared to the male patients (Table 3.7). Kemp et al. (2014b) also noticed that the female patients who had hip arthroscopy to treat chondrolabral pathology had lower hip abductor (*MD*=0.46 N·m.kg^{-1}, 95% CI [0.27, 0.65], p =0.001), adductor (*MD*=0.36 N·m.kg^{-1}, 95% CI [0.18, 0.54], p <0.001), extensor (*MD*=0.23 N·m.kg^{-1}, 95% CI [0.03, 0.43], p =0.033), external rotator (*MD*=0.27 N·m.kg^{-1}, 95% CI [0.15, 0.38], p <0.001) and internal rotator muscle strength (*MD*=0.17 N·m.kg^{-1}, 95% CI [0.08, 0.25], p <0.001) compared to the male patients. Therefore, the findings of this study are not so different from the findings of Kemp et al. (2014b) that female

arthroscopy patients do exhibit more weakness in their hip muscles compared to the males after HA.

Other studies have also shown that the females have weaker muscles. Suzuki et al. (2015) found reduced strength of the hip abductor (males= 21.65 ± 3.82 %BW, females= 18.48 ± 4.79 %BW, p =0.03) and external rotator muscles (males= 24.18 ± 3.90 %BW, females= 18.73 ± 3.75 %BW, p <0.01) of female basketball players compared to male players. The weakness in hip abduction of the females negatively correlated with the apparent knee valgus angle at initial contact (r_s= -0.46, p =0.03). Willson et al. (2012) discovered that the female runners have a 30% increase in muscle peak activation of the gluteus maximus compared to the male runners (males= 61.2 ± 25.5, females= 87.7 ± 41.8, p =0.028, effect size= 0.79). This particular finding is associated with females complaining from patellofemoral syndrome (Willson, et al., 2012). Taylor-Haas et al. (2014) stated that the increase in hip adduction moment found in females is associated with iliotibial band syndrome, patellofemoral pain syndrome and tibial stress fracture. The results of these studies emphasize on the importance of considering the sex based differences in strength. Ignoring such difference could lead to musculoskeletal problems of the lower limbs.

The anatomical variation in the hip and pelvis may play a role in explaining the sex differences in strength. Females have a higher ratio of distance between the hips to the width of the pelvis (males= 0.606 ± 0.034; females= 0.615 ± 0.034) and between pelvic width to pelvic height (males= 1.371 ± 0.068; females= 1.460 ± 0.084) compared to males. These variations could result in a biomechanical disadvantage in females (Brinckmann, et al., 1981).

The level of athletic activity is another factor that was found to affect postoperative HAB and HAD muscle strength (Table 3.8). The professional athletes were on average stronger in HAB and HAD measurements compared to the amateur athletes and inactive patients. These findings were expected since the professional athletes train and

compete at a higher level compared to the amateurs. Professional athletes can go back to playing at a higher rate due to the relief of hip pain, improved mechanics and increased muscle strength. About 96% of the Australian Football League players returned to high-level playing, and 62% of them were still playing after an average of 52.5 months from the operation (Amenabar & O'Donnell, 2013). Philippon et al. (2007) found a rate of 93% return to play in professional Hockey players. In a later study, it was reported that 100% of the Hockey players returned to Hockey drills after an average of 3.8 (range 1 to 5) months from the arthroscopy (Philippon, et al., 2010). No other studies looking at the differences between inactive patients, amateur and professional players were found in the literature. Therefore, further explanation and comparison of the results is not possible at this stage. The effect of physical activity and return to play are further discussed in section 3.7.3.

The data of the current sample were analysed for up to 24-weeks postoperatively. The changes in muscle strength are not limited to the medium-term follow-up period. Casartelli et al. (2014) found significant improvement in the strength of the HAB, HAD, HIR, HER, HFL, and HEX muscle groups at an average of 2.5 years of follow-up (Table 2.4). The lack of long-term studies investigating the changes in muscle strength is an issue that needs to be addressed by researchers. The gain or loss of muscle strength could have an effect on hip and knee kinematics and the level of function (Charlton, et al., 2016; Kemp, et al., 2014a).

Key Points

- Hip muscle strength show a statistically significant improvement after a minimum of eight weeks postoperatively.
- Female patients showed reduced muscle strength scores compared to the males. This issue need to be considered to prevent future complications.

- Professional athletes showed higher muscle strength values compared to the amateur players and inactive patients, which is not surprising.

3.7.2. Range of Motion

The presence of preoperative FAI may lead to a reduction in hip ROM (Kennedy, et al., 2009). Performing HA can improve the hip mechanics, squat range and ROM when conservative management fails (Lamontagne, et al., 2011; Emara, et al., 2011). The analysed sample of hip arthroscopy patients may have included patients diagnosed with FAI, hip dysplasia or both. However, the diagnostic information wasn't available.

The hip ROM values between HA patients and healthy individuals vary. In the current study, the average of all ROM measurements obtained pre- and postoperatively were less than the normative value of 45° for both motions on both sides (Roach & Miles, 1991). Kemp et al. (2014b) indicated that the HA patients had 5° reduction in HIR ROM (p =0.004) and 5° increase in HEX ROM (p =0.019) compared to a control group of healthy participants. The sample of patients included in the current study consisted of patients of varying ages, and 10% of them were classified as older adults. Joint arthritic changes, which become more evident with increasing age, are considered a leading cause of the reduced ROM (Kemp, et al., 2014a).

The results of the current study also showed that HIR and HER ROM did not show statistically significant changes post HA. HIR ROM reduced slightly on the 2nd week then gradually improved showing the highest gain in range at the 24th week. However, HER ROM remained almost the same on the 2nd week and gradually increased afterwards. The statistically insignificant change in ROM opposes the clinical improvements in ROM following hip arthroscopy. Such finding may be due to the reduction in sample size when the ANOVA tests were performed because of the missing values. Another factor could be the skewness of the sample towards adults and older adults or the presence of mixed

pathology of FAI and hip dysplasia. The precise reason behind this observation cannot be identified based on the current information.

However, the postoperative ROM changes found in this study are not that different from previous studies. Philippon et al. (2009) found no remarkable changes in HIR and HER ROM. However, HFL and HAD ROM demonstrated a noticeable improvement of 13° after 9 weeks of follow-up (Table 2.4). Nho et al. (2011) also reported no significant changes in HER (48.7° ± 10.1° to 49.5° ± 2.9°, p >0.05) and HIR ROM (21.7° ± 10.8° to 19.6° ± 10.4°, p >0.05) postoperatively in a group of 64 FAI patients (22.8 ± 6.2 years) (Table 2.4). In contrast to Philippon et al. (2009) and Nho et al. (2011), Brunner et al. (2009) found meaningful improvements in HIR (improved from 6° [range -20° to 45°] to 19° [range -5° to 45°], p <0.001) and HFL ROM (improved from 107° [range 60° to 130°] to 122° [range 70° to 145°], p <0.001) in 53 FAI patients (17 to 66 years) (Table 2.4). The results of Brunner et al. (2009) were different because of the ROM testing method, which was performed in a supine position instead of a sitting position. The latter position is universal and more commonly used to evaluate HER and HIR ROM.

In contrast to the ROM results, the analysis of FABER test distance showed a significant improvement in the joint's flexibility. The increase in the distance at the 2nd week postoperatively is normal due to pain, joint inflammation, and muscle guarding. By the end of the 24th week, there was a significant improvement in the distance compared to the baseline measurement, which was more noticeable on the right side. The reduction in the distance is equivalent to approximately 4 cm on the right side and 3 cm on the left, which are less than the 10 cm improvement reported by Philippon et al. (2009). Several factors can affect the hip flexibility measurement such as postoperative pain, muscle guarding, and apprehension. Moreover, age was found to affect the FABER test measurements based on the MLM results. The researcher did not come across other hip arthroscopy studies that considered the FABER test distance as a flexibility test in HA patients. Thus, this study is unique because it included the FABER test distance as

an outcome measure for hip arthroscopy patients and the data of this test were analysed over a period of 24 weeks postoperatively.

From a pure statistical point of view, the changes in FABER test distance contradict the insignificant changes in the ROM. Therefore, it is necessary to explain the differences between the two tests. The FABER test is performed in 30° of flexion, 30° of abduction and slight external rotation, which is close to the joint's loose-packed position. Placing the joint in or close to this position puts the least amount of stretch on the joint ligaments (Levangie, et al., 2011). Moreover, the FABER test can be affected by the status of the sacroiliac joint (Kokmeyer, et al., 2002). On the other hand, assessing HER ROM in a seated position is limited by the superior capsule and the iliofemoral ligament, while HIR ROM is constrained by the inferior capsule and the superior ischiofemoral ligament. Moreover, placing the joint in a flexed position while testing can change the length of many hip muscles and restrict the ROM (Simoneau, et al., 1998). The effect of position change on the muscle length, which affects the ROM in sitting, is a debatable topic. However, the variation in the testing and hip positions could explain the discrepancy in the findings between ROM and FABER test distance. The validity and reliability of applying the FABER test distance on FAI patients haven't been established yet and further research is required (Philippon, et al., 2007b).

The outcome of multilevel modelling indicated that age can influence HER ROM, HIR ROM, and the FABER test distance. Older patients aged above 60 years showed reduced HER and HIR ROM together with increased FABER distance on both sides. These results indicate that older adults are stiffer compared to the adult and adolescent patients. The loss of hip and other joints range of motion with ageing is normal because osteoarthritis is prevalent in 30% to 50% of the adults aged 65 years and above (Roach & Miles, 1991; Anderson & Madigan, 2014; Loeser, 2010). Ageing has been found to result in OA indirectly by making the body prone to OA triggering and worsening factors such as joint injury, obesity, and abnormal biomechanics. The wear-and-tear process in

the joint cartilage from ageing results in chondropathy, joint surface osteogenesis* and the development of bony spurs. When these factors are combined, they increase the likelihood of developing OA and reducing the range of motion (Loeser, 2010). Without a doubt, the presence of OA, as explored during radiographs or the arthroscopy procedure, can negatively affect the postoperative outcome (McCarthy & McMillan, 2013). In fact, patients presenting with severe OA often have lower survivorship rates resulting in Total Hip Arthroplasty (THA) procedures within a few years from the arthroscopy (Table 2.7) (Lindner, et al., 2013).

In addition to age, sex was also found to affect the ROM of the patients involved in the current study. Based on the findings of MLM, the female patients demonstrated increased HIR and reduced HER ROM compared to the male patients. Anatomically, females are characterised with a wider femoral anteversion† angle (14.7° ± 6.1°) compared to males (8.6° ± 4.7°) (Nguyen, et al., 2009). Moreover, increased femoral anteversion angle has been linked with decreased external rotation (r= -0.208; p =0.027) and increased internal rotation (r= 0.231; p =0.002) range of motion (Ejnisman, et al., 2013).

Hip arthroscopy results in a significant improvement in the symptoms associated with intra-articular impingement. Clinical improvement is linked with increased HFL and HAD but not HER and HIR ROM (Freke, et al., 2016). It is important to note that the age of the patients included in this study ranged from 15 to 75 years. Therefore, the changes in ROM of young patients may have been affected by the results of the older patients because the average scores of both groups were reported together. Excluding the patients older than 60 years may show different results. Based on this, the researcher does not recommend considering HER and HIR ROM only as markers for clinical improvement following hip arthroscopy. The range of motion tests should include other

* Production or growth of bone tissue.
† Rotation of the femoral neck in relation to a mediolateral axis passing through the femoral condyles.

movements such as HFL and HAB. Also, postoperative reassessment should be based functional tasks that reflect a true change in the patients' status such as performing deep squat exercise and sports specific tasks such as pivoting and twisting. The findings of these tests should be compared with baseline or pre-operative measurements.

Key Points

- Hip ROM did not show statistically significant difference following the surgery.
- Older age negatively affected the ROM and FABER test distance.
- The female patients demonstrated reduced HER and increased HIR ROM.
- The mixed pathology of patients could have affected the ROM findings. However, further analysis cannot be achieved at this stage.

3.7.3. Function and Pain

Analysis of the dataset in the current study also showed that pain and function significantly improved in all postoperative time points. These changes are considered clinically relevant because they meet the minimal clinically important difference (MCID) of 10 points for the MHHS and 1.37 points/cm change for the VAS (Hawker, et al., 2011a; Nielsen, et al., 2014).

The measured VAS pain scores of the current study were based on the level of pain during activity. Excision of the bone tissue overgrowth, the damaged labrum and the loose bodies can have a huge impact on reducing intra-articular pain (Lynch, et al., 2013). In the current study, there was a marked reduction in pain VAS scores from 7.14 ± 2.16 cm preoperatively to 1.70 ± 1.93 cm (p <0.001), which was also paired with statistically significant increases in muscle strength and function (Appendix B. 12). The pain VAS scores changes observed in this study are supported by the findings of previous research. Dippmann et al. (2014) found that pain, as measured using the VAS,

showed a statistically significant improvement from 5.79 ± 2.42 cm preoperatively to 3.00 ± 2.42 cm at three months (p <0.001), then to 2.26 ± 2.42 cm at six months (p =0.017). However, the pain did not show any significant improvement between 6 and 12 months (2.26 ± 2.42 cm to 2.79 ± 2.42 cm, p >0.05). Brunner et al. (2009) followed up with patients for an average of 2.4 years and also found a significant reduction in pain scores 5.7 (range: 1-9) to 1.5 (range: 0-6) points at follow-up using the numeric rating scale (NRS) (p <0.001). In addition, Nielsen et al. (2014) discovered a statistically significant drop in pain using the NRS from 5.0 ± 2.7 points to 3.7 ± 3.9 points (p <0.001) after 2 to 5 years from the arthroscopy. When the mean and standard deviation results of these studies and the current study were closely examined at the last follow-up, it is obvious that some of the patients had pain scores equal to or close to zero. Therefore, these observations support the use of hip arthroscopy in reducing intra-articular pain.

The MHHS is a commonly administered questionnaire for hip arthroscopy patients, and it was used as one of the main outcome measures to monitor the progression of the patients included in the present study. However, the MHHS scores were not obtained in the second week because physical activity was not allowed in the first few weeks (Enseki, et al., 2010). Thus, it was unable to measure the change in their functional abilities. The current study results show a clinically* and statistically significant increase in the MHHS scores from 65.41 ± 13.83 points preoperatively to 84.78 ± 13.01 points at the 24th week postoperatively (p <0.001). Dippmann et al. (2014) also found a clinically significant elevation in the MHHS score from 59.9 ± 1.9 points preoperatively to 74.5 ± 1.9 points at three months (p <0.001) and then to 80.1 ± 1.9 points at six months (p =0.004).

Hip arthroscopy patients were also followed for longer periods to monitor the change in function. Unfortunately, the dataset of the current study contained a high percentage of missing values in the 52nd week (one year) postoperatively. Therefore, analysing the

* Because the change in scores was more than 10 points, which is the minimal clinically important difference.

patients' functional progress beyond the 24th week was not possible. However, Kamath et al. (2009) found a clinically significant improvement in the MHHS scores from 56.8 preoperatively to 80.4 points at 2-years postoperatively (p <0.001) (further information is not available). Polesello et al. (2014) also found that preoperative MHHS scores showed a marked clinical change from 62.4 ± 13.4 points to 90.6 ± 11.6 points after 5-8 years from the HA (p <0.001) (Table 2.5). The short to medium-term clinically relevant improvement in the MHHS scores observed in the current study and the long-term changes observed in previous research strengthen the fact that the patient's function can significantly improve for a long period after hip arthroscopy.

Multilevel modelling indicated that sex and athletic activity have an impact on the measurements of function. The females scored less than the males, and the professional players scored less than the amateur players and inactive patients. In a retrospective case series study, Kamath et al. (2009) found that being a female, having a left sided procedure, and the preoperative activity level were positive predictors of good outcome. In contrast, Redmond et al. (2015b) noticed that being a male, increased BMI, the presence of back pain, and other factors were detected in the patients who had THA after the arthroscopy procedure (Table 2.7).

The drop in the functional level of the professional athletes may be due to their perception of physical activity. They may portray the level of physical activity in another way when compared to amateur players and inactive individuals. Therefore, they consider the lack of high-level activity, reduction in muscle strength and range of motion following the arthroscopy as a major drop in their functional abilities. On the other hand, giving a score of a 100 on the MHHS or other questionnaires may not necessarily indicate a full return to play and optimum functional levels. This ceiling effect impacts the analysis and interpretation of the results (Schilders, et al., 2011). The ceiling effect of the MHHS, subscales of HAGOS and other questionnaires may limit their clinical use with the athletes (Kemp, et al., 2013). Based on this discussion, the author recommends that a new and comprehensive patient-reported outcome questionnaire with a high

measurement ceiling should be developed to measure the changes in the functional abilities of professional athletes.

The results of data analysis imply that the duration of symptoms did not have a significant effect on explaining the variation of postoperative measurements. It was reported that the presence of symptoms longer than a year led to worse prognosis* (McCarthy & McMillan, 2013). However, Kamath et al. (2009) found that symptoms duration of >18 months was a good predictor of better postoperative MHHS scores (OR =1.667). In contrast, the results of another research suggest that symptoms duration of <18 months was associated with the patients who did better postoperatively (Byrd & Jones, 2010). In another study, the symptoms duration was negatively correlated with the iHOT-12 (r = -0.189, p =0.012) and the HAGOS quality of life subscale (r = -0.209, p =0.004) (Table 2.7) (Sansone, et al., 2016b). The conflicting conclusions indicate the need for further analysis of the effect of symptoms duration on the outcome.

Other researchers looked at the effects of symptoms duration on returning to play. Philippon et al. (2010) stated that the time from the injury to the surgery positively correlated with the time to return to play. Furthermore, the Hockey players who had symptoms for less than a year returned to sports within three months, whereas those who waited for more than a year to have the surgery returned to sports at an average of 4.1 months later (Philippon, et al., 2010). Similar results were reported in another study in which there was a significant difference between those with short periods (<12 months) and extended periods of symptoms (>12 months) (p <0.05). A higher percentage of return was recorded in the athletes with shorter periods compared to the other group (Table 2.7) (Sansone, et al., 2015).

The findings of the present study oppose the main findings reported by Philippon et al. (2010), Sansone et al. (2016b), and Sansone et al. (2015). The results of multilevel modelling, which is a more powerful statistical method to investigate the effect of factors

* Further information are not accessible.

on measurements, showed that the duration of symptoms had no statistically significant effect. However, the patients who waited more than two years before having the arthroscopy showed a reduction in the means of strength, range of motion and joint mobility with higher pain values. Even though, the results were accompanied with a large standard error. Accordingly, the findings of the effect of the duration of symptoms should be considered with caution.

Key Points

- MHHS and VAS scores showed significant improvements following the arthroscopy.
- Females had lower MHHS and higher VAS scores compared to the males and the professional athletes had slightly lower MHHS scores compared to the other two groups.
- The duration of symptoms did not affect postoperative outcome.

3.8. ***Study Limitations***

The current study has a number of limitations in the following domains:

- *Data collection and missing values:* The dataset contained a lot of missing values, especially at the 52nd week. Most of the patients failed to attend their last check-up at the 52nd week (1 year) from the surgery, which was previously reported in hip arthroscopy and arthroplasty patients (Clohisy, et al., 2008; McCarthy, et al., 2011). It is possible that the patients did not find it feasible nor necessary to attend the last assessment due to near full recovery and function. Excluding the measurements of the 52nd week negatively impacted the conclusions drawn from the dataset and limited the advantage of this extended period of follow-up.

- *Data extraction:* The data were extracted from the SOCRATES software. Even though the procedure was standardised to maximise the amount of information collected from the patients' files, data extraction was performed manually. Therefore, copying the data into the excel sheet could have been affected by human error resulting in a wrong or missing values.
- *Intra- and inter-operator variation and standardisation of testing:*
    - First, the measurements were taken by three clinicians (a surgeon and two physiotherapists) using the same methodological approach. However, some of the measurements such as muscle strength and ROM could have varied across the testers and timepoints.
    - Second, the exact location of the measurement tool such as the MicroFET2 or the goniometer may not have been consistent across the testing sessions, which could have affected the measurements (Krause, et al., 2007). However, the extent of measurements variability cannot be determined due to the absence of information regarding the tester's level of experience, sex, strength and accuracy in testing (Krause, et al., 2014).
    - Third, the procedures were performed by the same surgeon and the technique may not have varied greatly (McCarthy, et al., 2011).
- *Inter-patient variation and sample characteristics*: The dataset consists of a heterogeneous sample of patients with varying age groups, duration of symptoms and physical activity levels.
    - First, the hip arthroscopy procedure may have been performed on a wide range of patients complaining from various pathologies. In addition to patients with FAI and labral tears, the sample may have contained hip dysplasia patients, who are clinically characterised with hypermobility and do not necessarily show noticeable changes in the range of motion post-arthroscopy (Sturridge & Bankes, 2015). Also, some of the patients

may have been diagnosed with joint laxity or other uncommon pathologies (Bogunovic, et al., 2013). Nonetheless, all procedures were performed to resolve the source of pain and restore function (Daivajna, et al., 2015).

- Second, the information about the approach to treat labral pathology was missing. Since labral fixation patients show comparatively different results to labral resection patients, the impact of the procedure type on postoperative recovery could not be determined at this stage (Schilders, et al., 2011).
- Third, the patients' demographics such as the body mass and stature of was missing, which could have been helpful to normalise the values and reduce inter-patient variation (Keating & Matyas, 1996; Sisto & Dyson-Hudson, 2007).
- Fourth, there was an overrepresentation of adult (83%) and inactive (51%) patients (Table 3.3). Inactivity could have affected the postoperative muscle strength since physically active patients recover quicker than the sedentary patients (Kamath, et al., 2009; McCarthy, et al., 2011).

- *Outcome measures*: The MHHS was used as one of the primary outcome measures in this study. However, it is not a validated tool for hip arthroscopy patients and does not include a subscale for sports (Hinman, et al., 2014; Tran, et al., 2013). Moreover, a ceiling effect in measuring the changes in physical activity and function is noticed when the questionnaire was filled by athletes. Thus, it may not have been the best questionnaire for this study because the sample contained high-level athletes. Other questionnaires such the HAGOS, HOOS and HOS, which contain a subscale for sport, should have been applied along with the MHHS (Hinman, et al., 2014).
- *Interpretation of the data and study design*:

- First, applying the ANOVA tests in analysing the repeated measures may have weakened the power to detect statistically significant changes in the ROM. The sample size reduced to as low as 51 when the ROM was analysed because ANOVA tests follow a listwise deletion method (Field, 2013).
- Second, the researcher was informed that the data contained a heterogeneous sample of hip arthroscopy patients with mixed pathologies, but the number of cases of each pathology was unknown.
- Third, the retrospective analysis of the dataset may have been affected by selection bias, lack of a comparative control group and radiographic data (Tran, et al., 2013).
- Fourth, the dataset consisted of fixed pre- and postoperative measurement timepoints and it is unknown if other clinical changes have occurred in-between (Dippmann, et al., 2014).

3.9. *Conclusion*

Hip arthroscopy is a widely applied procedure to treat intra-articular pathologies. The analysis of the secondary data in the current study showed that hip muscle strength, function and pain could take at least 8 weeks to show statistically significant improvements. This finding was not clearly defined previously. Other researchers indicated that hip muscle weakness could persist for more than 2.5 years post hip arthroscopy. The exact relation and timeline of events between intra-articular hip pain and preoperative muscle weakness are difficult to establish at the moment.

Combining the findings of MLM in the current study show that the female, amateur, and inactive patients are weaker than the male patients, and high-level athletes. The strength differences between inactive patients and high level athletes were not surprising. These strength differences should be accommodated in the postoperative

rehabilitation programs. The rehabilitation protocols targeted to treat female patients should focus on increasing HAB and HER muscle strength to minimise HIR and knee valgus during movement. On the other hand, the rehabilitation programs of athletes should focus on high volume strength exercises since they start with stronger muscles compared to the other patients.

HER and HIR did not show statistically significant changes after the arthroscopy. It could be due to the drop in the sample size with repeated measures ANOVA, mixed pathology and sample characteristics. The ROM of older adults is more restricted in comparison to the younger patients. Arthritic changes with ageing play a significant role in limiting the range of motion in patients aged more than sixty years old. Regarding sex, female patients had a lower external rotation and higher internal rotation ROM compared to the male patients. The variation in ROM is due to the increased femoral anteversion angle found in females.

Pain and function as measured by the VAS and the MHHS respectively show significant improvement after the procedure. Professional athletes showed reduced function scores compared to the other patients. However, such athletes may perceive the reduction in function differently. Nonetheless, the reduced number of professional athletes compared to amateur players and inactive patients may have skewed the results. This issue could be avoided if the results were reported separately. More research should be performed in this area.

The modified Harris Hip score may not be the best outcome measure to test physical function and readiness to playing sports since most professional athletes can score very high. In other words, having a score greater than 90% does not reflect their ability to return to high-impact competitive sport. A new tool should be designed and validated to measure basic sports requirements such as sprint speed, endurance, functional range of motion and lower limbs strength using sport specific tests such as agility tests, pivoting and quick change in direction.

The findings of the current study add to the existing body of knowledge in multiple ways. First of all, it is the first study to include the measurements of muscle strength, range of motion, FABER test distance, function and pain over a period of 24 weeks. Second, the results can be more generalised because it includes patients in different age groups, levels of athletic activity and duration of symptoms. Third, it is the first study to compare the results of having hip arthroscopy on the right and left sides. Other studies did not consider this factor in the analysis. Lastly, using multilevel modelling to demonstrate that age, sex and level of athletic activity do affect postoperative measurements supports the use of this test in similar studies.

The analysis of this set of data can be performed with an alternative approach. In the current study, the patients were divided and compared based on the operation side. Alternatively, the patients with similar level of athletic activity or age range can be grouped together and compared to get a wider view of how hip arthroscopy can influence the strength, the range of motion, and function of these groups. To simplify it, the comparisons in ROM can be based on the age group, in which the measurements of the younger patients are compared against the measurements of the older patients. Moreover, muscle strength could be considered a parameter of postoperative recovery. In other words, benchmarks should be set for muscle strength in a set time points such as a percentage of strength or number of repetitions compared to preoperative measurements or the contralateral side. Such criteria have to be specific to the age group, the level of athletic activity and sex of the patient.

Muscle strength testing of the patients included in this study was performed using a manually stabilised hand-held dynamometer. Although the application of hand-held dynamometry testing in the clinic is quick and provides essential information, the reliability and objectivity of testing could be improved. Therefore, two stabilisation platforms were designed and discussed in the next chapter of the thesis (Figure 1.1).

Chapter 4 Designing Process of the Assessment Platforms

4.1. *Chapter Overview*

Hip arthroscopy, physiotherapy and other interventions may result in significant changes to muscle strength. The results of the first study showed that the hip abductor and adductor muscle strength can significantly change following hip arthroscopy. Therefore, a valid and reliable method of hip muscle strength testing, which can be quickly administered has become a necessity. This is one of the main reasons behind introducing two hip muscle strength assessment platforms that were designed by the author to improve the reliability of strength testing using a hand-held dynamometer. The design features and steps based on the concepts of testing will be briefly described in this chapter. It is out of the scope of this thesis to discuss them in detail.

4.2. *Significance of the Platforms*

Now, manual muscle testing (MMT) is a widely applied clinical assessment tool to evaluate muscle strength based on a five main grading scores. However, MMT is criticised by many researchers as having a high degree of subjectivity and variation in scores for up to 25% because the test relies on the estimation of force applied by the examiner (Bohannon, 2002; Frese, et al., 1987; Wadsworth, et al., 1987). In the recent years, clinicians adapted the use of various types of hand-held dynamometers (HHD), which are valid and reliable in testing muscle strength (Kawaguchi & Babcock, 2010; Lu, et al., 2011; Thorborg, et al., 2010a). Yet, the use of HHDs in testing can be affected by the location of application, the examiner's strength, and method of stabilisation (Katoh & Yamasaki, 2009; Krause, et al., 2014). Isokinetic devices are also used in muscle strength testing, but they are expensive to purchase and impractical (Drouin, et al., 2004; Zapparoli & Riberto, 2016).

The use of hand-held dynamometers stabilisation platforms can overcome the issues related to the manual application of HHDs in measuring hip muscle strength before and after hip arthroscopy. Therefore, the Groin Abductor-Adductor Strength Platform (GrAASP) was designed to test the strength of the hip abductor and adductor muscle groups in the frontal plane. After phase one of testing in the second study was performed, the design was modified to include more muscle groups. The modified platform was named the Hip Strength Assessment Platform (STAPH), which allows testing of the hip abductor, adductor, flexor, extensor, internal and external rotator and the hamstring muscle groups. In this chapter, the designing process of the GrAASP, STAPH and the stabilising tools will be briefly discussed as it is not the focus of this thesis.

The two platforms were introduced to improve the objectivity of screening using the dynamometer and to evaluate whether the results are more consistent and have less testing error or not. Also, the platforms can be assembled and utilised in sports, physiotherapy, and orthopaedic clinics to test the strength of various hip muscle groups and bridge the gap between HHDs and isokinetic devices. The significance of these two platforms is further discussed in Chapter 7.

4.3. *Idea Development and Stakeholders*

The idea of developing a frame or a board to test hip and groin muscle strength objectively was introduced in February 2014. The previously mentioned platforms in section 2.7 were reviewed and critiqued.

The process of designing, manufacturing, and testing the board involved several related parties. The researcher (MH), who introduced the idea, supervised the designing process and performed the necessary validation and reliability studies. The supervisors, Professors Cooke and Schilders, provided guidance throughout the design and study phases together with redirection towards the potential sources of error. Dr Claire Griffiths

provided valuable input in the statistical methods that were used to compare the results of testing using the platforms against a gold standard. Equal Adventure Ltd, represented by Dr Suresh Paul*, provided the designing and manufacturing teams to draw and assemble the 3D models.

4.4. *Development of the Platforms*

4.4.1. Overview

The designing process of the platform and dynamometer support was founded on three concepts. The first concept is based on using a frame for stabilisation; therefore, it is referred to as the frame in section 4.4.2. It was followed by the second concept, in which the frame was replaced by a flat board and a dynamometer support system. When the support system is attached to the board, it is referred to as the platform. The idea behind the third concept is to redesign the platform to test more muscle groups.

4.4.2. Concept 1

Version P1.0

In the first concept, a metal frame made from aluminium fixes to a standard wooden physiotherapy treatment bed. The dimensions of the frame can be modified to accommodate different body sizes and types of treatment beds (Figure 4.1). The dynamometer is attached to vertical rods, which provide the required stabilisation. The central rod supports the

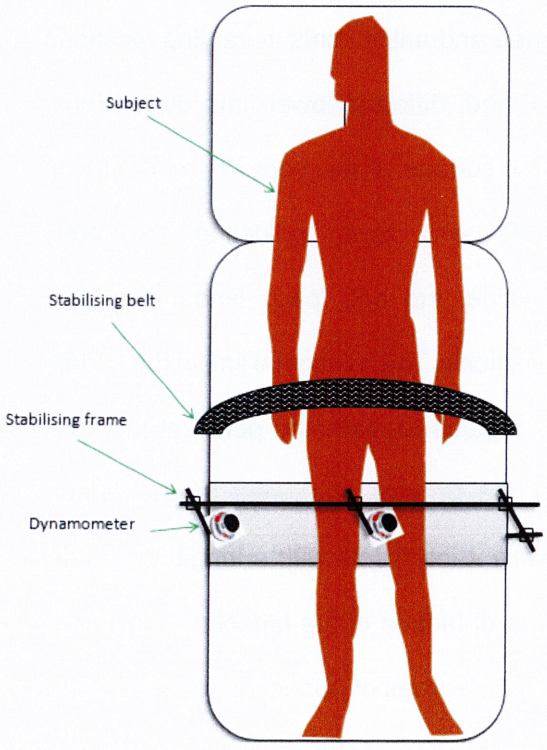


Figure 4.1 – Drawing of the board version P1.0 (top view)

* PhD in Inclusive Adventure by Design and founder of Equal Adventure

dynamometer in hip adduction (HAD) strength assessment, while the side rods provide an attachment point for the dynamometer in hip abduction (HAB) strength assessment. However, attaching the metal frame to the table in a complex and non-convenient manner together with its weight and size may hinder the practicality of this approach.

4.4.3. Concept 2

Version P2.0

The metal frame was replaced with a flat board to allow easy application and portability. The participant lies in a supine position and pushes against the dynamometer, which slides in a groove instead of being fixed in certain locations. The dynamometer is placed near the knee and ankle joints in varying locations to meet different lower limb dimensions. The surface of the board is marked with lines and axes indicating the body midline, hip and lower limbs alignment positions. The horizontal line in the middle is a guide to place the pelvis at the level of the greater trochanter of the femur, while it intersects with a longitudinal line at a right angle to form the anterior-posterior axis of motion of the hip joint (Figure 4.2). The starting position of the lower limb is set at 10 degrees of abduction from the midline for consistent results. The decision to place the hip in 10° of abduction is based on the research performed by Neumann et al. (1987), in which this specific angle showed the highest torque values for the HAB muscles.

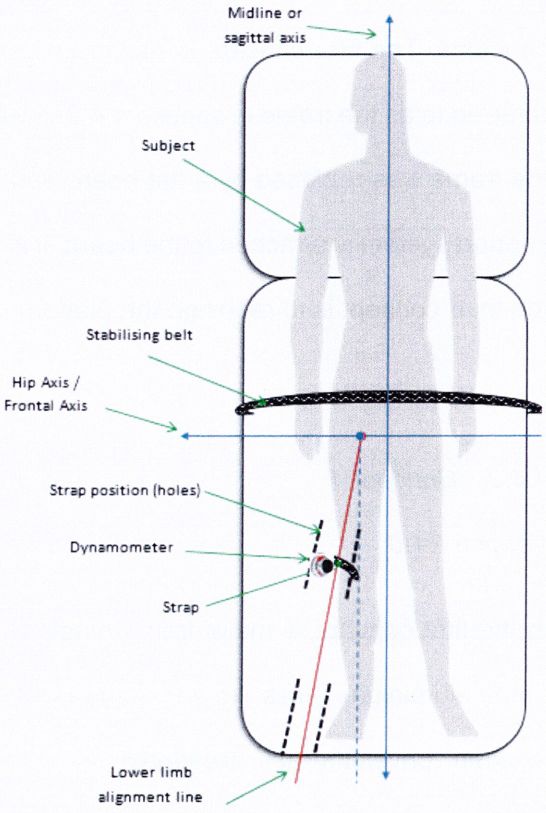


Figure 4.2 – Drawing of the board version P2.0 (top view).

The design was drawn using 3D *SolidWorks* computer-aided design software (Dassault Systèmes, Vélizy, France). The concept is based on testing both hips without repositioning. However, after asking several volunteers of different body sizes to participate in a quick test, it was decided to change the design because it does not allow bilateral hip testing.

Version P2.1

The principle of this modified version is that each side of the board is used to assess the right or left hips separately. Volunteers were asked to sit on the cardboard model to simulate testing and check if the design was suitable for practical application (Figure 4.3). It was noticed that the ischial tuberosity of the opposite side is positioned outside the board, which may be uncomfortable for some individuals.

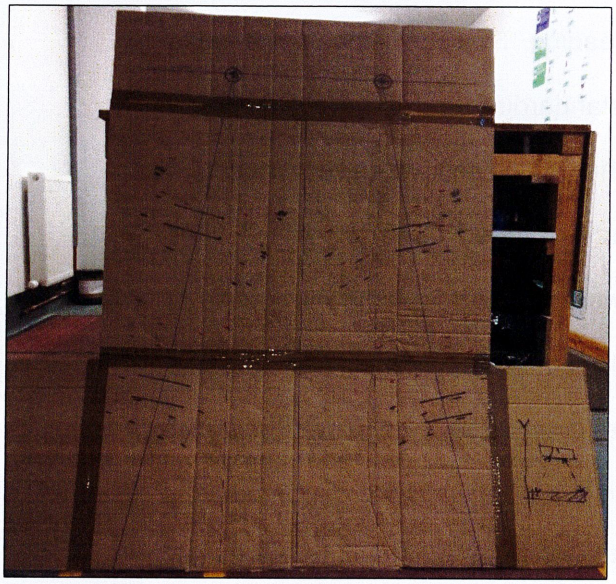


Figure 4.3 – Cardboard model of version P2.1.

Version P2.2

In this design, extra space was added to accommodate the ischial tuberosity of the non-involved side. The groove was shifted to the centre of the board approximately 35 cm away from the left edge. Unnecessary areas were trimmed and cut to reduce the size (Figure 4.4).

Version P2.3

Extra spaces were cut, and the dimensions were amended. A reference measurement line in centimetres was added to the hip line to improve the objectivity of testing. Another parallel line was added close to the top edge for easier reading. These two lines are used to standardise the starting position of the participants when reassessment is performed. Also, reference lines were drawn parallel to the groove in 10 cm increments to measure the lever arm from the greater trochanter to the position of the dynamometer. These lines run for 100 cm along both sides of the groove and are based on hip height normative values (*M* ± *SD)* (50th% 92 ± 5.0 cm) (Pheasant and Haslegrave, 2006).

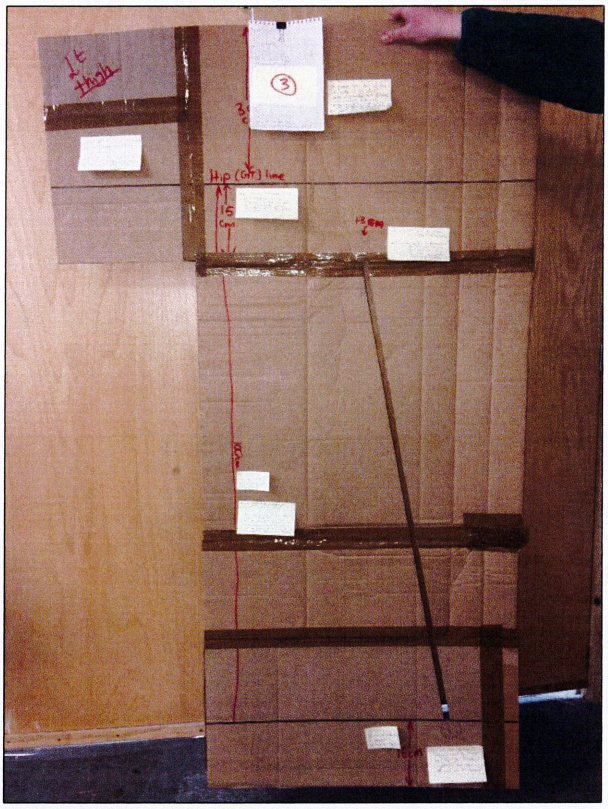


Figure 4.4 - Cardboard model of version P2.2.

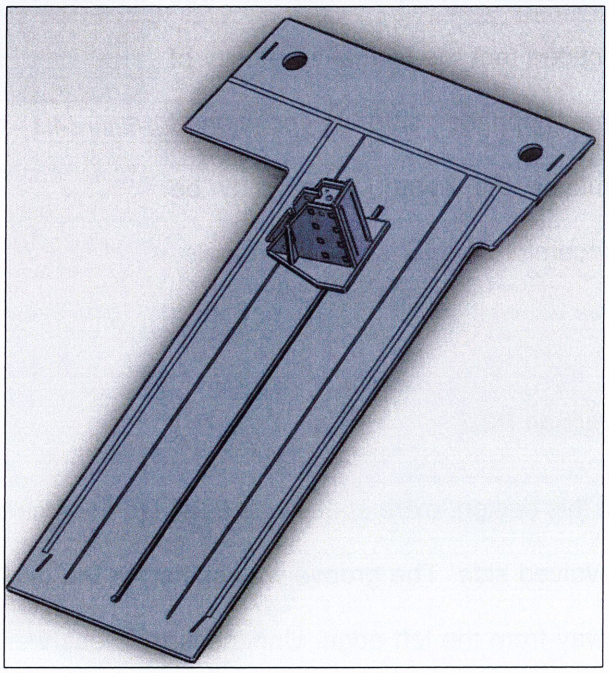


Figure 4.5 - 3D SolidWorks drawing of version 2.3.

Since the initial idea was to place the board on a physiotherapy bed for assessment purposes, two straps were added to the board to fix it securely against the bed. Regarding the groove, the ends were cut in the shape of a circle measuring ∅ 1.3 cm, which allows easier entry and exit of the dynamometer support. For storage purposes, holes were

made at the top third to hang the board on the wall when not in use. The final design was made in 3D using SolidWorks (Figure 4.5).

Before moving forward with the manufacturing phase, a small pilot study was required to determine the potential sources of error in measurements. Based on the results of testing, it was decided to use the board on the floor instead of the bed, place the participants in a supine position rather than long sitting, and to make a few changes to the board design and dimensions. The study details and results can be found in Appendix C. 1 to C. 4.

Version P2.4 (The GrAASP)

This version is designed so that the hip of interest is aligned with a horizontal reference line. It is followed by repositioning the lower limb and sliding the dynamometer in a groove at 10° of abduction* (Figure 4.6). The dimensions of the board and reference lines were based on *BodySpace* by Pheasant and Haslegrave (2006). The fiftieth percentile (50th %) reference values of the British adult males are considered because they are higher than the values of females.



Figure 4.6 – The (GrAASP) showing the stabilising straps, head and back sponge support and the dynamometer support.

The hip line, on which the greater trochanter will be aligned, was positioned 30 cm away from the top edge based on the abdominal depth† reference values (27.0 ± 3.2 cm) (Pheasant & Haslegrave, 2006). Fifteen

* Will be explained in the Development of the Dynamometer Support section
† Definition: The maximum horizontal distance between a virtual vertical line at the back to front of the abdomen in sitting position

centimetres were considered between the hip line and the starting point of the groove to avoid placing the dynamometer very proximal to the groin area. The starting point was positioned 20 cm away from the left edge and was 100 cm long based on the normative values of hip height* (50th% 92.0 ± 5.0 cm; 95th% 100.0 ± 5.0 cm) (Pheasant & Haslegrave, 2006). A distance of 16 cm was considered between the end of the groove and the bottom edge to avoid breakage while testing. Two straps were added in the middle third to stabilise the non-tested side at knee level (Figure 4.7). The final dimensions of this model were H 166 x W 80 cm.

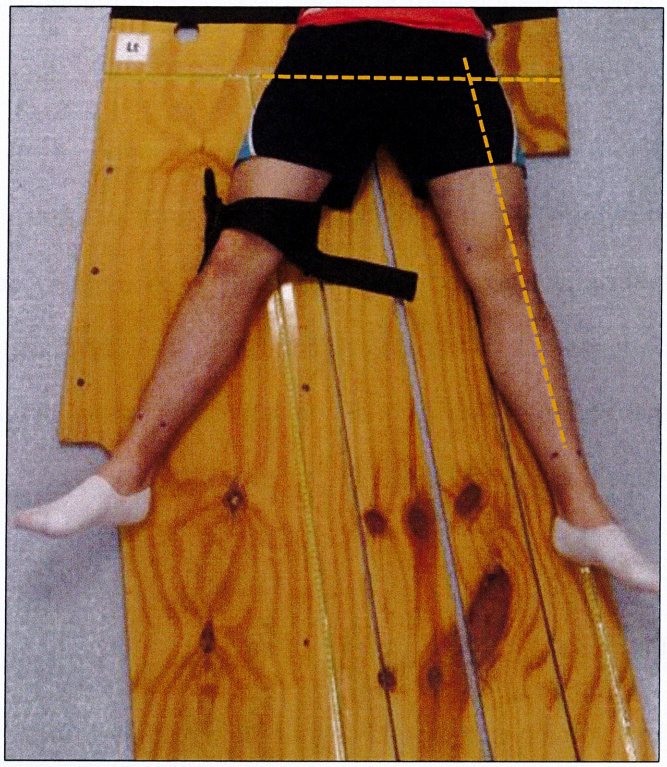


Figure 4.7 – A participant positioned to have his left hip tested. The left hip is aligned over the intersections between the horizontal and hip lines.

Extra space was added on the left side of the board to accommodate the stabilising straps. The board was later manufactured of wood with the dynamometer support grooves and holes for the straps. It was appropriately marked with the side and muscle group. Moreover, a measurement tape was added on the hip horizontal line to act as guidance. Another tape was added along the grooves to be used as a reference for the location of the dynamometer. The board was named the Groin Abductor-Adductor Strength Platform (GrAASP). The validity and reliability of GrAASP will be discussed in Chapter 7.

* Definition: The maxim vertical distance from the floor to the greater trochanter in standing position

4.4.4. Concept 3

Version P3.0

After performing the validity and reliability study of the GrAASP, several issues were raised regarding the design of the platform, the dynamometer support, and the attachment. A modified version of GrAASP was introduced to resolve the problems related to the GrAASP. The new model, which is based on an entirely new concept, can be adjusted to test the hip abductor (HAB) and adductor (HAD) muscles from a supine position. The hip flexor (HFL) and extensor (HEX) muscles can be tested from a side-lying position. The platform also allows testing the hip internal rotator (HIR), the external rotator (HER) and the hamstring (HAM) muscles from a sitting position. The new design was renamed into the Hip Strength Assessment Platform (STAPH) (Figure 4.8).



Figure 4.8 - The hip strength assessment platform (STAPH) – version P3.0.

The STAPH is made of two standard wooden boards with legs (54 x 54 cm). The two boards attach to a third board with stainless steel extensions, which are supported on the floor by short legs. Aligning the boards side to side with each other allows testing the HAB, HAD, HFL, and HEX. Once the participants are positioned appropriately for testing, the body weight prevents the three wooden boards from moving. They can also be stacked on each other to test the HIR, HER and HAM using the long legs (Figure 4.9). The dynamometer holder or the upright is positioned in appropriate locations on the testing platform and is secured in place using capturing pins.

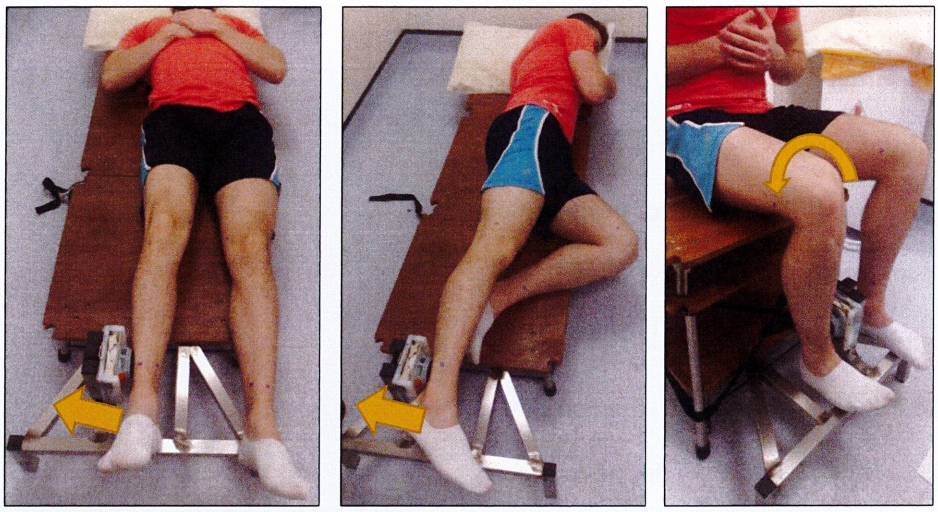


Figure 4.9 - A participant aligned in the supine, side lying and sitting positions to test hip abductors (left), extensors (centre) and internal rotators (right) consecutively (yellow arrows point at the direction of push).

4.5. *Development of the Dynamometer Support*

Using a stabilised dynamometer while testing is recommended. Therefore, the support system was designed to stabilise the dynamometer on the platform during testing.

4.5.1. Concept 1

Version S1.0

The first concept was based on previously published work in which other researchers have used a belt to stabilise the dynamometer (Figure 4.10) (Bohannon, et al., 2011; Katoh, et al., 2011; Kramer, et al., 1991). It was soon excluded due to impracticality and complexity.



Figure 4.10 - Concept 1 of the dynamometer support – version S1.0.

4.5.2. Concept 2

Version S2.0

In this version, the idea was to attach the dynamometer to a wooden support. A bolt runs from the sides of the support through the T-groove of the platform and ends with a nut. The whole system can be tightened against the other side of the board using a revolving handle on the top (Figure 4.11). However, the support became cumbersome and bulky and changes had to be made to the model.

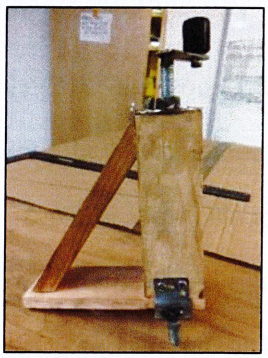


Figure 4.11 – Version S2.0 of the dynamometer support.

Version S2.1

In this model, the wooden parts were replaced by a series of parallel and intersecting planes made of plastic. An anterior adjustable plane was added to the model. The plane slides upward and downward to change the position of the dynamometer in relation to the lower extremity. The dynamometer will be held in position in front of the sliding plane (Figure 4.12). This design was excluded due to potentially increasing the manufacturing costs.

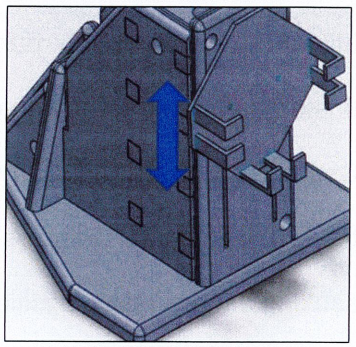


Figure 4.12 – Version S2.1 of the dynamometer support.

4.5.3. Concept 3

Version S3.0

The new support system was made primarily from a metal frame with a plastic box attached to it. Instead of having a triangular back support, two metal rods were welded together at 90 degrees. The sides of the base have been extended to

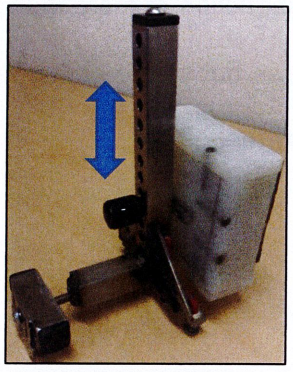


Figure 4.13 – Version S3.0 of the dynamometer support.

form a triangle, while holes were made in the vertical rod to adjust the position of the plastic box. The back contains a screw to fix the support system tightly against the board.



Figure 4.14 - Stabilising the support system during abduction testing.

The position of the support is modifiable using the sliding mechanism and can be securely fixed in place. The vertical position of the plastic box, which contains the dynamometer, can be adjusted to match the height of the limb (Figure 4.13). The dimensions of the plastic box accommodate the MicroFET2 hand-held dynamometer (Hoggan Health Industries, Utah, USA).

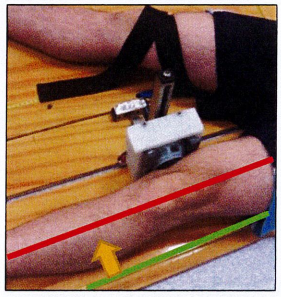


Figure 4.15 - The difference (yellow arrow) between the starting position based on 10° of abduction (green line) and the actual testing position (red line).

After performing the pilot study (Chapter 5), it was observed that the support system did not provide enough stability, and the MicroFET2 was wobbly during testing. The plastic box was unstable and had to be tightened now and then (Figure 4.14). Combined with the instability of the MicroFET2 inside the box, the tested limb was sliding up and down as the participants were seeking a more stable position.

Another issue was faced during testing. The hip had to be taken into some flexion to raise the leg for proper alignment with the MicroFET2. Moreover, a gap was sometimes formed between the device and skin. Occasionally, the limb had to be taken into hip abduction of more than 10° (Figure 4.15). In many circumstances, the support system was not stable enough, and had to be manually stabilised by the tester.

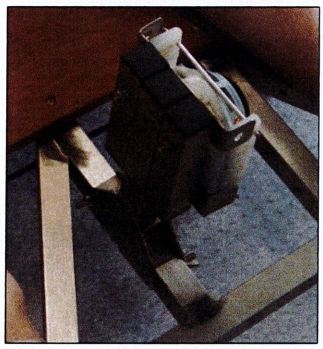


Figure 4.16 - Version S3.1 of the dynamometer support.

Version S3.1

This version of the support was designed and built to be attached to the STAPH (concept 3). The support consists of

an upright with a triangular side support that can be fixed to the STAPH using bolts and wing nuts. The position of the upright can be changed to match the testing muscle (Figure 4.16). Extra holes were made in the frame of the testing platform for this purpose.

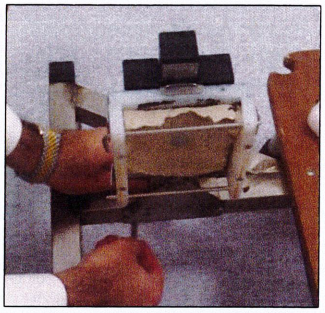


Figure 4.17 - The MicroFET2 support box (version 3.2) showing the mould and the brackets.

The upright attaches to the MicroFET2 support in a precisely measured fitting between two other stainless-steel bars. The space is tight enough to hold the box in place and provide the necessary stability. A clay mould was made to match the round back of the MicroFET2, which is secured in place using two brackets (Figure 4.17).

The designing process of the GrAASP and the STAPH went through a lot of stages and modifications to reach the current versions. A summary of the designing processes can be found in Figure 4.18 and Figure 4.19. Further images can be found in Appendix C. 5 to C. 17.

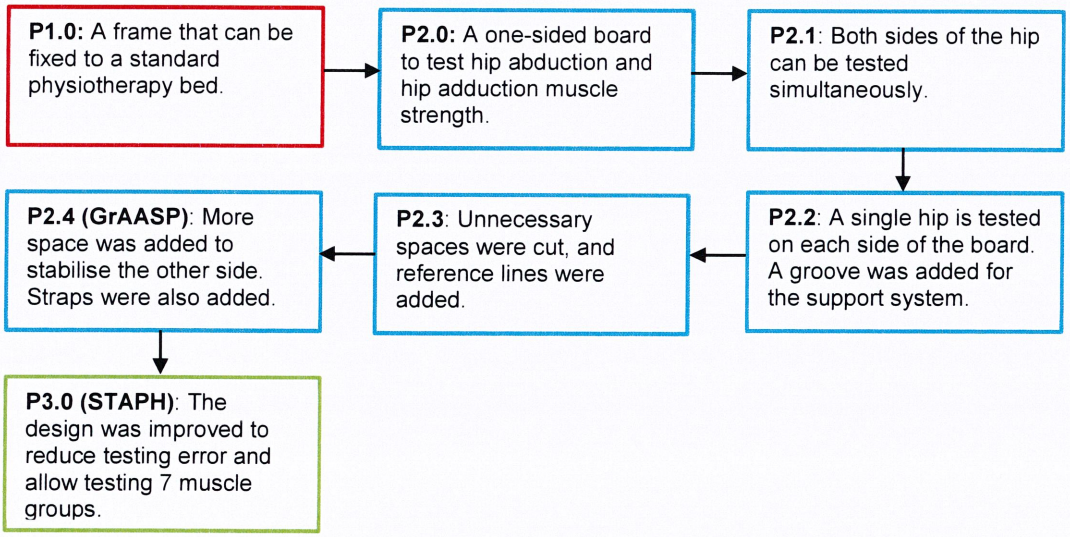


Figure 4.18 – A summary of the platform designs and concepts (GrAASP, groin abductor-adductor strength platform; STAPH, hip strength assessment platform; red box represents concept 1; blue boxes represent concept 2; green box represents concept 3; P, platform).

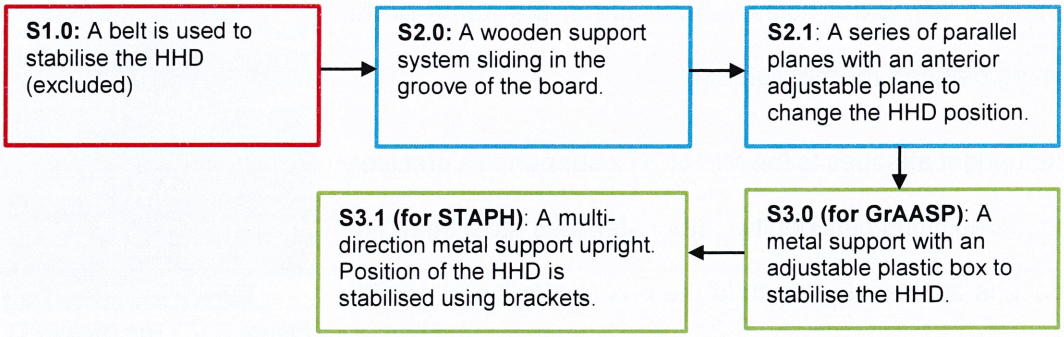


Figure 4.19 – A summary of the stabilising system designs and concepts (GrAASP, groin abductor-adductor strength platform; STAPH, hip strength assessment platform; red box represents concept 1; blue boxes represent concept 2; green boxes represent concept 3; S, support).

Chapter 5 Study Two: Methodology

5.1. *Chapter Overview*

The second study of this thesis is divided into two phases. Phase one of testing was performed to establish the validity and assess the reliability of the Groin Abductor-Adductor Strength Platform (GrAASP) in evaluating the hip muscle strength of healthy active individuals. The relative reliability of using the Hip Strength Assessment Platform (STAPH) in assessing various hip muscle groups was measured in phase two. The study significance, ethical consideration, participation and methodology of these two phases are also explained in this chapter.

5.2. *Research Significance, Contribution and Questions*

5.2.1. Research and Clinical Significance

Enough strength of the lower extremities muscles has been positively linked with the activities of daily living, walking velocity, and the quality of life (Blacklock, et al., 2007; Brown, et al., 1995; Goodpaster, et al., 2006; Rasch, et al., 2010). The patients who experience groin pain, patellofemoral pain syndrome (PFPS), iliotibial band syndrome (ITBS), and anterior cruciate ligament (ACL) injuries have reduced hip muscle strength (Baldon, et al., 2009; Bittencourt, et al., 2016; Taylor-Haas, et al., 2014; Tyler, et al., 2010). Also, the hip muscle strength of hip arthroscopy patients was also found to be reduced pre- and postoperatively (Casartelli, et al., 2011; Nepple, et al., 2015). Therefore, using an objective method such as hand-held dynamometry (HHD) in assessing hip muscle strength is necessary to monitor any alterations that may affect function and results in injury (section 1.2). Plus, the amount and pattern of change should be quantified to aid in planning the postoperative rehabilitation protocols.

Ideally, the muscle strength assessment technique should have sufficient validity and reliability and is less affected by extraneous factors. Using the HHD in testing was found

to be affected by the location and method of stabilisation (Krause, et al., 2014; Krause, et al., 2007). To overcome this issue, several platforms were introduced to enhance the objectivity and reliability of testing. Even though, the uses of the current platforms are limited due to their design, size and complexity (section 2.7).

The present study is a continuation of the first study emphasising on measuring hip muscle strength using a HHD and stabilising platforms on a sample of healthy individuals. In the present study, the GrAASP and STAPH were introduced to improve the objectivity and reliability of hip muscle strength testing of healthy individuals using a HHD. The STAPH, which is the improved version of the GrAASP, is the first platform that can be configured to test seven muscle groups of the hip and thigh.

The results of this study gave valuable information on which of the two stabilisation methods (manual or fixed) and placement locations (proximal or distal) applied in HHD testing are closer to the measurements obtained on a Biodex. Moreover, the results showed which of these methods is more reliable with less testing error. It is the first study to include testing using the Biodex, two methods of stabilisation and location of placement relevant to the hip joint. In addition, the test starting position of the hip abductor and adductor muscles was compared at three different angles and two speeds to determine which position and speed gives more consistent results.

The results of this research add to the existing knowledge around using a stabilised dynamometer in assessing hip muscle strength. The progression of cases and clinical decision making are all based on accurate and thorough assessment procedures. The lack of a precise, practical and a quick measurement test that reflects actual muscle strength may negatively influence the outcome and physical activity resulting in injuries. Therefore, using the platforms in a clinical setting would aid clinicians in obtaining more consistent results in assessment.

The obtained reference values from testing using these platforms will be beneficial. They may set the foundation for future related studies that are based on a similar participants

characteristic. Currently, there is a lack of published values related to hip muscle strength as the current focus is on the quadriceps and hamstring muscle groups. The strength ratios of the agonist and antagonist muscle groups can also be calculated, and the deviation from such values or ratios may indicate the presence of muscular weakness.

5.2.2. Phase One Research Concept, Novel Contribution and Strengths:

- The GrAASP is designed to test the hip abductor (HAB) and hip adductor (HAD) muscle groups using two stabilisation methods on two locations, which is unique to this study.
- It is the first study to include direct comparisons between three angles of abduction in isometric testing and two isokinetic speeds. In this phase, the torque values of the HAB and HAD muscle groups were measured isometrically at three angles of hip abduction (0°, 15° and 30°) and isokinetically at two speeds ($60°.sec^{-1}$ and $120°.sec^{-1}$) using the Biodex (an isokinetic device).
- Muscle strength was also measured by attaching the microFET2 (a hand-held dynamometer) to the GrAASP and by holding it in hand. The collected MicroFET2 data using both methods were compared against those obtained by the Biodex to assess the concurrent validity* of the platform.
- It is the first study to compare the utilisation of a fixed and manual stabilisation methods on a proximal and distal placement sites simultaneously. A proximal (knee level) and distal (ankle level) placement locations of the dynamometer were compared in assessing HAB and HAD muscle strength. Moreover, a fixed (using the GrAASP) and manual (hand-held) stabilisation methods were also compared. The findings may aid in deciding which of the two stabilisation methods and placements are more reliable.

* How well the measurements obtained by an assessment tool agree with another tool at almost the same time (Kolber and Cleland, 2005).

- The relative reliability was evaluated by performing two consecutive MicroFET2 testing sessions.
- The muscle strength ratios between the HAB and HAD groups were also calculated at three angles of hip abduction and two speeds of testing. No previously published research included such data.

5.2.3. *Phase Two Research Concept, Novel Contribution and Strengths:*

- The STAPH was used as the assessment tool of this phase. The new platform can be configured in multiple ways to test the HAB, HAD, hip flexor (HFL), and hip extensor (HEX) muscles groups. Additionally, the hip internal (HIR) and external rotator (HER) muscles can be tested. Lastly, the platform also allows for testing the knee flexor (KFL) muscle group. The STAPH can be classified as the first platform which permits the testing of seven muscle groups.
- For this phase of the study, only the distal location using a fixed method of stabilisation will be considered. This decision is based on the results of phase one of testing (section 6.2).
- The second phase of testing was carried out to assess the relative reliability of using the STAPH in the measurement of muscle strength.
- If the STAPH showed to be a reliable device, it will vastly improve the objectivity of hip muscle strength testing.
- The muscle strength ratios were also calculated in this phase.

5.2.4. *Research Questions (RQ)*

1. What are the hip and knee muscles torque values of a sample of healthy individuals when tested using the Biodex and STAPH? [Phases one and two].

2. What are the torque differences between the dominant and non-dominant sides, and between sexes as measured by the Biodex and the STAPH? [Phases one and two].
3. What are the differences in HAB and HAD isometric muscle torque values at 0°, 15° and 30° of hip abduction as measured by the Biodex? [Phase one].
4. What are the differences in HAB and HAD concentric isokinetic muscle torque values at $60°.sec^{-1}$ and $120°.sec^{-1}$ as measured by the Biodex? [Phase one].
5. What are the ratios of strength between opposing muscle groups in healthy individuals when measured by the Biodex, GrAASP and STAPH? [Phases one and two].
6. Which of the two stabilisation methods of MicroFET2 (fixed and manual) as measured by the GrAASP have the strongest agreement with the measurements obtained by the Biodex? [Phase one].
7. Which of the two MicroFET2 placement locations (proximal and distal) as measured by the GrAASP have the strongest agreement with the measurements obtained by the Biodex? [Phase one].
8. What is the level of agreement between measurements obtained in the two MicroFET2 sessions using the GrAASP and STAPH? [Phases one and two].
9. Which of the two assessment platforms is more reliable in testing hip muscle strength? [Phases one and two].

5.3. ***Ethical Considerations and Participation***

5.3.1. Ethical Approval

The ethical approval to start phases one and two of the study were obtained on the 18th of June, 2015 and on the 25th of May, 2016 respectively. Both applications were signed by the director of studies (DoS) and the Local Research Ethics Coordinator (LREC) from Leeds Beckett University (Appendix D. 1).

5.3.2. Personal Data and Anonymity

Data collection was achieved through two questionnaires and physical assessment of muscle strength, which were considered primary data and not available in the public domain (Appendix D. 2 and Appendix D. 3). Where photographic or videotape recordings were taken, the face and identity of the participants were omitted or covered. This procedure was applied to maintain their anonymity if the material is presented in a journal, lecture or conference.

Once data was stored on a secured, and an encrypted USB stick, hard copies of the data collection forms were shredded. Because the participants agreed to allow the data to be used in other or future research, as mentioned in the consent form, the personal and contact details will be erased upon completion of the PhD and the raw anonymised data will be stored securely for further research (see section 5.3.4).

5.3.3. Human Participants

Participants aged 18 to 45 were recruited using emails and posters containing all the required information. The following criteria was applied in both phases and had to be met to be eligible to participate:

- They should be apparently healthy in general and not complaining from a current lower back, pelvic or lower extremity pain.
- They should not have experienced any recent injury to the hip, groin or knee regions in the past 3 months.
- They should be engaged in any sports activity involving the lower extremities for ≥3 training sessions per week and a total of ≥3 hours per week (Feiring, et al., 1990).

The participants were asked to remain active as normal, avoid extraneous lower limb exercises the day before any session and to maintain their average caloric intake. Any participant who did not meet the minimal inclusion criteria or did not follow the

instructions between the sessions was excluded (Appendix D. 4). These criteria were applied to ensure that the participants were healthy, have strong hip and thigh muscles, and to reduce the variability between them. No at risk, vulnerable or under 16 participants were involved in the study.

5.3.4. Recruitment and Participation

The recruitment of participants was through posters distributed around campus (Appendix D. 5 and Appendix D. 6). Moreover, an email was sent to the staff of the School of Sport and the undergraduate Sport and Exercise Science students asking them to voluntarily participate in the study under the consent of the course leader. A clear information sheet was given to the participants in hard or soft copies if they were interested in participating (Appendix D. 7 and Appendix D. 8). Informed consent was obtained in the first session before proceeding with data collection (Appendix D. 9 and Appendix D. 10). The participants could withdraw from the study at any stage without giving reasons.

5.3.5. Risks and Benefits

There were minimal controlled risks associated with participating in the study. The testing was non-invasive and only required the application of resistance against the MicroFET2 or the Biodex. However, the participants may have felt some soreness in the buttock or thigh regions due to the increased number of completed repetitions. The type of experienced soreness is similar to post-exercise muscular fatigue or ache.

The tester went through appropriate training sessions to mitigate the risks associated with using any of the equipment. The Biodex goes through regular safety check-ups and calibration procedures by the Learning Support Officers. The MicroFET2 is new and has been recently received from the manufacturer in May 2015.

The results of testing were sent to the participants after being briefly analysed and discussed. They were also offered one physiotherapy consultation or advice on a sports related physical complaint within one month by the researcher (MH). He is an experienced physiotherapist with more than three years of clinical experience. He has a Master's degree in Rehabilitation and specialises in musculoskeletal related problems including sports injuries. Detailed risk assessment and safeguards can be found in Appendix D. 11 and Appendix D. 12.

5.4. *Measurement Tools*

Two measurement tools were used to evaluate muscular force and torque. The Biodex System 4 Pro (Biodex Medical Systems, Inc, NY, USA) isokinetic dynamometer was used in phase one only (Figure 5.1). The MicroFET2 hand-held dynamometer (Hoggan Health Industries, Utah, USA) was used in phases one and two (Figure 5.1).



Figure 5.1 – The MicroFET2 hand-held dynamometer (left) and the Biodex System 4 Pro (right).

5.5. *General Testing Protocol*

5.5.1. Screening

1) The interested participants were scheduled as appropriate. After reading the information sheet, they were given the consent form to read, sign and were given the opportunity to ask questions if needed (Appendix D. 9 and Appendix D. 10). Next, they were given the physiology screening questionnaire followed by the dynamometer study specific questionnaire to complete (Appendix D. 2 and Appendix D. 3). Filling these forms determined their eligibility to participate and their risk category. The participants who were classified in the low or moderate risk categories were included because the study does not test maximum cardiorespiratory function (Appendix D. 4).

2) The principle assessor (MH) measured stature (using Seca 225 stadiometer, Germany) by asking the participant to stand on the base (without shoes) with the shoulders, back, pelvis and heels in contact with the upright. The positioning was followed by taking a deep breath, and the head plate was lowered to measure stature in meters (m) to the nearest 0.01 m (Figure 5.2). Body mass was measured by asking the participant to step on an electrical scale, look forward and remain still. It was measured in Kilogrammes (kg) to the nearest 0.1 kg. From these two measurements, the Body Mass Index (BMI) was calculated using the formula:



Figure 5.2 - Stature measurement using a stadiometer.

$$BMI = \frac{Mass\ (kg)}{Stature\ (m^2)}$$

3) Waist circumference was measured by asking the participants to remove any bulky clothes while leaving light clothes on and assume a relaxed standing position with the arms folded across the chest. A retractable, non-

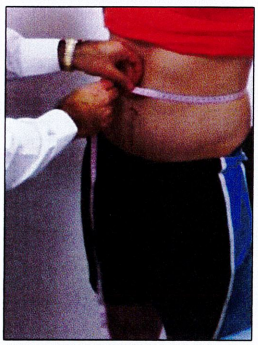


Figure 5.3 - Waist circumference measurement.

extensible and flexible anthropometric tape was passed around the abdomen half way between the lower costal border and top of the iliac crest perpendicularly to the vertical axis of the body (WHO, 2008). The participants were asked to breathe normally, and the measurements were taken at the end of a normal expiration to the nearest 0.5 cm (Figure 5.3).

4) An electronic sphygmomanometer was used to obtain systolic and diastolic blood pressure. The participants were asked to sit on a chair and extend the left elbow on a table. The measurement cuff was wrapped around the arm about 2 cm above the elbow joint. Two fingers were inserted between the cuff and skin to make sure it was not too tight. The waist circumference and blood pressure measurements were taken as part of the physiology screening procedures. The methodology followed in this study was adapted from previously published work and recommendations (Katoh, et al., 2011; Kollock Jr, et al., 2013; Thorborg, et al., 2013b).

5.5.2. Warm-up

The assessor asked the participants to warm up using a stationary cycle ergometer (Resistance: 1.5 kg, 80 RPM, 10 min) or jog outdoors (supervised for 10 min) before any assessment. After that, the hip and thigh muscles were held in a static stretching position for 20 seconds x 2 repetitions. Proper positioning and stretching were insured by MH.

5.5.3. Randomisation and Set-up

The participants and machine set up were according to the operation manual provided by the manufacturer. The starting side was randomised to standardise the effect of fatigue on the measurements and to avoid systemic influence. The starting test side, muscle group, MicroFET2 stabilisation method and placement location were also

randomly assigned by selecting folded papers from a box. The dominant side was determined by asking the participant the following: "Which leg would you use to kick a ball?" or if it was not applicable, "Which is your preferred leg during physical activity?".

5.5.4. *Skin Marking*

- *Phase One:* The important anatomical landmarks were located in a supine position with the hips in 0° of flexion and 0° of rotation and the knees in 0° of flexion. One black mark was placed on each of the following sites for consistency in positioning the MicroFET2: 5 cm proximal to the medial (MFC) and lateral femoral condyles (LFC), and 5 cm proximal to the proximal edge of the medial (MM) and lateral malleoli (LM) (Figure 5.4) (Thorborg, et al., 2013a). The points were selected proximal to anatomical landmarks for proper stabilisation of the dynamometer and to avoid any discomfort during testing. Figure 5.5 shows the medial and lateral points.

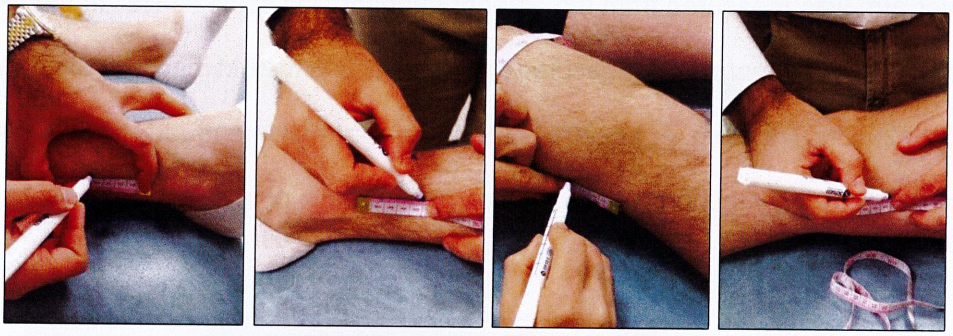


Figure 5.4 – (From left to right) Placing a mark 5 cm proximal to the LM, ML, LFC and MFC.

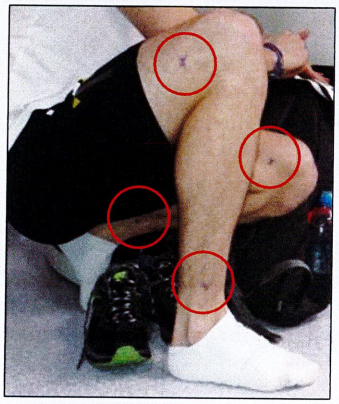


Figure 5.5 - Side view showing the skin landmarks of phase one.

- *Phase Two:* Only the distal location of stabilisation was considered in this phase of testing. A mark was placed 5 cm proximal to the LM (fibula). This point, which was regarded as a reference, was followed by placing parallel marks on the shin (anteriorly), the tibia (medially) and the Achilles tendon (posteriorly) (Figure 5.6).

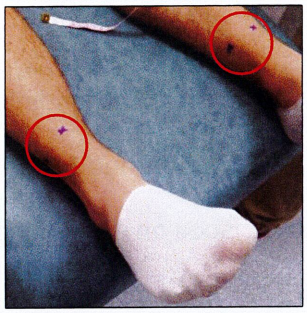


Figure 5.6 - The skin marks used in phase two of testing (blue crosses).

5.5.5. Lever Arm Measurement

The lever arm, which is used in torque calculations, was measured as follows: The short lever arm was measured from the anterior superior iliac spine (ASIS) to the LFC mark, and the long lever arm was measured from the ASIS to the LM mark (Figure 5.7) (Fenter, et al., 2003b; Thorborg, et al., 2010a). Both lever arms were measured in phase one of testing, while only the long lever arm was measured in phase two.

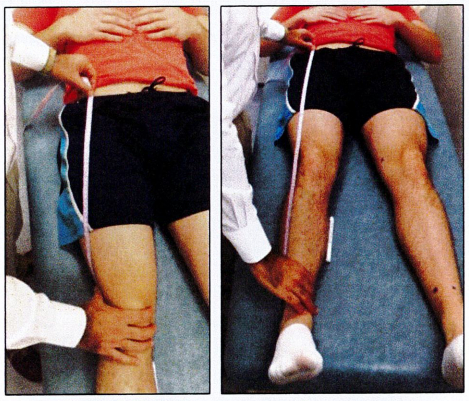


Figure 5.7 - Short lever arm measurement (left) and long lever arm measurement (right).

5.6. *Phase One Testing Protocol*

5.6.1. Biodex Testing

1) The testing procedure was explained to the participant after performing the warm-up and skin marking. The appropriate extensions were attached to the Biodex, and the customised protocol of testing was selected. Next, the participants were appropriately positioned in a standing position, and the lower edge of the attachment was placed at the level of the mark proximal to the LFC (Figure 5.8).

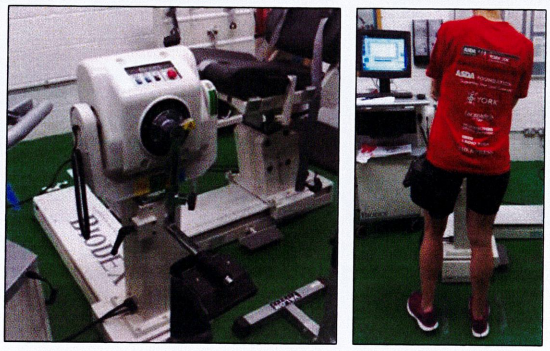


Figure 5.8 –Biodex set-up to test the right hip (left); A participant getting familiar with isometric testing at 0° (right).

2) *Testing:* The participants were allowed 3 submaximal (less than 50% of maximum effort) familiarisation attempts at the beginning of each testing speed. The selected speeds were $0°.sec^{-1}$ (isometric), $60°.sec^{-1}$ (concentric, slow contractile velocity), and $120°.sec^{-1}$ (concentric, intermediate contractile velocity) (Amundsen, et al., 1990). During Maximum Voluntary Contraction (MVC), the participants were verbally instructed and encouraged to perform at their best. The isometric MVC was performed three times for 5 seconds (at $0°.sec^{-1}$ at the angles of 0°, 15° and 30°) for the HAB and HAD muscles. The second set of testing involved concentric MVC testing at $60°.sec^{-1}$ and $120°.sec^{-1}$ for five repetitions each for both muscle groups (Figure 5.9). A rest period of 60 seconds was given between repetitions, and the highest recorded value was used in the

analysis (Katoh, et al., 2011; Kawaguchi & Babcock, 2010). The Biodex testing was performed in the first session. In the following two sessions, isometric muscle strength was measured using a calibrated MicroFET2 dynamometer (Figure 5.1). The full protocol is illustrated in Figure 5.12.

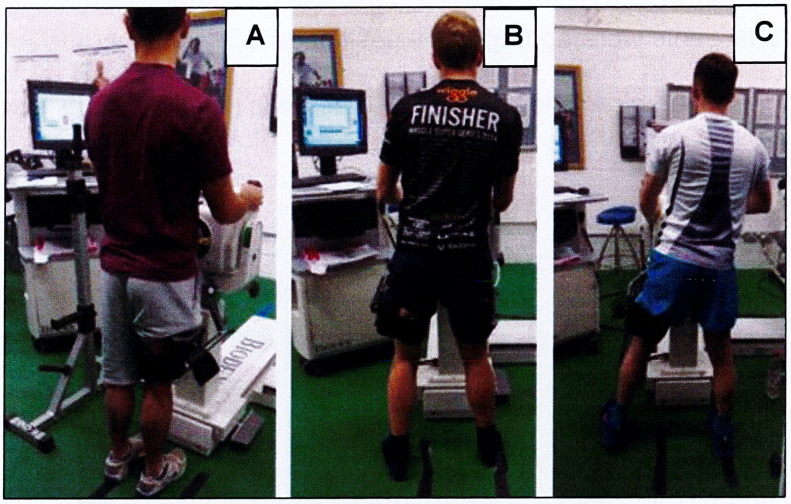


Figure 5.9 – Biodex isometric testing: (A) at 0 degrees of abduction; (B) at 15 degrees of abduction; (C) at 30 degrees of abduction.

5.6.2. *MicroFET2 Testing Sessions*

1) *Preparation:* The second testing session was performed within 7 days of the first session. Following the warm-up and skin marking, the testing procedure was explained to the participant. The third session was conducted within 14 to 30 days of the second session.
2) *Positioning and familiarisation:* The participants were asked to lay down in a supine position on the GrAASP. An imaginary axis of motion passing through the hip at the greater trochanter was aligned with the intersection between the hip and limb reference lines to position the hip joint appropriately. Next, a strap was placed around the opposite limb for stability (Figure 5.10). The MicroFET2 pad was positioned perpendicular to the skin, with the skin mark located centrally. The participants were not allowed to hold on to the platform to stabilise

themselves. Once the setup was completed, the participant was given 3 submaximal familiarisation attempts followed by a rest period of 60 seconds. To measure HAB muscle strength, the MicroFET2 was positioned on the LFC (proximal placement) and the LM (distal placement). For HAD muscle strength, the MicroFET2 was positioned on the MFC (proximal placement) and the MM (distal placement).

3) *Testing:* The participants were instructed to gradually build up and produce MVC, avoid sudden forceful contractions and lifting off the heel. For each strength measurement, the participant was asked to perform 3 isometric MVCs for 5 seconds, followed by 30 seconds rest period and the highest recorded value was considered in calculations (Fenter, et al., 2003b; Katoh, et al., 2011; Thorborg, et al., 2010a).

4) *Testing Sequence:* The placement location and stabilisation method were randomly selected before testing. Then, the placement location was changed (e.g. change from distal to proximal). It was followed by testing the opposite muscle group using the same method and sequence of placement if applicable (Figure 5.10). Both stabilisation methods were applied when appropriate to minimise repeatedly changing the participant's position. However, the same muscle group was not tested twice in a row. Once completed, the opposite side was tested using the same sequence. While the opposite side was being tested, the first side was effectively at rest for at least 3 minutes (5 sec x 3 attempts x 2 locations x 2 muscle groups plus resting time in between sets) before the second round of measurements (i.e. another stabilisation method on the first side if needed). If more rest was required, the participant was given two extra minutes. The full testing protocol of phase two is illustrated Figure 5.13.

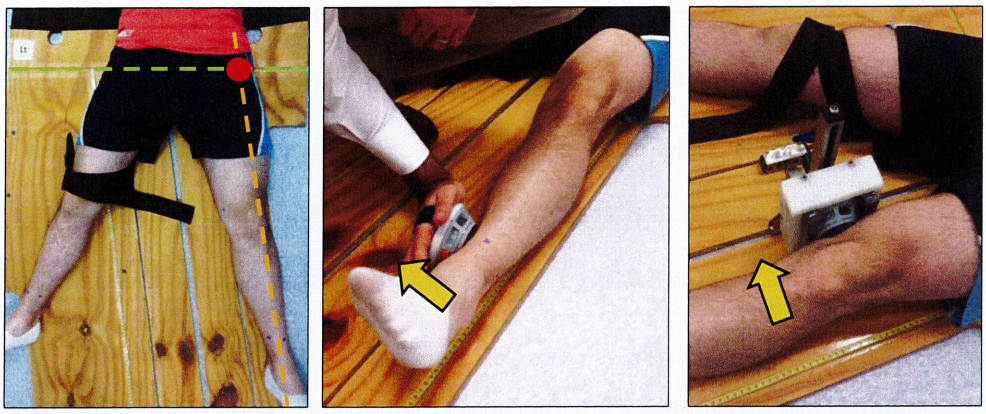


Figure 5.10 – Left hip alignment on the reference lines to test hip adduction strength while the opposite limb is stabilised (left); Left hip adduction strength measurement on a distal location using the manual method (middle); Left hip adduction strength measurement on a proximal location using the stabiliser (right) (The green and orange dashed lines represent the hip and limb reference lines respectively; the red dot represent the axis of motion; the yellow arrows represent the direction of push).

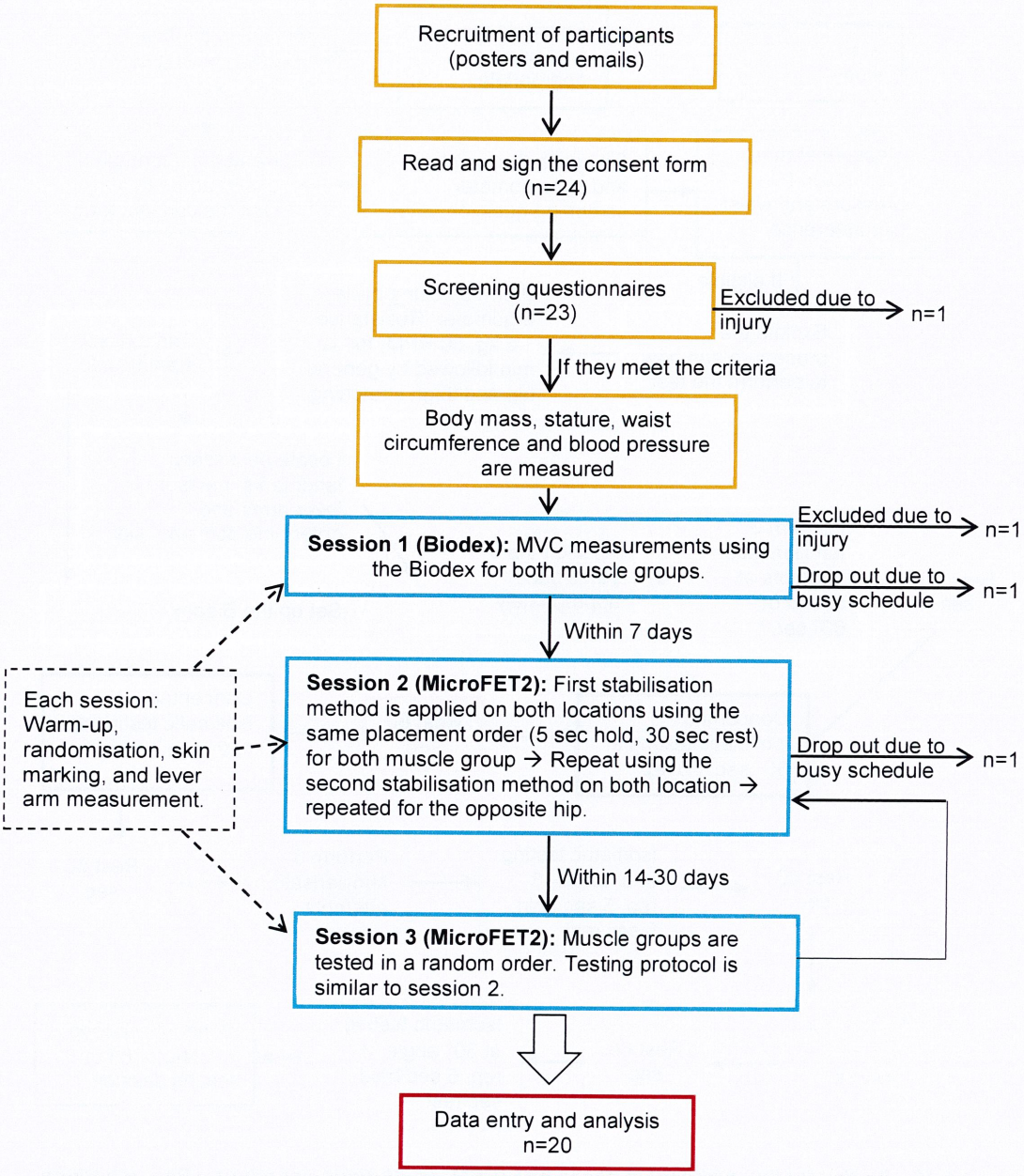


Figure 5.11 - General design of phase one testing (Orange boxes represent recruitment and screening phase; blue boxes represent testing phase; red box represents analysis phase; the dashed box represents pre-test preparation before each session).

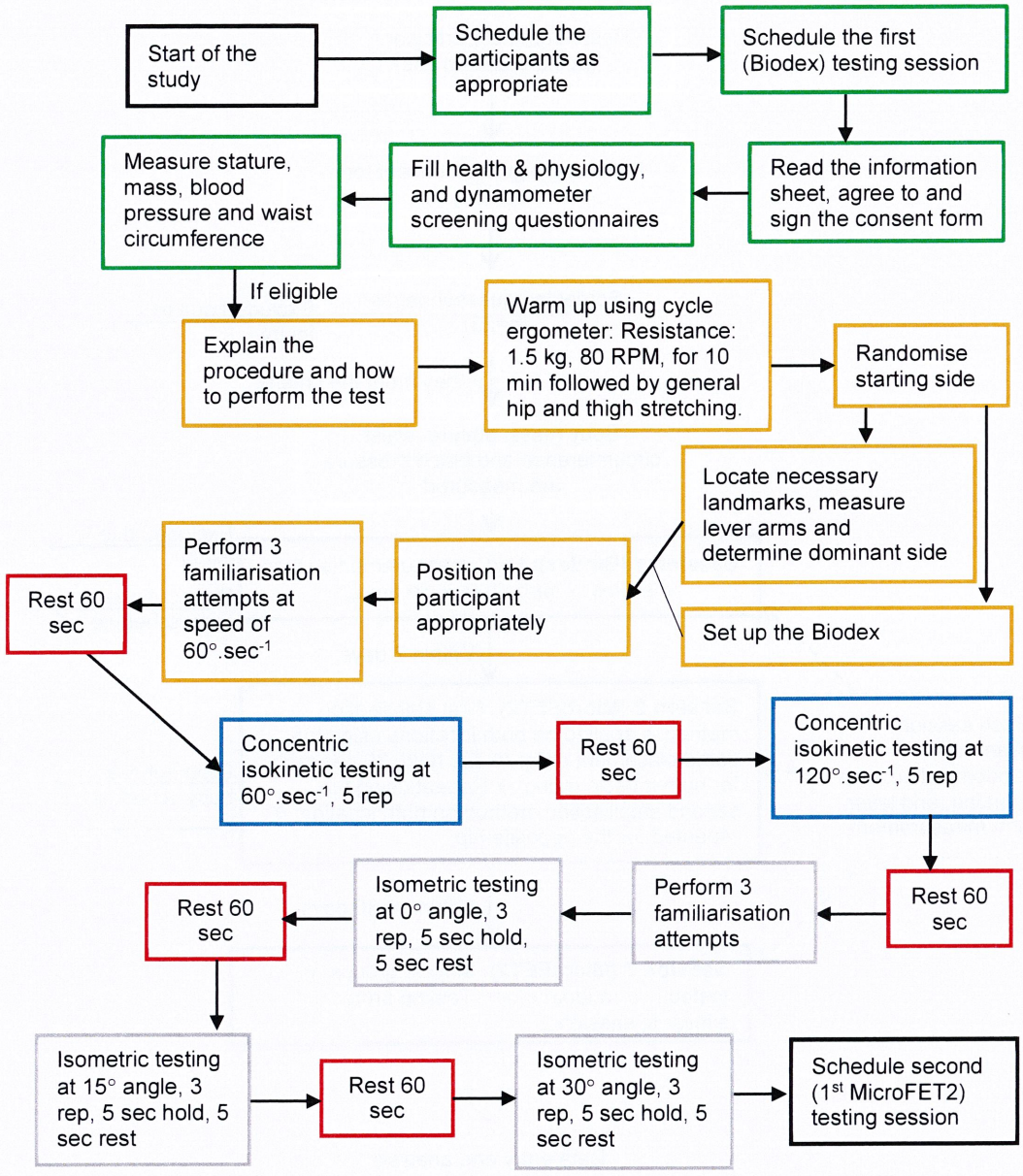


Figure 5.12 - Biodex testing protocol in phase one (RPM, revolutions per minute; Rep, repetition; sec, second; green boxes represent recruitment and screening phase; orange boxes represent test preparation phase; blue boxes represent isokinetic tests; grey boxes represent isometric tests and red boxes represent rest periods).

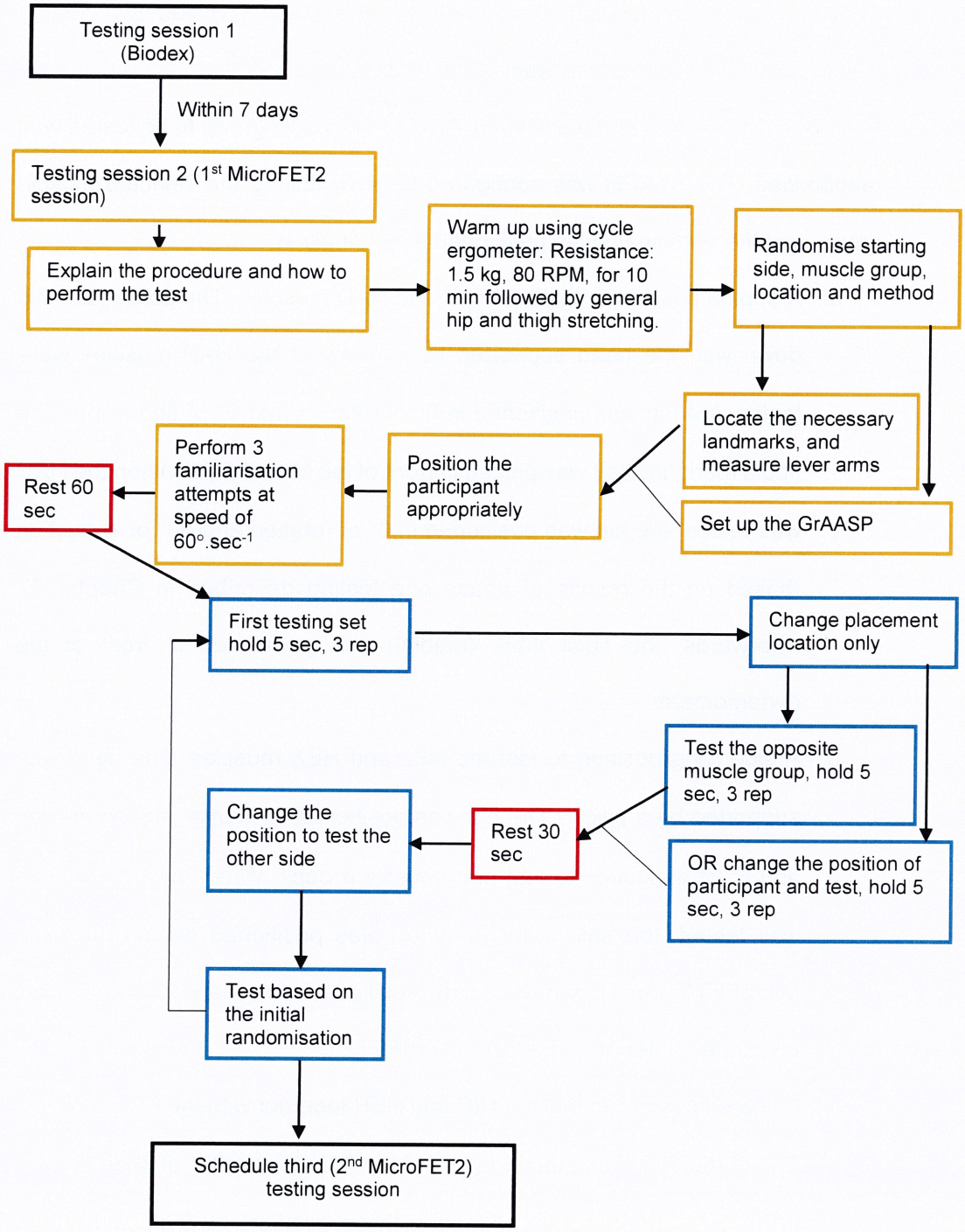


Figure 5.13 – MicroFET2 testing protocol in phase one (GrAASP, groin abductor-adductor strength platform; RPM, revolutions per minute; rep, repetitions; Orange boxes represent the preparation phase; Blue boxes represent the HDD testing stage; Red boxes represent the rest periods).

5.7. *Phase Two Testing Protocol*

1) *Preparation:* Following the warm-up and skin marking, the testing procedure was explained to the participant. This phase of the study involved two testing sessions that were performed within 7 to 10 days.
2) *Platform configuration and positioning:* The initial muscle group to be tested was randomised. The STAPH was configured to allow testing the selected muscle group, and the participant was positioned accordingly in:
    a. A supine position to test the HAB and HAD muscles. The participant laid down with the head supported by a pillow. If the HAB muscles were tested, the hip was positioned in 0° of rotation and 0° of abduction. The fibula mark (lateral) was placed in front of the MicroFET2. When the HAD was tested, the hip was positioned in 0° of rotation and 30° of abduction (based on the results of phase one testing described in Chapter 6). Afterwards, the tibial mark (medial) was positioned in front of the dynamometer.
    b. A side lying position to test the HFL and HEX muscles. The head was supported by a pillow. The hip was positioned in neutral at 0° of flexion and 0° of abduction to test both muscle groups. When the HFL muscle was tested, the shin mark (anterior) was positioned behind the fixed MicroFET2. For HEX muscle, the Achilles tendon mark (posterior) was positioned anterior to the fixed MicroFET2.
    c. A seated position to test the HIR and HER together with the KFL muscles. The participant was placed in 90° of hip flexion and 90° of knee flexion. If the joints could not be aligned in these positions, 2.5 cm wide foam pads were added under the pelvis until the desired position was achieved. The third wooden section was added on top of the other two wooden platforms for taller participants (Figure 5.14). After proper positioning, the fibula point was positioned in front of the MicroFET2 to

test the HIR. When the HER muscles were tested, the tibia mark (medial) was positioned in front of the MicroFET2. The Achilles tendon mark (posterior) was positioned in front of the MicroFET2 to test the KFL muscles.



Figure 5.14 - Using the foam pad (left) and the wooden section (right) to achieve proper hip and knee position.

3) *Familiarisation:* Once the setup was completed, the participant was given 3 submaximal familiarisation attempts followed by a rest period of 60 seconds.
4) *Testing:* The participants were instructed to gradually build up and produce MVC and avoid sudden forceful contractions. They were also instructed to put the hands across the chest when applicable and avoid stabilising themselves using hands or the contralateral leg. For each strength measurement, the participants were asked to perform 3 isometric MVCs for 5 seconds, followed by 30 seconds rest period and the highest recorded value was considered in calculations (Fenter, et al., 2003b; Katoh, et al., 2011; Thorborg, et al., 2010a).
    a. *HAB testing:* The participants were asked to keep the knee straight, avoid rotating the hip and to push outward against the MicroFET2 (Figure 5.15).
    b. *HAD testing:* The participants were asked to keep the knee straight, avoid rotating the hip and to push inward against the MicroFET2 (Figure 5.15).

c. *HFL testing:* The participants were given tactile feedback before testing to activate the HFLs. They were asked to keep the knee straight, tighten the quadriceps and push forward against the MicroFET2 (Figure 5.15).

d. *HEX testing:* The participants were given tactile feedback before testing to activate the HEXs. They were asked to keep the knee straight, tighten the hamstring and push backwards against the MicroFET2. (Figure 5.15).

e. *HIR testing:* The participants were asked to contract the gluteus medius and push the foot outwards against the MicroFET2 to internally rotate the hip (Figure 5.16).

f. *HER testing:* The participants were asked to push the foot inwards against the MicroFET2 to externally rotate the hip (Figure 5.16).

g. *KFL testing:* The participants were asked to bend the knee and push backwards against the MicroFET2 (Figure 5.16).

5) *Testing Sequence:* Each muscle group was followed by testing the antagonist muscle group (e.g. abductors then adductors) to minimise repeatedly changing the participant's position. After that, the MicroFET2 and the participant's positions were modified to test the opposite side. When both sides are tested, other muscle groups were tested in random. The general design of phase two testing protocol is illustrated in Figure 5.17. The protocol of MicroFET2 testing sequence is shown in Figure 5.18.

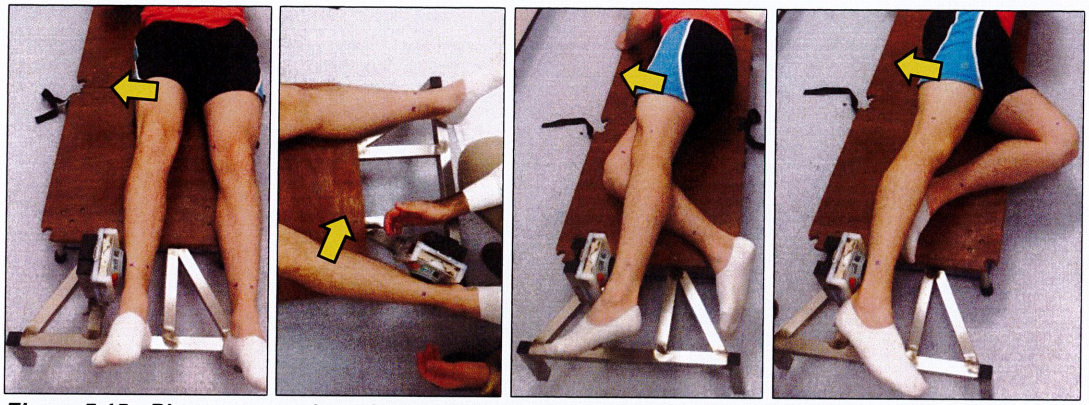


Figure 5.15 - Phase two testing of the hip abductor, adductor, flexor and extensor muscle groups (left to right) (yellow arrows represent the direction of push)

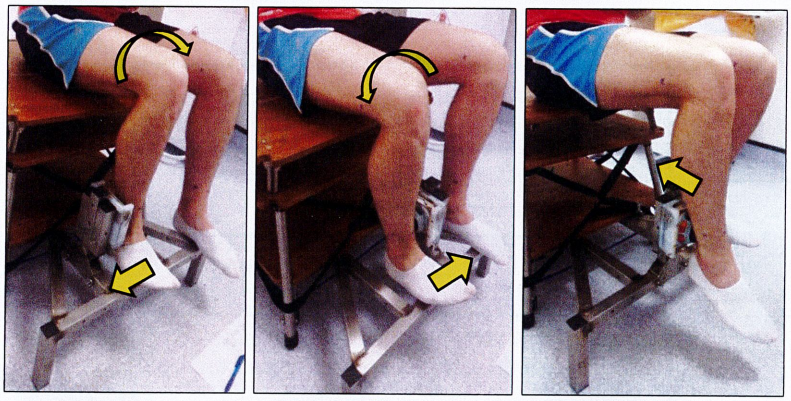


Figure 5.16 - Phase two testing of the hip internal rotator, external rotator and knee flexor muscle groups (left to right) (yellow arrows represent the direction of push).

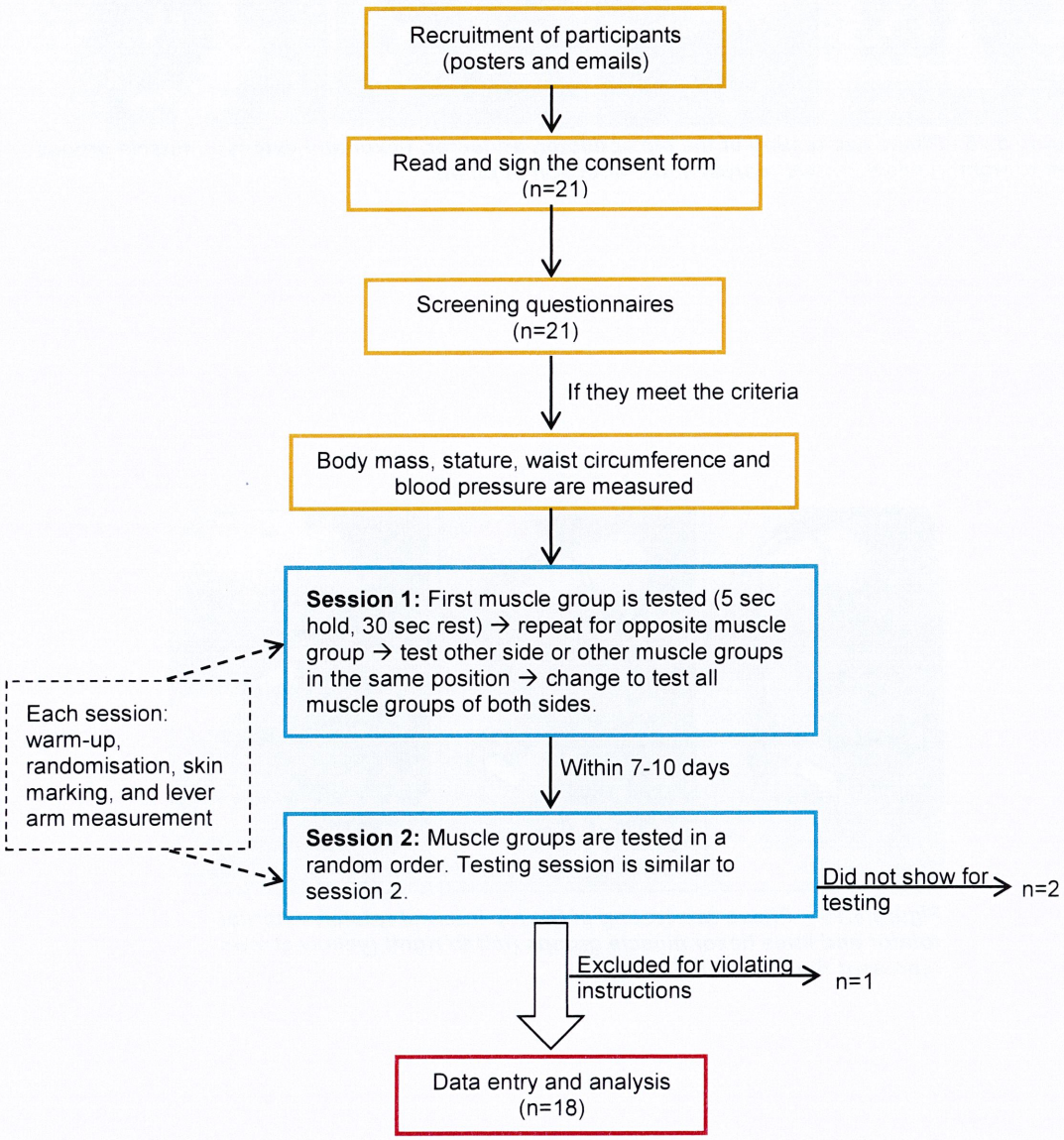


Figure 5.17 - General design of phase two testing (Orange boxes represent recruitment and screening phase; blue boxes represent testing phase; red box represents analysis phase; the dashed box represents pre-test preparation before each session).

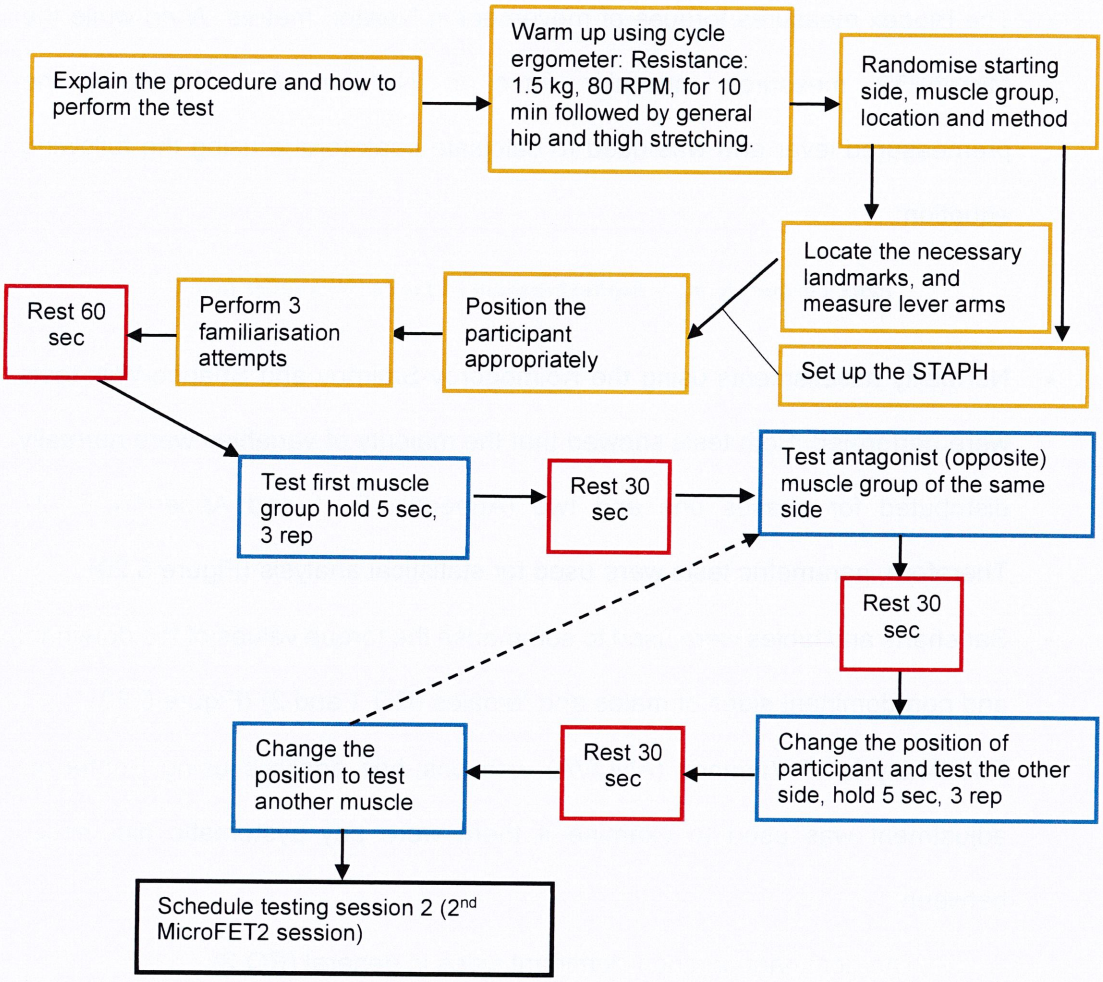


Figure 5.18 - MicroFET2 testing protocol in phase two (STAPH, hip strength assessment platform; RPM, revolutions per minute; rep, repetitions; Orange boxes represent the preparation phase; Blue boxes represent the MicroFET2 testing stage; Red boxes represent the rest periods).

5.8. *Statistical Analysis*

- The collected data were analysed using IBM Statistical Package for the Social Sciences (SPSS) version 22 (IBM, Chicago, IL). Demographical information including age, sex, Body Mass Index (BMI) and the recreational sport of interest was recorded.
- The Biodex measures torques of movement in Newton metres ($N{\cdot}m$) while the MicroFET2 measures muscular force in Newton (N). Therefore, the premeasured lever arm was used to calculate peak torque using the following equation:

$$Peak\ Torque\ (N.m) = highest\ measured\ force\ (N)\ x\ lever\ arm\ (m)$$

- Normality assessments using the Kolmogorov-Smirnov and Shapiro-Wilk tests were performed. Both tests showed that the majority of variables were normally distributed for phases one and two (Appendix E. 1 and Appendix E. 2). Therefore, parametric tests were used for statistical analysis (Figure 5.20).
- Bar charts and tables were used to summarise the torque values of the dominant and non-dominant sides of males and females (RQ 1 and 2) (Figure 5.20).
- The Analysis of Variance (ANOVA) with post-hoc analysis using Bonferroni adjustment was used to examine if there were any systematic differences between:
    1. The dominant and non-dominant sides in general (RQ 2).
    2. The dominant and non-dominant sides based on sex (RQ 2).
    3. The three hip angles in isometric testing and the two speeds in concentric-isokinetic testing (RQ 3 and 4) (Figure 5.20).

  The partial eta squared (ηp^2) is an outcome of ANOVA tests and was used to estimate effect size, which measures the magnitude of the difference between scores and practical significance regardless of the scale used (Lakens, 2013). The results were interpreted as 0 - 0.1, weak effect; 0.1 -

0.3, modest effect; 0.3 - 0.5, moderate effect and >0.5 was considered strong effect (Levine & Hullett, 2002).

- The ratio of muscle strength (RQ 5) was calculated by dividing strength of the agonist's muscle by the antagonist's muscle. Given that the lever arm was fixed for each participant, it was cancelled together with the torque unit. It results in a strength ratio rather than torque ratio. All statistical methods used to analyse the data are displayed in Figure 5.20.

$$Ratio\ of\ muscle\ strength = \frac{Agonist\ m.\ strength}{Antagonist\ m.\ strength}$$

- Bland-Altman plots were used to determine the level of agreement between various measurements, sessions and methods with paired *t*-tests to compare the scores (RQs 6, 7, 8 and 9) (Figure 5.20). The average of two measurements of the dominant side only was plotted against the difference between them to display a scatter plot. In phase one of testing, the results of Biodex testing were compared to the results of the first MicroFET2 testing session only.
    - The mean of difference or bias (MD or $\bar{d}$) was calculated by taking the average of the difference between two measurements. The 95% confidence interval of the agreement limits was calculated using the following formula (Bland & Altman, 2003):

$$95\%\ Confidence\ Interval\ of\ \bar{d} = \bar{d} \pm 1.96 \times SD\ of\ the\ differences$$

    - The coefficient of repeatability (*CR*), or the smallest real difference (*SRD*) considers both the random and systematic errors. It includes 95% of the differences between the two measurements and was calculated using the formula (Bland & Altman, 2003; Vaz, et al., 2013):

$$Coefficient\ of\ repeatability\ (CR) = 1.96 \times Standarad\ Deviation\ (SD)$$

- The percentage of measurement error (*ME* %) was calculated using the following formula to estimate the inaccuracy of the second measurement compared to the first (Hanneman, 2008):

$$Measurement\ Error\ (ME) = \frac{Confidence\ limit\ (range\ of\ confidence\ interval) of\ d}{Established\ or\ first\ device\ mean\ of\ measurements} \times 100$$

- The percentage of change (%) in scores in the second MicroFET2 session was compared to the first session using the following formula:

$$Percentage\ of\ change\ (\%) = \frac{2nd\ MicroFET2\ test\ scores - 1st\ MicroFET2\ test\ scores}{1st\ MicroFET2\ test\ scores} \times 100$$

- Interpretation of the plots was based on the recommendations of previously published work (Bland & Altman, 2003; Giavarina, 2015; Hanneman, 2008). The lowest mean of difference, agreement limit range and measurement error were considered collectively to draw the conclusion about the best location of application, method of stabilisation and platform. Although agreement tests are the best procedures to assess how much scores agree with each other, Pearson correlation (*r*) was calculated in phase one between the Biodex and the 1st MicroFET2 testing to compare the findings with the results of previously published research and will only be mentioned later in the discussion chapter.

- The coefficient of variation (CV) was calculated to measure how widely scores are dispersed and distributed around the mean. It is measured by taking the ratio between the standard deviation to the average scores of the dominant and non-dominant sides in the two MicroFET2 sessions (RQ 9). Therefore, it is considered a measure of relative variability and precision (Hopkins, 2000).

$$Coefficient\ of\ variation\ (CV\ \%) = \frac{Standard\ deviation\ (SD)}{Mean\ (M)}$$

- Sensitivity analysis using multilevel modelling (MLM) was used to explore the variability in the participants' scores in the two MicroFET2 sessions of phase one only (RQ 8). The models were built to analyse the effects of having two sessions and sex on the measurements. The comparison at Level one was performed to examine within-participants (repeated measures) variability (ε_{oj}) and at Level 2 to examine between-participants variability (μ_{oj}) (Figure 5.19). Building up the models and interpretation of the results were based on recommendations by others (Field, 2013; Heck, et al., 2014; Steele, 2008). The analysis of MLM results is similar to the procedure followed in Chapter 3 (section 3.5.4). Model 1 was considered a null model, in which no factors were considered. It was used to establish within- and between-patients' variabilities. This model was followed by two models, each contained an additional factor to explain the within- and between-patients' variabilities in scores (Table 5.1). For further details, the reader is advised to refer to Appendix G. 3. The models are summarised in Table 5.1.
    - Each model was compared with the null model. The conclusion was based on Model 3, which includes both factors. Interpretation of the results of Model 3 was as follow:
    - The degrees of freedom (*df*) and -2 Log Likelihood (*-2LL)* change from the null Model to Model 3 were compared against the Chi Square (χ^2) distribution table. After adding the factors, if the change in *df* and *-2LL* was more than the critical value (df_{change} = 2, $\chi^2_{0.05}$ = 5.991) it was considered a significant change at p =0.05. Therefore, this model is considered to have a statistically significant effect on explaining within- and between-participants' variability in scores. A copy of the Chi Square table can be found in Appendix B. 10.
        1. The estimate (regression coefficient) of a group or factor is considered to have a significant effect at p <0.05.

2. The reference groups in the coefficients table were set as the second session and being a female. Further information can be found under each table.

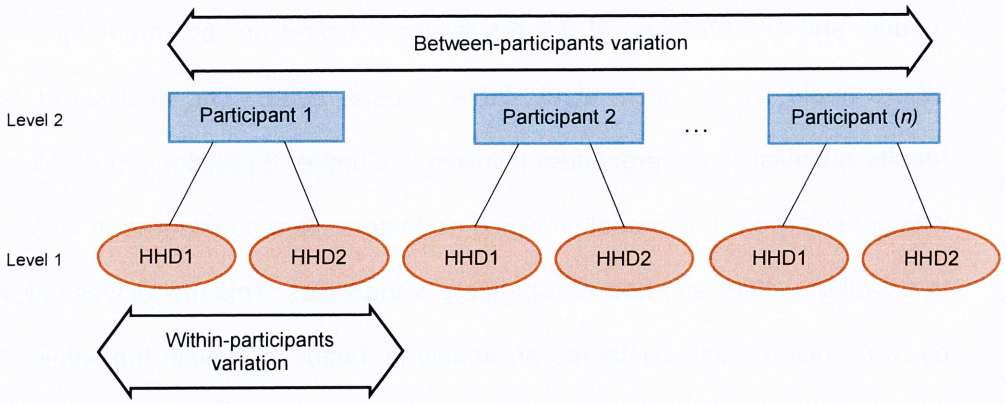


Figure 5.19 - Structure of the multilevel modelling (Orange circles represent level 1 and the blue boxes represent level 2 of the model)

Table 5.1 - The models used in multilevel modelling.

	Model 1	Model 2	Model 3
Factors	-	Two sessions	Two sessions
	-	-	Sex
Variation	ε_j & μ_j	ε_j & μ_j	ε_j & μ_j

ε_j, within patients variability; μ_j, between patients variability; Note: Model 1 is a null model and does not contain any factors (Heck, et al., 2014).

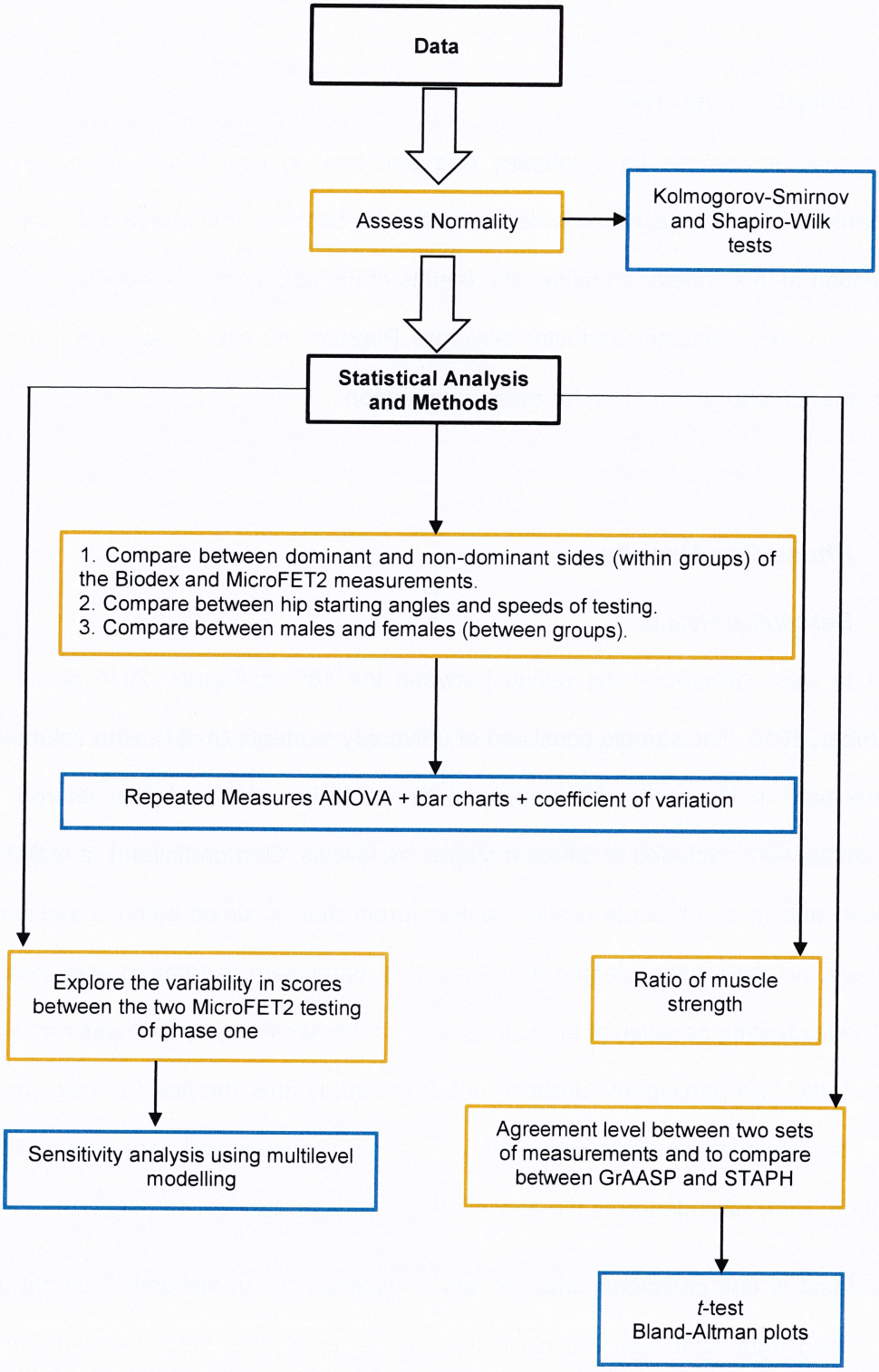


Figure 5.20 - Statistical methods used to analyse the data of phases one and two (ANOVA, analysis of variance; orange boxes represent aim of the test reflecting the research questions; blue boxes represent the statistical tests used; GrAASP, groin abductor-adductor strength platform; STAPH, hip strength assessment platform).

Chapter 6 Study Two: Results

6.1. *Chapter Overview*

In this chapter, the results of phases one and two of testing were reviewed and summarised. The findings were divided into sections based on the phase of testing. The last section of this chapter combines the results of hip abduction and adduction testing using the Groin Abductor-Adductor Strength Platform (GrAASP) and Hip Strength Assessment Platform (STAPH) for close comparison.

6.2. *Phase One Results*

6.2.1. Descriptive Results

The data was collected in the period between the 18th of August, 2015 and 20th of November, 2015. The sample consisted of university students (n=24), who volunteered to participate in the study. After applying the inclusion and exclusion criteria, four participants were excluded at different stages as follows: One participant (a male) was excluded due to a sub-acute grade-II ankle sprain that occurred before participation. Therefore, no data was collected from him. One participant (a female) was excluded after Biodex testing because of an acute grade-I ankle sprain injury that was not related to the study. Two participants dropped out of the study after the first (a male) and the second sessions (a female) due to a busy schedule. Consequently, any data collected from them were excluded from the analysis.

The dataset of one participant was not fully completed due to discomfort on the distal point during right adduction muscle testing. As a result, this measurement was not obtained. The final number of participants with complete data from the three testing sessions was nineteen (Table 6.1).

Overall, the enrolled participants had a mean age of 22.2 ± 4.2 years (*Mean* ± *SD*), had an average stature of 1.73 ± 0.06 *m* and an average body mass of 69.80 ± 9.87 *kg*. The sample's mean BMI was 23.19 ± 2.19 $kg.m^{-2}$, with the males having a higher BMI (24.29 ± 1.90 $kg.m^{-2}$) than the females (21.84 ± 1.76 $kg.m^{-2}$) due to higher mass in relation to stature. The average duration between the two consecutive MicroFET2 sessions was 23.8 ± 9.0 days. The dominant side was the right in 18 participants, while it was the left side in 2 participants (Table 6.2).

The participants were active for at least 3 hours a week and playing a variety of sports. Fitness training (*n*=3) and running (*n*=3) were the two most commonly played sports followed by lacrosse (*n*=3). Golf, swimming, Thai boxing and other sports (*n*=1 each) were played by other participants (Table 6.3).

Table 6.1 - Number of complete and partial datasets in phase one.

Dataset Description	***n***
Complete dataset (all three testing sessions)	19
Partial dataset (all measurements except right adduction distal testing)	1
Total (*n*)	20

Table 6.2 - Baseline characteristics of the participants in phase one.

Parameters		**Males**	**Females**	**Overall**
Age (years)		21.7 ± 3.6	23.8 ± 4.9	22.2 ± 4.2
Stature (*m*)		1.76 ± 0.04	1.69 ± 0.06	1.73 ± 0.06
Body mass (*kg*)		75.62 ± 8.07	62.78 ± 6.92	69.80 ± 9.87
Body Mass Index (BMI, $kg.m^{-2}$)		24.29 ± 1.90	21.84 ± 1.76	23.19 ± 2.19
Days between MicroFET2 sessions		22.8 ± 9.6	25.0 ± 8.6	23.8 ± 9.0
Dominant Side (*n*)	Right	10	8	18
	Left	1	1	2
Total (*n*)		11	9	20

Values are presented as mean ± standard deviation (*M* ± *SD*)

Table 6.3 - Sports played by the participants in phase one.

Sport	*n*
Fitness Training (gym workout)	3
Running	3
Lacrosse	3
Football (soccer)	2
Basketball, Triathlon, golf, swimming, netball, CrossFit, cycling, cricket and Thai boxing	1 each
Total (*n*)	20

6.2.2. Sides Comparison from Biodex Testing (RQs 1 and 2)

Isometric hip abduction (HAB) torque did not change much between 0°, 15° and 30° on the dominant side (DS). However, the torque values of the non-dominant side (NS) reduced as the hip angle increased from 116.55 ± 36.57 *N·m* at 0° to 109.94 ± 35.24 *N·m* at 30° (Figure 6.1). The DS was stronger by 0.4% at 15° and 4.4% at 30°. On the other hand, hip adduction (HAD) muscle torque on the DS and NS increased as the hip angle increased. Figure 6.2 shows that HAD torque values on the DS were less than those on the NS by 4.6% at 0°, and higher than the NS by 2.1% at 15° and by 0.5% at 30° (Figure 6.2). Further details of isometric torque values are listed in Appendix E. 3.

In isokinetic testing, HAB muscle torque on both sides increased as the speed increased from 60°.sec^{-1} (DS, 130.59 ± 38.49 *N·m*; NS, 124.15 ± 38.12 *N·m*) to 120°.sec^{-1} (DS, 135.30 ± 49.32 *N·m*; NS, 135.38 ± 46.61 *N·m*) (Figure 6.3). In contrast, HAD muscle torque on both sides dropped as the speed increased from 60°.sec^{-1} (DS, 81.41 ± 48.09 *N·m*; NS, 98.20 ± 45.79 *N·m*) to 120°.sec^{-1} (DS, 62.52 ± 43.95; NS, 72.88 ± 36.34 *N·m*) (Figure 6.4). The DS was stronger by 1.9% at 60°.sec^{-1}, while it was 2.1% weaker at 120°.sec^{-1} in HAB testing. On the other hand, the NS was stronger than the DS by 20.6% at 60°.sec^{-1} and 16.6% at 120°.sec^{-1} in HAD testing. The details of isokinetic torque values are listed in Appendix E. 3.

Key Results

- Isometric HAB muscle torque did not significantly change but isometric HAD muscle torque changed from 0° to 15° to 30°.
- Isokinetic HAB muscle torque increased, but isokinetic HAD muscle torque decreased as the speed changed from 60°.sec^{-1} to 120°.sec^{-1}.
- No statistical differences were found between the DS and NS in both modes.

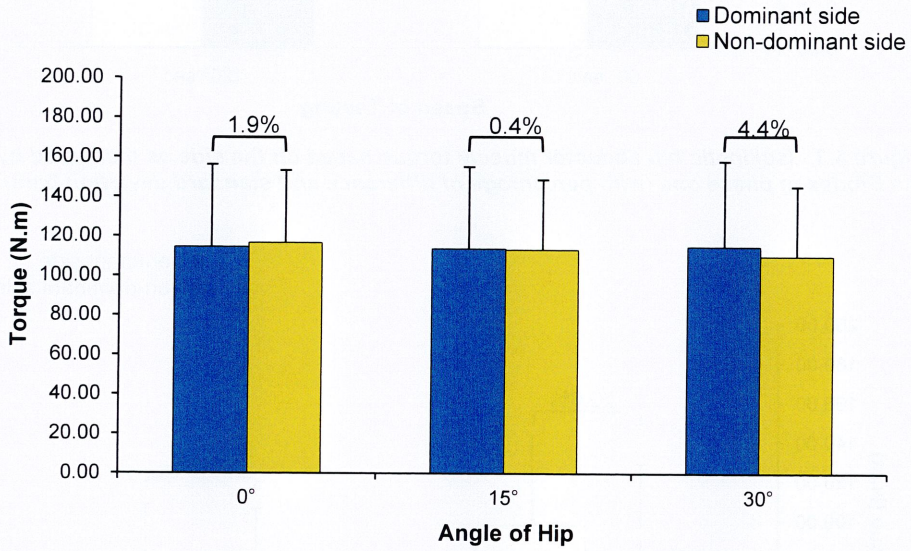


Figure 6.1 - Isometric hip abductor muscle torque based on the side as measured by the Biodex in phase one (with percentage of difference and standard deviation bars).

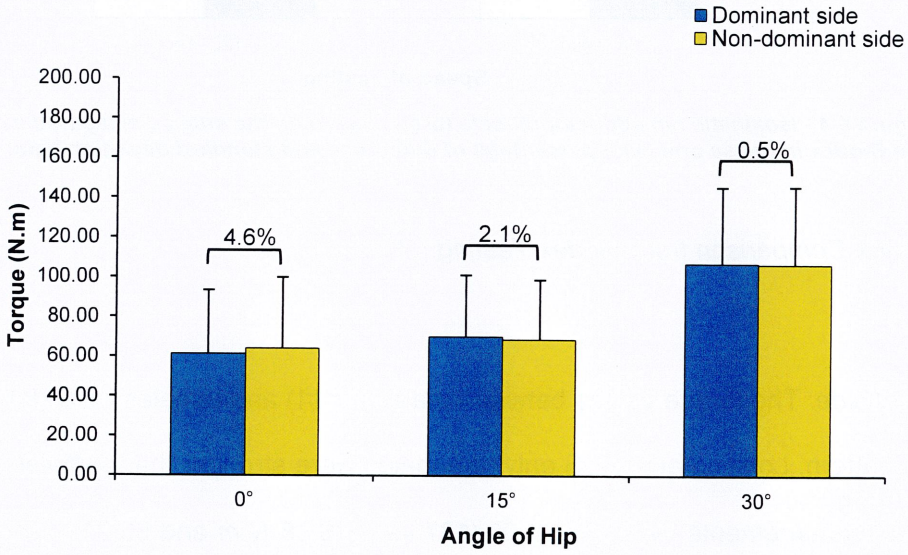


Figure 6.2 – Isometric hip adductor muscle torque based on the side as measured by the Biodex in phase one (with percentage of difference and standard deviation bars).

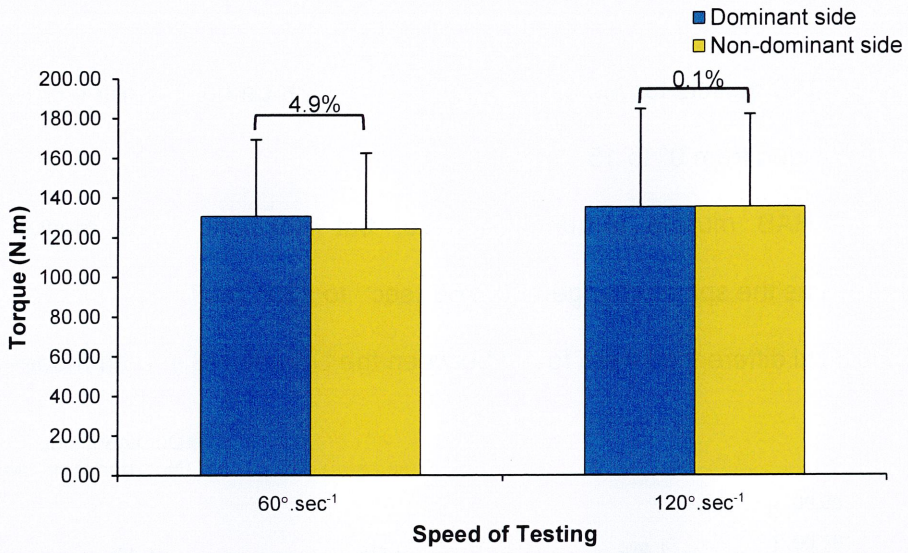


Figure 6.3 - Isokinetic hip abductor muscle torque based on the side as measured by the Biodex in phase one (with percentage of difference and standard deviation bars).

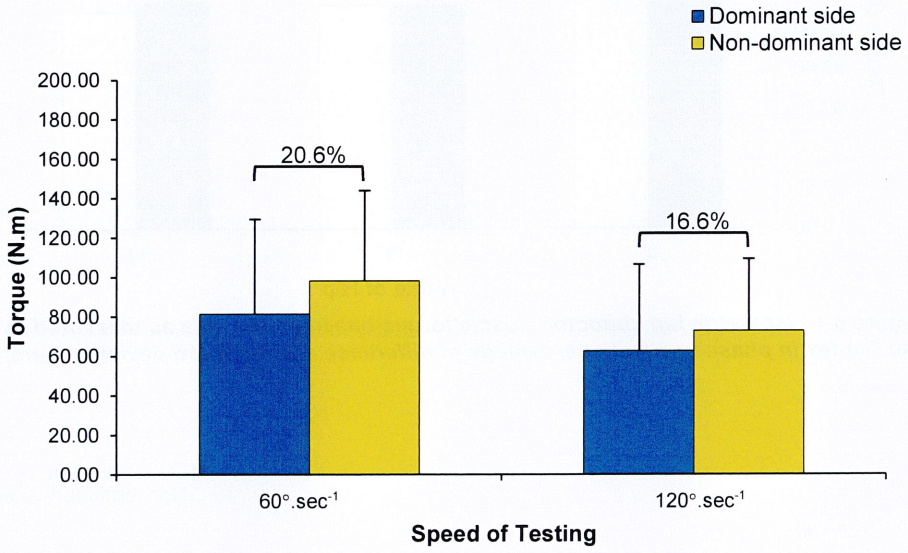


Figure 6.4 - Isokinetic hip adductor muscle torque based on the side as measured by the Biodex in phase one (with percentage of difference and standard deviation bars).

6.2.3. Sex Comparison from Biodex Testing (RQ 2)

Dominant side: The torque values between males (n=11) and females (n=9) followed a specific pattern. Looking at the DS only, the males were stronger than females in HAB isometric measurements by 37.3% at 0° (137.49 ± 38.68 *N·m* and 86.21 ± 25.05 *N·m* respectively), 28.3% at 15° (130.26 ± 39.99 *N·m* and 93.46 ± 35.56 *N·m* respectively) and 31.4% at 30° (133.86 ± 41.50 *N·m* and 91.83 ± 34.23 *N·m* respectively) (Figure

6.5). The males also scored higher by 49.5% at 0° (78.69 ± 33.42 and 39.73 ± 11.14 *N·m* respectively), 30.2% at 15° (80.96 ± 32.32 *N·m* and 56.49 ± 24.99 *N·m* respectively) and by 38.0% at 30° (128.92 ± 32.29 *N·m* and 79.97 ± 27.23 *N·m* respectively) in isometric HAD testing (Figure 6.6).

In isokinetic measurements, the torque values of both muscle groups changed as the speed increased from 60°.sec^{-1} to 120°.sec^{-1}. The males scores slightly increased from 152.04 ± 30.62 *N·m* to 154.46 ± 46.29 *N·m* for HAB testing and dropped from 88.56 ± 56.07 *N·m* to 71.06 ± 51.61 *N·m* for HAD testing (Figure 6.7 and Figure 6.8). The same was found in the females scores which showed an increase from 104.38 ± 30.67 *N·m* to 111.87 ± 44.44 *N·m* for HAB testing, while the HAD testing scores dropped from 72.66 ± 37.49 *N·m* to 52.09 ± 32.18 *N·m* (Figure 6.7 and Figure 6.8). The difference between males and females was 31.3% and 27.6% for the HAB, 18.0% and 26.7% for the HAD muscles at 60°.sec^{-1} and 120°.sec^{-1} respectively. All values and further information can be found in Appendix E. 4.

The repeated measures ANOVA revealed statistically significant differences between males and females in isometric HAB and HAD tests at 0°, isometric HAD tests at 30°, and isokinetic HAB testing at 60°.sec^{-1} (Table 6.4). The strongest effect in isometric HAB was at 0° (F [1, 19] =11.73, ηp^2= 0.395, p =0.009), and in isometric HAD it was at 30° (F [1, 19] =11.12, ηp^2= 0.420, p =0.006). While for isokinetic HAB it was observed at 60°.sec^{-1} (F [1, 19] =11.97, ηp^2= 0.399, p =0.006). There were no significant differences between both sexes with weak effect size in isokinetic HAD measurements (p =0.954) (Table 6.4).

Non-dominant side: The pattern of torque differences between both sexes is similar compared to the DS. However, the percentage of difference was much lower on the NS in comparison to the DS. The lower differences were observed in isometric HAB and HAD, and isokinetic HAB tests. All values and further information can be found in Appendix E. 4.

The results of the repeated measures ANOVA showed different findings compared to the DS (Table 6.5). The strongest effect size in isometric HAB was at 15° (F [1, 19] =13.78, ηp^2= 0.434, p =0.006), and in isometric HAD it was at 0° (F [1, 19] =10.23, ηp^2= 0.362, p =0.015). There were no significant differences between both sexes with weak effect size in isokinetic HAB and HAD measurements (p =818) (Table 6.5).

Key Results

- The male participants were stronger than the females in all measurements of the HAB and HAD muscle groups on both sides.
- The DS was stronger than the NS in all HAB measurements.
- On the DS, the strongest effect sizes were between isometric HAB at 0°, isometric HAD at 30°, and isokinetic HAB at 60°.sec^{-1}.

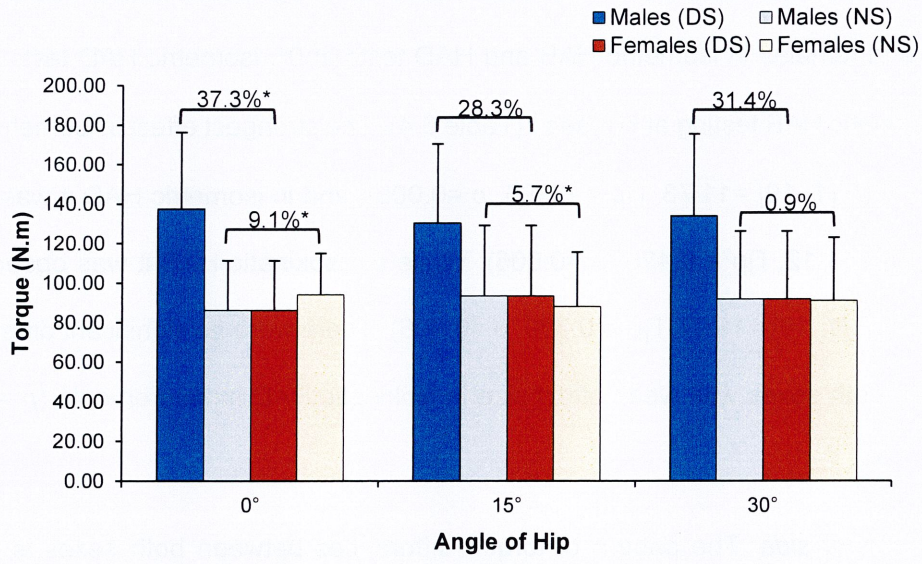


Figure 6.5 - Isometric hip abductor muscle torque based on gender and side as measured by the Biodex in phase one (with percentage of difference and standard deviation bars; DS, dominant side; NS, non-dominant side; *, significant difference between males and females at 0.05).

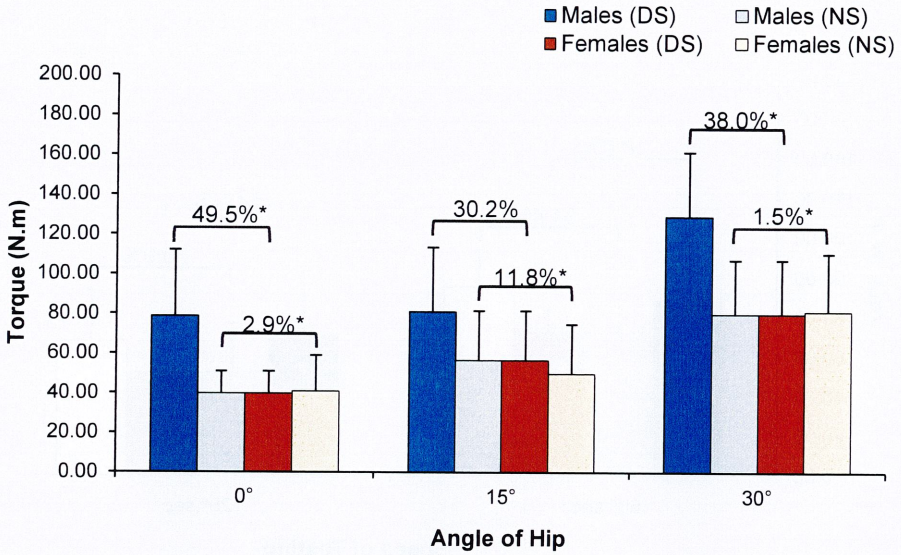


Figure 6.6 - Isometric hip adductor muscle torque based on gender and side as measured by the Biodex in phase one (with percentage of difference and standard deviation bars; DS, dominant side; NS, non-dominant side; *, significant difference between males and females at 0.05).

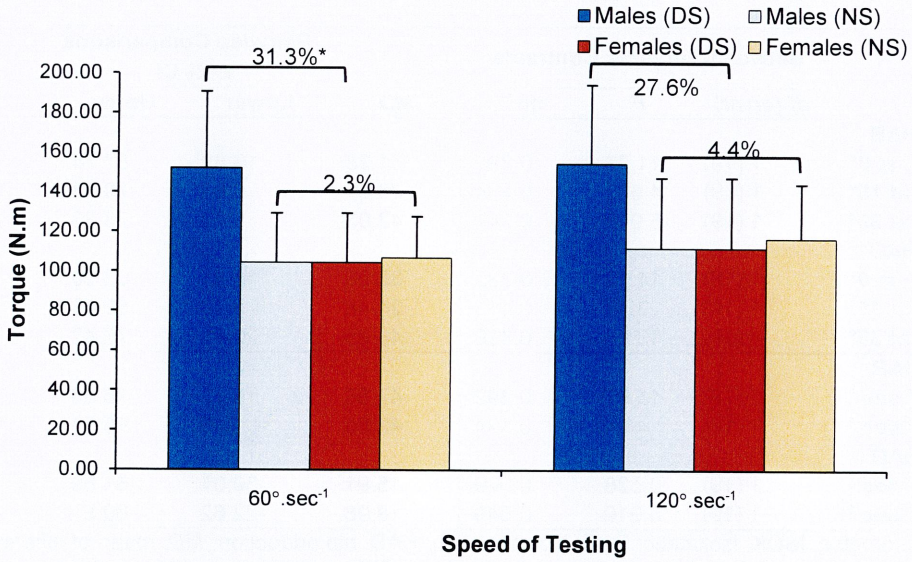


Figure 6.7 - Isokinetic hip abductor muscle torque based on gender and side as measured by the Biodex in phase one (with percentage of difference and standard deviation bars; DS, dominant side; NS, non-dominant side; *, significant difference between males and females at 0.05).

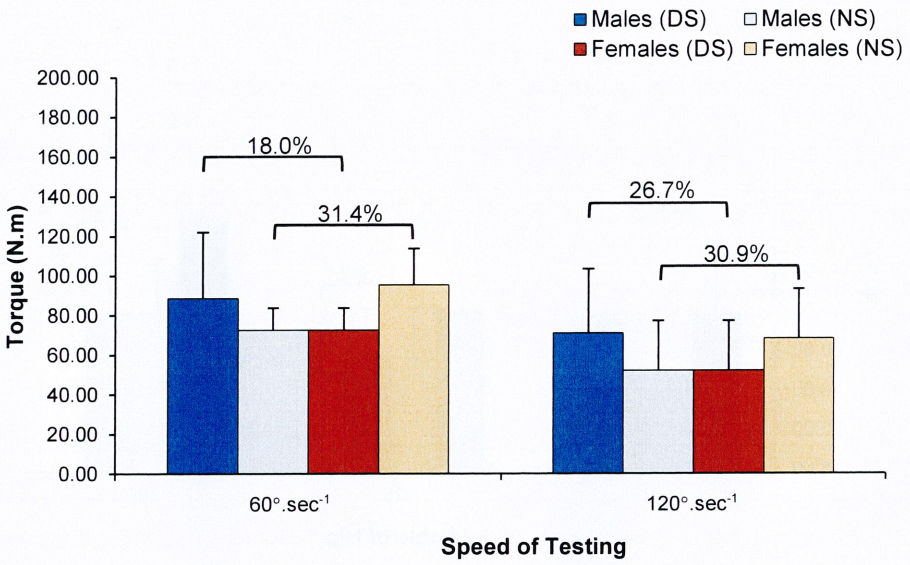


Figure 6.8 - Isokinetic hip adductor muscle torque based on gender and side as measured by the Biodex in phase one (with percentage of difference and standard deviation bars; DS, dominant side; NS, non-dominant side).

Table 6.4 - Repeated measures ANOVA test results comparing between male and female measurements in phase one testing (dominant side only).

	Between-Subjects Contrasts			Pairwise Comparisons 95% CI			
	df (error)	*F*	ηp²	*MD*	Lower	Upper	*p*
ISOM HAB							
at 0°	1 (19)	11.73	0.395	51.28	19.82	82.74	0.009*
at 15°	1 (19)	4.623	0.204	36.81	0.84	72.78	0.135
at 30°	1 (19)	5.918	0.247	42.03	5.73	78.33	0.078
ISOM HAD							
at 0°	1 (19)	11.12	0.382	38.96	14.41	63.50	0.012*
at 15°	1 (19)	3.46	0.161	24.48	-3.18	52.13	0.237
at 30°	1 (19)	13.05	0.420	48.95	20.48	77.42	0.006*
ISOK HAB							
at 60°.sec^{-1}	1 (19)	11.97	0.399	47.66	18.72	76.60	0.006*
at 120°.sec^{-1}	1 (19)	4.34	0.194	42.60	-0.346	85.54	0.104
ISOK HAD							
at 60°.sec^{-1}	1 (19)	0.528	0.029	15.91	-30.07	61.89	0.954
at 120°.sec^{-1}	1 (19)	0.919	0.049	18.98	-22.62	60.57	0.702

ISOM, isometric; ISOK, isokinetic; HAB, hip abduction; HAD, hip adduction; *MD, m*ean of difference; df, degrees of freedom; ηp², partial eta squared; *, significant difference between males and females at 0.05.

Table 6.5 - Repeated measures ANOVA test results comparing between male and female measurements in phase one testing (non-dominant side only).

	Between-Subjects Contrasts			Pairwise Comparisons 95% CI			
	df (error)	*F*	ηp^2	*MD*	Lower	Upper	*p*
ISOM HAB							
at 0°	1 (19)	8.712	0.326	40.92	11.79	70.05	0.027*
at 15°	1 (19)	13.78	0.434	45.67	19.82	71.53	0.006*
at 30°	1 (19)	5.98	0.249	34.48	4.86	64.10	0.075
ISOM HAD							
at 0°	1 (19)	10.23	0.362	42.04	14.42	69.66	0.015*
at 15°	1 (19)	9.068	0.335	34.05	10.29	57.80	0.021*
at 30°	1 (19)	9.981	0.357	45.86	15.36	76.36	0.015*
ISOK HAB							
at 60°.sec^{-1}	1 (19)	3.915	0.179	31.57	-1.95	65.08	0.126
at 120°.sec^{-1}	1 (19)	2.860	0.137	33.81	-8.20	75.82	0.216
ISOK HAD							
at 60°.sec^{-1}	1 (19)	0.054	0.003	4.92	-39.44	49.28	0.818
at 120°.sec^{-1}	1 (19)	0.260	0.014	8.50	-26.50	43.50	0.616

6.2.4. Angle and Speed Comparisons from Biodex Testing (RQs 3 and 4)

The findings of repeated measures ANOVA showed no significant differences between the three hip angles in isometric HAB testing (p =0.598). On the other hand, there was a statistically significant difference in isometric HAD (F [2, 38] =40.80, p <0.001), isokinetic HAB testing (F [1, 19] =4.411, p =0.049) and isokinetic HAD testing (F [1, 19] =22.22, p <0.001). However, pairwise comparisons imply that the only statistical difference was between 0° and 30° of hip abduction in isometric HAD testing (*MD*= 44.06, 95% CI [28.49, 59.63]) (Table 6.6).

The strongest effect sizes were found between 0° and 30° of hip abduction (ηp^2= 0.773), and between 0° and 15° of hip abduction in isometric HAD testing (ηp^2= 0.744). In isokinetic testing, only HAD testing showed a strong effect size (ηp^2= 0.539) (Table 6.6).

Key Results

- There were statistically significant differences between the three angles of hip abduction in isometric HAD testing only.
- Strong effect sizes were seen in isometric and isokinetic HAD testing.

Table 6.6 - Repeated measures ANOVA test results comparing between the hip angles of abduction in isometric testing and speeds in isokinetic testing of phase one testing.

	Within-Subjects Contrasts			Pairwise Comparisons 95% CI			
	df (error)	*F*	ηp^2	*MD*	Lower	Upper	*p*
ISOM HAB							
0°, 15° and 30°	2 (28)	0.521	0.027	-	-	-	0.598
0° vs 15°	1 (19)	0.524	0.027	-2.01	-9.28	5.27	0.478
0° vs 30°	1 (19)	0.588	0.030	-3.04	-13.44	7.37	0.453
ISOM HAD							
0°, 15° and 30°	2 (28)	40.80	0.682	-	-	-	<0.001*
0° vs 15°	1 (19)	55.21	0.744	6.68	-6.76	20.11	0.623
0° vs 30°	1 (19)	64.74	0.773	44.06	28.49	59.63	<0.001*
ISOK HAB							
60°.sec^{-1} vs 120°.sec^{-1}	1 (19)	4.411	0.188	3.18	-5.37	11.73	0.049*
ISOK HAD							
60°.sec^{-1} vs 120°.sec^{-1}	1 (19)	22.22	0.539	-13.57	0.045	27.10	<0.001*

ISOM, isometric; ISOK, isokinetic; HAB, hip abduction; HAD, hip adduction; *MD, m*ean of difference; df, degrees of freedom; ηp^2, partial eta squared; *, significant difference at 0.05.

6.2.5. Muscle Strength Ratios from Biodex and MicroFET2 Testing (RQ 5)

The strength ratios between the HAD and HAB muscle groups (HAD/HAB) were calculated on the dominant side. The ratios from Biodex isometric testing showed a gradual increase from 0.55 ± 0.25 to 0.73 ± 0.55, and 1.11 ± 0.86 for 0°, 15°, and 30° respectively. In contrast, the ratios obtained from isokinetic testing showed the opposite results. At 60°.sec^{-1}, the ratio was 0.72 ± 0.57, and dropped to 0.61 ± 0.71 at 120°.sec^{-1} (Table 6.7). In the first MicroFET2 testing session and using the fixed method, the ratios were 1.78 ± 0.98 proximally and 1.07 ± 0.35 distally. In the second session, the ratios increased slightly to 1.54 ± 0.44 for the proximal placement and 1.11 ± 0.19 for the distal placement (Table 6.7).

Key Results

- From Biodex testing, the change in muscle strength ratio was related to the change in hip abduction angle and speed of testing.
- From MicroFET2 testing, the proximal location of testing showed higher ratios compared to the distal location.

Table 6.7 - Ratios of hip muscle strength obtained by the Biodex and MicroFET2 in phase one (dominant side only).

Mode		Ratio of HAD / HAB	*M*	*SD*
Biodex	**Isom**	at 0°	0.55	0.25
		at 15°	0.73	0.55
		at 30°	1.11	0.86
	Isok	at 60°.sec^{-1}	0.72	0.57
		at 120°.sec^{-1}	0.61	0.71
MicroFET2 S1	**Man**	Prox	1.19	0.33
		Dist	1.19	0.40
	Fix	Prox	1.78	0.98
		Dist	1.07	0.35
MicroFET2 S2	**Man**	Prox	1.20	0.28
		Dist	1.17	0.21
	Fix	Prox	1.54	0.44
		Dist	1.11	0.19

Isom, isometric; isok, isokinetic; man, manual; fix, fixed; S, session; HAB, hip abduction; HAD, hip adduction; *M, m*ean; *SD,* standard deviation.

6.2.6. Abduction Isometric Tests on the Biodex compared with MicroFET2 Scores (RQs 6 and 7)

The results of limits of agreement tests were consistent between HAB isometric testing on the Biodex and using the MicroFET2. Statistically significant differences in torque measurements were found between 0°, 15°, and 30° of HAB isometric testing and the proximal placement location using both methods of stabilising the MicroFET2:

- Isometric at 0° with HAB manual proximal (Pair 1, Table 6.8).
- Isometric at 0° with HAB fixed proximal (Pair 3, Table 6.8).

- Isometric at 15° with HAB manual proximal (Pair 5, Table 6.8).
- Isometric at 15° with HAB fixed proximal (Pair 7, Table 6.8).
- Isometric at 30° with HAB manual proximal (Pair 9, Table 6.8).
- Isometric at 30° with HAB fixed proximal (Pair 11, Table 6.8).

The differences were not statistically significant when the distal location of placement was considered (p =268) (Pair 4, 6, 8, 10 and 12, Table 6.8). The measurements obtained using the dynamometer were less than those acquired using the Biodex leading to negative means of differences (mean of dynamometer testing results – mean of Biodex testing results). The smallest measurement error (*ME*= ±21.11%) and the narrowest range of agreement limits were between HAB fixed distally using the MicroFET2 and isometric testing at 0° (*MD*= -9.21 $N{\cdot}m$, 95% LoA [-61.57, 43.14], p =0.130) (Pair 4, Table 6.8; Figure 6.9). The results of limits of agreement tests are summarised in Table 6.8. For all Bland-Altman plots, refer to Appendix E. 5.

Key Results

- There were statistically significant differences between HAB isometric Biodex testing and the proximal placement location of the MicroFET2 only (using both methods of stabilisation).
- The smallest *ME* was between isometric testing at 0° on the Biodex and the fixed method distally for the MicroFET2.

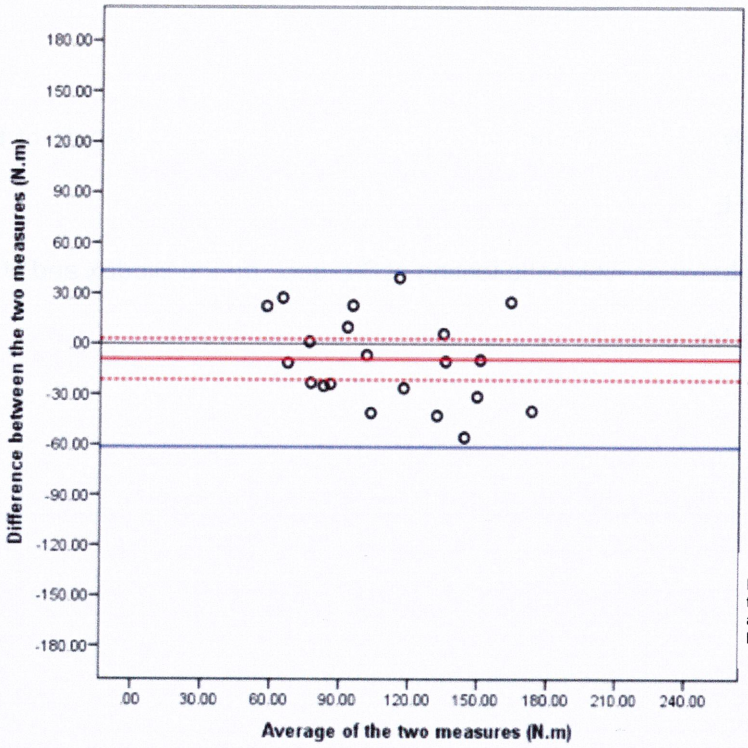


Figure 6.9 - Bland-Altman plot between abduction isometric at 0 and abduction fixed distal in phase one (pair 4 from Table 6.8; red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits).

Note: Scale of the plot is based on the biggest difference (180 N·m) and highest average of scores (240 N·m) reported in both phases.

6.2.7. Abduction Isokinetic Tests on the Biodex compared with MicroFET2 Scores (RQs 6 and 7)

Isokinetic HAB measurements at 60°.sec^{-1} and 120°.sec^{-1} were all significantly different from HAB measurements obtained by the MicroFET2 (Table 6.8). The tests of limits of agreement revealed that the measurements obtained using the MicroFET2 were less than those obtained using the Biodex, resulting in negative means of differences across all comparisons. The smallest measurement error (*ME*= ±17.54%) was between isokinetic testing at 60°.sec^{-1} and HAB manual distally (*MD*= -27.93 *N·m*, 95% LoA [-76.95, 21.09], p <0.001) (Comb 14, Table 6.8; Figure 6.10). The results of limits of agreement tests are summarised in Table 6.8. For all Bland-Altman plots, refer to Appendix E. 5.

Key Results

- There were statistically significant differences between HAB measurements of the Biodex and the MicroFET2.
- The smallest *ME* was between isokinetic testing at 60°.sec^{-1} on the Biodex and HAB manual distally for the MicroFET2.

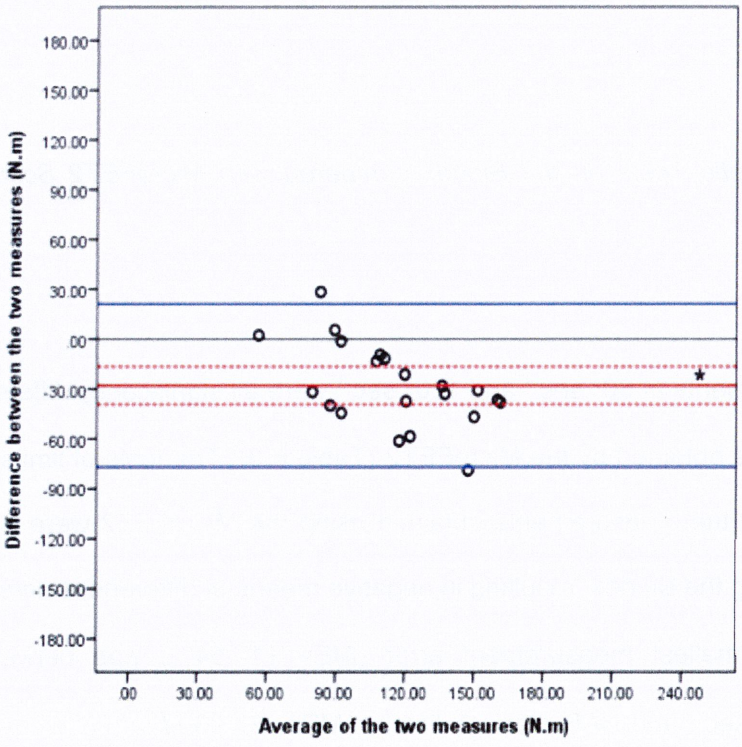


Figure 6.10 - Bland-Altman plot between abduction isokinetic at 60°.sec^{-1} and abduction manual distal in phase one (pair 14 from Table 6.8; red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits; *, significant difference at 0.05).

Note: Scale of the plot is based on the biggest difference (180 N·m) and highest average of scores (240 N·m) reported in both phases.

Table 6.8 – Limits of agreement tests between hip abduction measurements of the Biodex and the 1st MicroFET2 testing session in phase one (dominant side only).

Pair	Biodex	MicroFET2	Ms1 (*N·m*)	Ms2 (*N·m*)	Paired *t*-test *t*	Paired *t*-test *p*	Mean of Difference (Bias) Mean (*N·m*)	*SD* (±)	95% CI Lower	95% CI Upper	LLA (*N·m*)	ULA (*N·m*)	Limit range	Error (%)
1	HAB Isometric at 0°	HAB MAN PROX	115.19	66.26	-6.560	<0.001*	-48.93	34.18	-64.49	-33.37	-115.92	18.06	133.99	27.02
2		HAB MAN DIS	115.19	101.86	-2.105	0.048*	-13.33	29.03	-26.55	-0.12	-70.23	43.57	113.80	22.95
3		HAB FIX PROX	115.19	62.74	-6.686	<0.001*	-52.45	35.95	-68.82	-36.09	-122.91	18.01	140.92	28.41
4		HAB FIX DIS	115.19	105.98	-1.581	0.130	-9.21	26.71	-21.37	2.94	-61.57	43.14	104.70	21.11
5	HAB Isometric at 15°	HAB MAN PROX	113.19	66.26	-6.451	<0.001*	-46.93	33.34	-62.11	-31.76	-112.28	18.41	130.69	26.81
6		HAB MAN DIS	113.19	101.86	-1.829	0.082	-11.33	28.39	-24.26	1.59	-66.98	44.32	111.30	22.84
7		HAB FIX PROX	113.19	62.74	-5.809	<0.001*	-50.45	39.80	-68.57	-32.33	-128.46	27.56	156.02	32.01
8		HAB FIX DIS	113.19	105.98	-1.139	0.268	-7.21	29.02	-20.42	5.99	-64.08	49.66	113.74	23.34
9	HAB Isometric at 30°	HAB MAN PROX	114.83	66.26	-6.626	<0.001*	-48.58	33.60	-63.87	-33.28	-114.42	17.27	131.69	26.63
10		HAB MAN DIS	114.83	101.86	-2.077	0.051	-12.98	28.63	-26.01	0.06	-69.09	43.13	112.23	22.70
11		HAB FIX PROX	114.83	62.74	-5.715	<0.001*	-52.10	41.77	-71.11	-33.08	-133.96	29.77	163.74	33.12
12		HAB FIX DIS	114.83	105.98	-1.352	0.191	-8.86	30.02	-22.52	4.81	-67.69	49.98	117.67	23.80
13	HAB Isokinetic at 60°.sec^{-1}	HAB MAN PROX	129.79	66.26	-9.897	<0.001*	-63.53	29.42	-76.92	-50.14	-121.19	-5.87	115.32	20.64
14		HAB MAN DIS	129.79	101.86	-5.118	<0.001*	-27.93	25.01	-39.32	-16.55	-76.95	21.09	98.04	17.54
15		HAB FIX PROX	129.79	62.74	-8.033	<0.001*	-67.05	38.25	-84.46	-49.64	-142.03	7.92	149.95	26.83
16		HAB FIX DIS	129.79	105.98	-3.956	0.001*	-23.81	27.59	-36.37	-11.26	-77.89	30.26	108.15	19.35
17	HAB Isokinetic at 120°.sec^{-1}	HAB MAN PROX	133.84	66.26	-7.431	<0.001*	-67.58	41.68	-86.56	-48.61	-149.28	14.11	163.39	28.35
18		HAB MAN DIS	133.84	101.86	-3.959	0.001*	-31.99	37.02	-48.84	-15.13	-104.55	40.58	145.13	25.18
19		HAB FIX PROX	133.84	62.74	-6.656	<0.001*	-71.10	48.96	-93.39	-48.82	-167.06	24.85	191.91	33.30
20		HAB FIX DIS	133.84	105.98	-3.256	0.004*	-27.87	39.22	-45.72	-10.01	-104.74	49.01	153.76	26.68

Ms1, first measurement; Ms2, second measurement; Prox., proximal; Dis., distal; HAB, hip abduction; MAN, manual stabilisation; FIX, fixed stabilisation, *M*, mean; *SD*, standard deviation; CI, confidence interval; LLA, lower limit of agreement; ULA, upper limit of agreement; *, significant difference between the two measurements at 0.05.

6.2.8. Adduction Isometric Tests on the Biodex compared with MicroFET2 Scores (RQs 6 and 7)

The limits of agreement tests showed significant differences between the values obtained using the Biodex and those collected using the MicroFET2 in:

- Isometric at 0° with HAD manual proximal (Pair 21, Table 6.9).
- Isometric at 0° with HAD manual distal (Pair 22, Table 6.9).
- Isometric at 0° with HAD fixed proximal (Pair 23, Table 6.9).
- Isometric at 0° with HAD fixed distal (Pair 24, Table 6.9).
- Isometric at 15° with HAD manual distal (Pair 26, Table 6.9).
- Isometric at 15° with HAD fixed proximal (Pair 27, Table 6.9).
- Isometric at 15° with HAD fixed distal (Pair 28, Table 6.9).
- Isometric at 30° with HAD manual proximal (Pair 29, Table 6.9).
- Isometric at 30° with HAD manual distal (Pair 30, Table 6.9).

Means of differences (bias) between the two devices were mainly positive indicating that measurements obtained using the MicroFET2 were higher than those of the Biodex. In isometric testing, the smallest MD and ME were between isometric testing at 30° and fixed method proximally (*MD*= -4.43 *N·m*, 95% LoA [-68.38, 59.52], *p* =0.541), and fixed method distally (*MD*= 12.80 *N·m*, 95% LoA [-42.84, 68.43], *p* =0.731) (Pair 31 and 32, Table 6.9; Figure 6.11 and Figure 6.12). The results of limits of agreement tests are summarised in Table 6.9. For all Bland-Altman plots, refer to Appendix E. 5.

Key Results

- Hip adduction scores obtained from the MicroFET2 measurements were higher than those obtained from the Biodex.
- The two smallest MD and ME values were found between isometric testing at 30° on the Biodex and the fixed method for the MicroFET2 when applied using both placement locations.

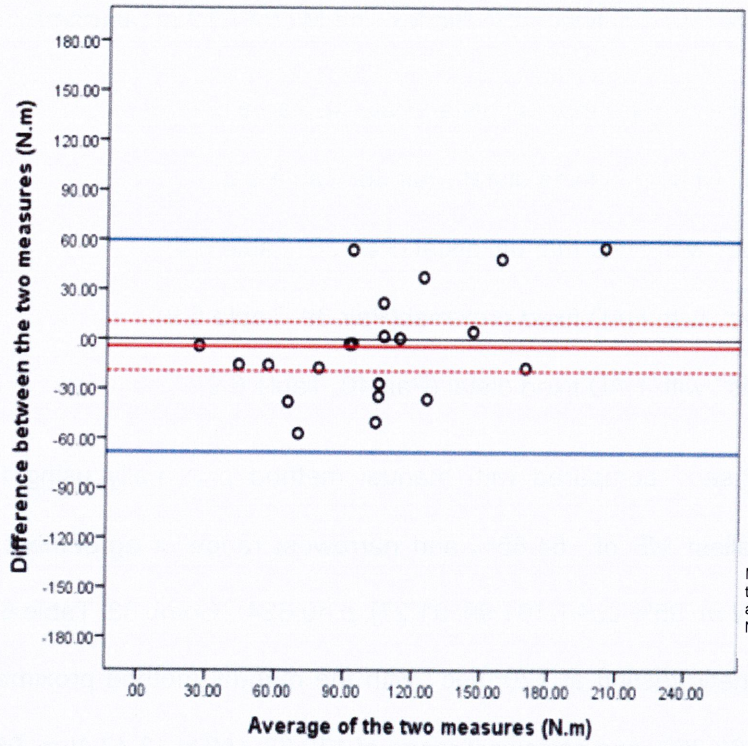


Figure 6.11 - Bland-Altman plot between adduction isometric at 30° and adduction fixed proximal in phase one (pair 31 from Table 6.9; red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits).

Note: Scale of the plot is based on the biggest difference (180 N·m) and highest average of scores (240 N·m) reported in both phases.

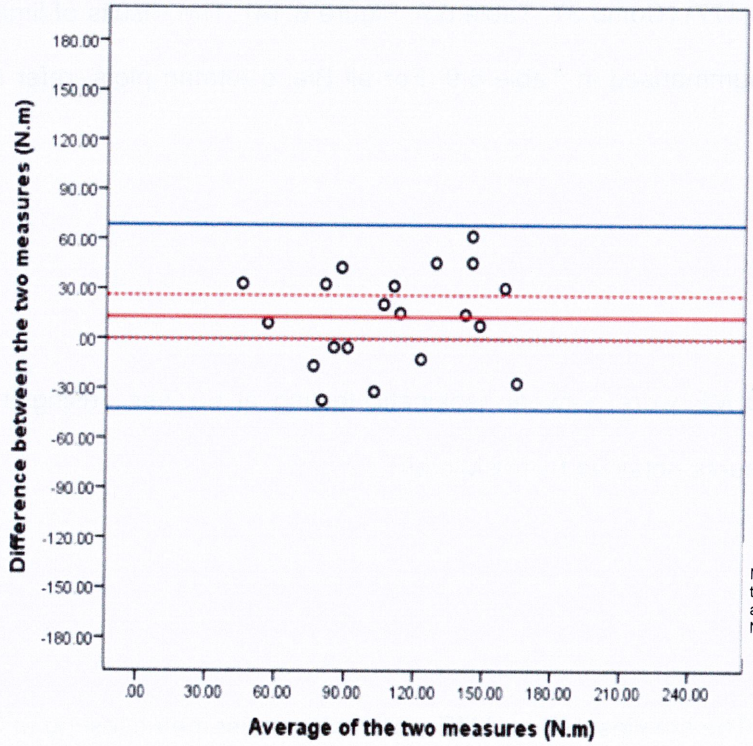


Figure 6.12 - Bland-Altman plot between adduction isometric at 30° and adduction fixed distal in phase one (pair 32 from Table 6.9; red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits).

Note: Scale of the plot is based on the biggest difference (180 N·m) and highest average of scores (240 N·m) reported in both phases.

6.2.9. Adduction Isokinetic Tests on the Biodex compared with MicroFET2 Scores (RQs 6 and 7)

Statistical differences were found between the Biodex and MicroFET2 measurements:

- Isokinetic at 60°.sec^{-1} with HAD manual distal (Pair 34, Table 6.9).
- Isokinetic at 60°.sec^{-1} with HAD fixed distal (Pair 36, Table 6.9).
- Isokinetic at 120°.sec^{-1} with HAD manual distal (Pair 38, Table 6.9).
- Isokinetic at 120°.sec^{-1} with HAD fixed proximal (Pair 39, Table 6.9).
- Isokinetic at 120°.sec^{-1} with HAD fixed distal (Pair 40, Table 6.9).

Isokinetic testing at 60°.sec^{-1} compared with manual method proximally using the MicroFET2 had the smallest *ME* of ±54.65% and narrowest range of agreement of 193.25 *N·m* (*MD*= -5.36 *N·m*, 95% LoA [-101.99, 91.27], *p* =0.624) (Comb 33, Table 6.9; Figure 6.13). Also, isokinetic testing at 120°.sec^{-1} with the manual method proximally had the smallest *ME* of ±63.35% and narrowest range of 172.67, (*MD*= 13.47 *N·m*, 95% LoA [-72.87, 99.88], *p* =0.177) (Comb 37, Table 6.9; Figure 6.14). The results of limits of agreement tests are summarised in Table 6.9. For all Bland-Altman plots, refer to Appendix E. 5.

Key Results

- The smallest MD and ME were between isokinetic testing at 60°.sec^{-1} using the Biodex and measurements obtained by the MicroFET2.

Overall Key Results:

- Hip abduction testing: The smallest MD and ME were between isometric testing at 0° on the Biodex and the fixed MicroFET2 distally for the MicroFET2. They were also observed between isokinetic testing at 60°.sec^{-1} on the Biodex and the manual MicroFET2 distally.

- Hip adduction testing: The smallest MD and ME in were between isometric testing at 30° on the Biodex and the fixed MicroFET2 proximally and distally. Isokinetically, they were paired with testing at 60°.sec^{-1} in general.

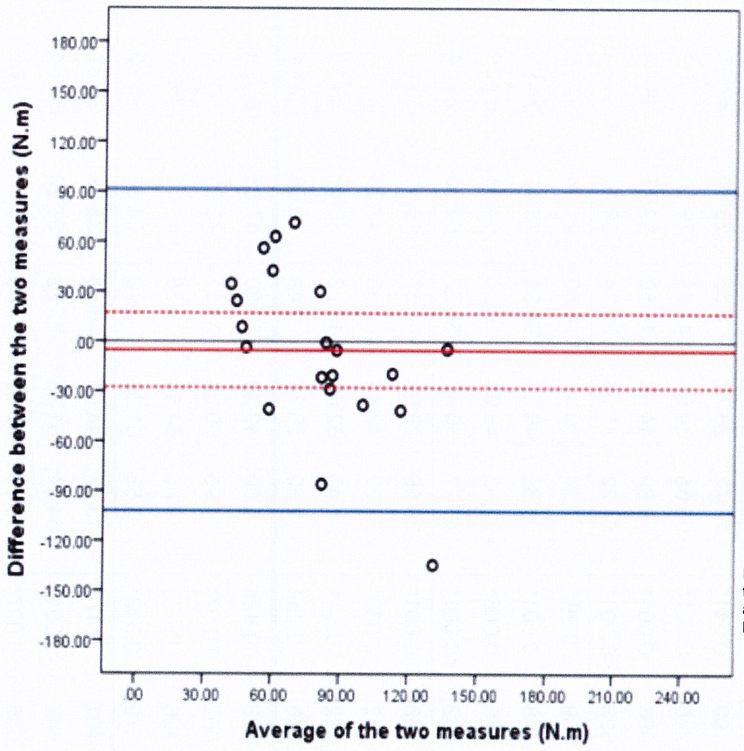


Figure 6.13 - Bland-Altman plot between adduction isokinetic at 60.sec^{-1} and adduction manual proximal in phase one (pair 33 from Table 6.9; Red line illustrates mean of difference; Dotted red lines illustrate 95% confidence interval of the mean of difference; blue lines illustrate agreement limits).

Note: Scale of the plot is based on the biggest difference (180 N·m) and highest average of scores (240 N·m) reported in both phases.

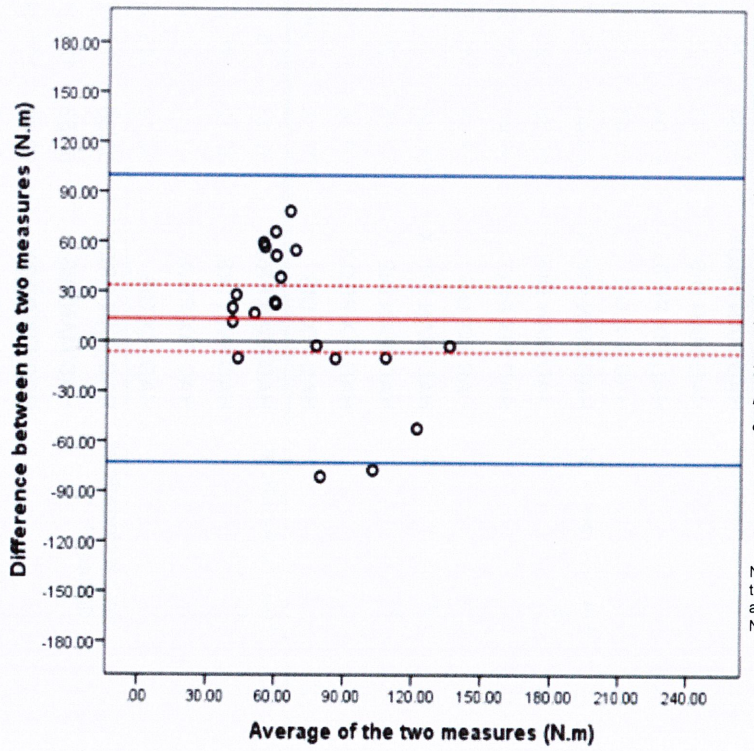


Figure 6.14 - Bland-Altman plot between adduction isokinetic at 120.sec^{-1} and adduction manual proximal in phase one (pair 37 from Table 6.9; Red line illustrates mean of difference; Dotted red lines illustrate 95% confidence interval of the mean of difference; blue lines illustrate agreement limits).

Note: Scale of the plot is based on the biggest difference (180 N·m) and highest average of scores (240 N·m) reported in both phases.

Table 6.9 - Limits of agreement tests between hip adduction measurements of the Biodex and the 1st MicroFET2 testing session in phase one (dominant side only).

Pair	Biodex	MicroFET2	Ms1 (*N·m*)	Ms2 (*N·m*)	Paired *t*-test *t*	Paired *t*-test *p*	Mean of Difference (Bias) Mean (*N·m*)	*SD* (±)	95% CI Lower	95% CI Upper	LLA (*N·m*)	ULA (*N·m*)	Limit range	Error (%)
21	HAD Isometric at 0°	HAD MAN PROX	63.04	76.77	2.190	0.041*	13.72	28.72	0.65	26.80	-42.57	70.02	112.59	41.48
22		HAD MAN DIS	61.43	125.07	8.248	<0.001*	63.65	34.51	47.49	79.80	-3.99	131.28	135.27	52.59
23		HAD FIX PROX	63.04	101.93	3.988	0.001*	38.89	44.68	18.55	59.23	-48.70	126.47	175.16	64.53
24		HAD FIX DIS	61.43	115.65	7.628	<0.001*	54.22	31.79	39.34	69.10	-8.09	116.53	124.61	48.44
25	HAD Isometric at 15°	HAD MAN PROX	69.40	76.77	1.109	0.281	7.37	30.47	-6.50	21.24	-52.35	67.09	119.44	39.97
26		HAD MAN DIS	68.38	125.07	5.448	0.001*	56.69	46.53	34.91	78.47	-34.51	147.89	182.41	63.70
27		HAD FIX PROX	69.40	101.93	3.046	0.006*	32.53	48.94	10.26	54.81	-63.39	128.45	191.84	64.20
28		HAD FIX DIS	68.38	115.65	5.802	<0.001*	47.27	36.43	30.22	64.31	-24.14	118.67	142.80	49.87
29	HAD Isometric at 30°	HAD MAN PROX	106.36	76.77	-4.755	<0.001*	-29.59	28.52	-42.57	-16.61	-85.49	26.31	111.79	24.41
30		HAD MAN DIS	102.85	125.07	2.342	0.030*	22.22	42.43	2.36	42.08	-60.94	105.38	166.32	38.61
31		HAD FIX PROX	106.36	101.93	-0.622	0.541	-4.43	32.63	-19.28	10.42	-68.38	59.52	127.90	27.93
32		HAD FIX DIS	102.85	115.65	2.016	0.731	12.80	28.38	-0.49	26.08	-42.84	68.43	111.26	25.83
33	HAD Isokinetic at 60°.sec^{-1}	HAD MAN PROX	82.13	76.77	-0.498	0.624	-5.36	49.30	-27.80	17.08	-101.99	91.27	193.25	54.65
34		HAD MAN DIS	80.12	125.07	3.002	0.0 07*	44.96	66.98	13.61	76.30	-86.32	176.23	262.56	78.26
35		HAD FIX PROX	82.13	101.93	1.406	0.175	19.80	64.54	-9.58	49.18	-106.70	146.30	253.01	71.54
36		HAD FIX DIS	80.12	115.65	2.579	0.018*	35.53	61.60	6.70	64.36	-85.20	156.26	241.47	71.97
37	HAD Isokinetic at 120°.sec^{-1}	HAD MAN PROX	63.30	76.77	1.401	0.177	13.47	44.05	-6.58	33.52	-72.87	99.80	172.67	63.35
38		HAD MAN DIS	60.83	125.07	4.643	<0.001*	64.24	61.88	35.28	93.20	-57.04	185.52	242.57	95.22
39		HAD FIX PROX	63.30	101.93	3.036	0.007*	38.63	58.30	12.09	65.17	-75.65	152.90	228.55	83.85
40		HAD FIX DIS	60.83	115.65	4.510	<0.001*	54.82	54.35	29.38	80.25	-51.71	161.34	213.06	83.63

Ms1, first measurement; Ms2, second measurement; M, mean; Prox., proximal; Dis., distal; HAB, hip abduction; HAD, hip adduction; MAN, manual stabilisation; FIX, fixed stabilisation, *M,* mean; *SD,* standard deviation; CI, confidence interval; LLA, lower limit of agreement; ULA, upper limit of agreement; *, significant difference between the two measurements at 0.05.

Table 6.10 - Pearson Correlation (r) between measurements of the Biodex and the 1st MicroFET2 session in phase one (dominant side only).

Ms1 - Biodex	Ms2 - HHD1	Pearson *r*	Sig. *p*
HAB Isometric at 0°	**HAB MAN PROX**	0.554	0.009*
	HAB MAN DIS	0.711	<0.001*
	HAB FIX PROX	0.512	0.018*
	HAB FIX DIS	0.757	<0.001*
HAB Isometric at 15°	**HAB MAN PROX**	0.584	0.005*
	HAB MAN DIS	0.724	<0.001*
	HAB FIX PROX	0.383	0.086
	HAB FIX DIS	0.707	<0.001*
HAB Isometric at 30°	**HAB MAN PROX**	0.631	0.002*
	HAB MAN DIS	0.751	<0.001*
	HAB FIX PROX	0.354	0.116
	HAB FIX DIS	0.705	<0.001*
HAB Isokinetic at 60°/sec	**HAB MAN PROX**	0.644	0.002*
	HAB MAN DIS	0.755	<0.001*
	HAB FIX PROX	0.365	0.104
	HAB FIX DIS	0.703	<0.001*
HAB Isokinetic at 120°/sec	**HAB MAN PROX**	0.535	0.013*
	HAB MAN DIS	0.666	0.001*
	HAB FIX PROX	0.283	0.213
	HAB FIX DIS	0.595	0.004*

Ms1 - Biodex	Ms2 - HHD1	Pearson *r*	Sig. *p*
HAD Isometric at 0°	**HAD MAN PROX**	0.517	0.016*
	HAD MAN DIS	0.328	0.146
	HAD FIX PROX	0.518	0.016*
	HAD FIX DIS	0.379	0.091
HAD Isometric at 15°	**HAD MAN PROX**	0.391	0.079
	HAD MAN DIS	0.052	0.822
	HAD FIX PROX	0.385	0.085
	HAD FIX DIS	0.311	0.169
HAD Isometric at 30°	**HAD MAN PROX**	0.653	0.001*
	HAD MAN DIS	0.036	0.877
	HAD FIX PROX	0.778	<0.001*
	HAD FIX DIS	0.302	0.183
HAD Isokinetic at 60°/sec	**HAD MAN PROX**	0.155	0.502
	HAD MAN DIS	-0.241	0.294
	HAD FIX PROX	0.149	0.520
	HAD FIX DIS	-0.114	0.623
HAD Isokinetic at 120°/sec	**HAD MAN PROX**	0.233	0.310
	HAD MAN DIS	-0.237	0.301
	HAD FIX PROX	0.254	0.267
	HAD FIX DIS	-0.051	0.827

Ms1, first measure; Ms2, second measure; Prox., proximal; Dis., distal; HAB, hip abduction; HAD, hip adduction; MAN, manual stabilisation; FIX, fixed stabilisation, *, significant correlation at 0.05.

6.2.10. Location of Application and Method of Stabilisation for the MicroFET2 (RQ 8)

The distal and proximal application sites of the MicroFET2 together with the two stabilisation methods were compared between the two testing sessions. There were significant increases in all measurements for the second session, supported by positive means of differences (mean of measurement in the 2nd session – mean of measurement in the 1st session) (Table 6.11).

In HAB testing, the highest difference in scores was 34.4% in using the manual method proximally (*MD*= 22.63 *N·m,* 95% LoA [-31.69, 76.95], *ME*= ±39.4%, *p* =0.002) (Pair 41, Table 6.11; Figure 6.15). On the other hand, the smallest difference of 11.4% was observed in using the fixed method distally (*MD*= 12.09 *N·m,* 95% LoA [-24.34, 48.52], *ME*= ±16.4%, *p* =0.009) (Pair 44, Table 6.11; Figure 6.16). The smallest coefficient of variation (CV) was noted in using the fixed method distally (11.30%), while the biggest CV was in using the manual method proximally (21.73%) (Table 6.12).

In terms of HAD testing, applying a manual method proximally had the highest difference in scores of 33.6% (*MD*= 25.77 *N·m,* 95% LoA [-24.89, 76.42], *ME*= ±31.5%, *p* <0.001) (Pair 45, Table 6.11; Figure 6.18). In contrast, the smallest difference of 12% was associated with the distal location using a fixed method (*MD*= 13.87 *N·m,* 95% LoA [-26.85, 54.60], *ME*= ±17.3%, *p* =0.009) (Pair 48, Table 6.11; Figure 6.17). The smallest CV was also associated with the manual (15.16%) and fixed methods (16.83%) distally (Table 6.12). For all Bland-Altman plots, refer to Appendix E. 6. For all torque values, refer to Appendix E. 7.

Key Results

- Measurements in the second MicroFET2 testing session showed statistically significant higher values than the first session.

- The smallest MEs of HAB and HAD testing were noted distally using the fixed method.

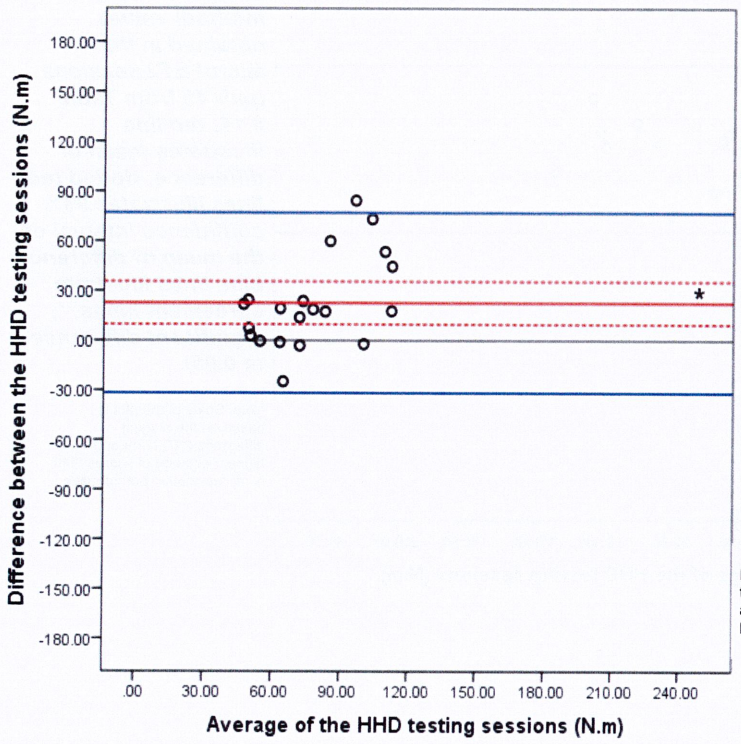


Figure 6.15 - Bland-Altman plot of abduction testing in phase one (proximal using the manual method) values obtained in the MicroFET2 sessions (pair 41 from Table 6.11; red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits; *, significant difference at 0.05).

Note: Scale of the plot is based on the biggest difference (180 N·m) and highest average of scores (240 N·m) reported in both phases.

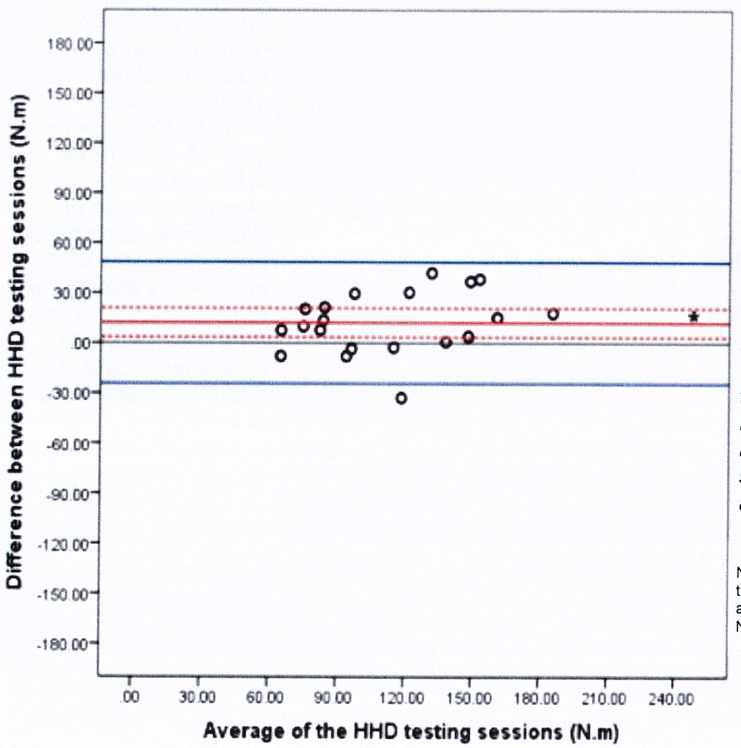


Figure 6.16 - Bland-Altman plot of abduction testing in phase one (distal using the fixed method) values obtained in the MicroFET2 sessions (pair 44 from Table 6.11; red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits; *, significant difference at 0.05).

Note: Scale of the plot is based on the biggest difference (180 N·m) and highest average of scores (240 N·m) reported in both phases.

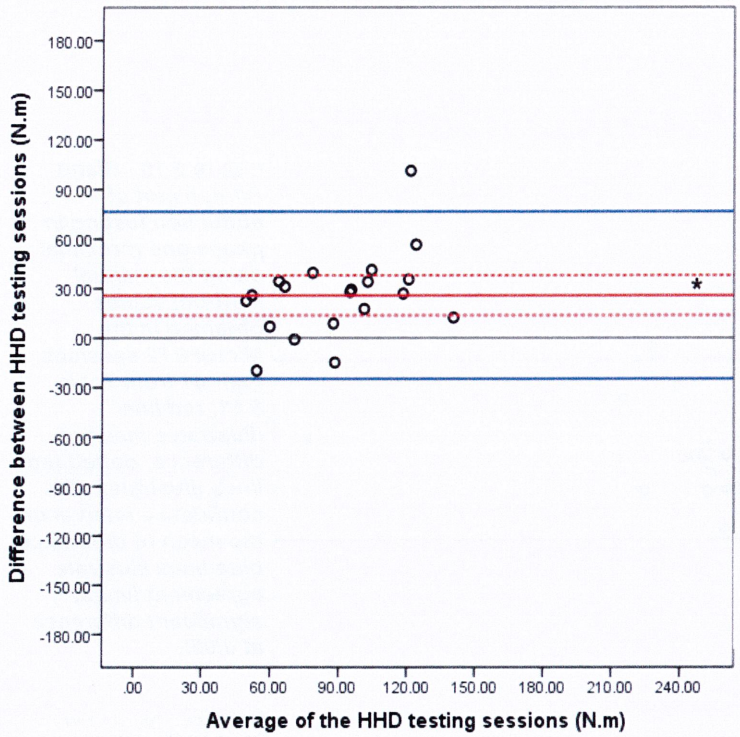


Figure 6.17 - Bland-Altman plot of adduction testing in phase one (proximal using the manual method) values obtained in the MicroFET2 sessions (pair 45 from Table 6.11; red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits; *, significant difference at 0.05).

Note: Scale of the plot is based on the biggest difference (180 N·m) and highest average of scores (240 N·m) reported in both phases.

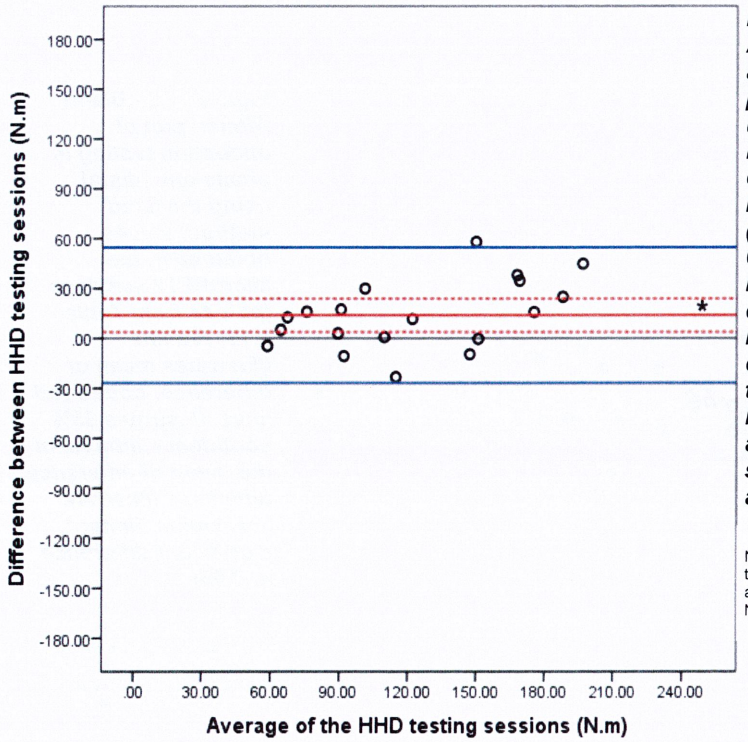


Figure 6.18 - Bland-Altman plot of adduction testing in phase one (distal using the fixed method) values obtained in the MicroFET2 sessions (pair 48 from Table 6.11; red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits; *, significant difference at 0.05).

Note: Scale of the plot is based on the biggest difference (180 N·m) and highest average of scores (240 N·m) reported in both phases.

Table 6.11 - Limits of agreement tests between the results of the two MicroFET2 sessions in phase one (dominant side only).

Pair	Measurement	S1 (*N·m*)	S2 (*N·m*)	Paired *t*-test *t*	Paired *t*-test *p*	Mean of Difference (Bias) Mean (*N·m*)	*SD* (±)	95% CI Lower	95% CI Upper	LLA (*N·m*)	ULA (*N·m*)	Limit range	Change (%)	Error (%)	CoR (*N·m*)
41	HAB MAN PROX	65.76	88.39	3.652	0.002*	22.63	27.71	9.66	35.60	-31.69	76.95	108.63	34.4	39.4	54.32
42	HAB MAN DIS	101.74	124.16	3.429	0.003*	22.43	29.24	8.74	36.11	-34.89	79.74	114.64	22.0	26.9	57.32
43	HAB FIX PROX	62.85	83.66	2.562	0.019*	20.82	36.34	3.81	37.82	-50.41	92.04	142.45	33.1	54.1	71.22
44	HAB FIX DIS	106.06	118.15	2.909	0.009*	12.09	18.59	3.39	20.79	-24.34	48.52	72.87	11.4	16.4	36.43
45	HAD MAN PROX	76.80	102.56	4.459	<0.001*	25.77	25.84	13.67	37.86	-24.89	76.42	101.30	33.6	31.5	50.65
46	HAD MAN DIS	124.56	139.80	2.603	0.018*	15.24 †	25.52	2.94	27.54	-34.78	65.27	100.05	12.2	19.8	50.02
47	HAD FIX PROX	101.15	132.62	2.596	0.018*	31.47	54.22	6.10	56.84	-74.79	137.73	212.53	31.1	50.2	106.26
48	HAD FIX DIS	115.61	129.48	2.911	0.009*	13.87 †	20.78	3.86	23.89	-26.85	54.60	81.45	12.0	17.3	40.72

S, session; M, mean; Prox., proximal; Dis., distal; HAB, hip abduction; HAD, hip adduction; MAN, manual stabilisation; FIX, fixed stabilisation, *N·m*, newton metres; mean; *SD*, standard deviation; CI, confidence interval; LLA, lower limit of agreement; ULA, upper limit of agreement; CoR, coefficient of repeatability; †, one participant was excluded due to incomplete dataset; *, significant difference between the two measurements at 0.05.

Table 6.12 – The Coefficient of Variation between the two MicroFET2 sessions in phase one (dominant side only).

Muscle	Measurement	CV
HAB	MAN PROX	21.73%
	MAN DIS	16.40%
	FIX PROX	25.35%
	FIX DIS	11.30%
HAD	MAN PROX	24.22%
	MAN DIS	15.16%
	FIX PROX	25.59%
	FIX DIS	16.83%

Prox., proximal; Dis., distal; HAB, hip abduction; HAD, hip adduction; MAN, manual stabilisation; FIX, fixed stabilisation; CV, coefficient of variation

6.2.11. Sensitivity Analysis for the Two MicroFET2 Sessions (RQ 8)

The sensitivity analysis using MLM was used to explain within- and between-participants' variances in scores. In the regression model, the grand mean for the measurement was represented by the intercept (β_{oj}). A positive regression coefficient in MLM indicates that the reference group or measurement had higher scores than the other group(s) and the vice versa. The ANOVA model showed within- and between-participants variation in scores.

The output from model 3 showed significant effects of performing two testing sessions over the measurements obtained in the MicroFET2 testing sessions (Table 6.14).

- Proximal location of placement: the results showed an increase of 22.63 *N·m* (Standard error [SE]= 6.03 *N·m*) for HAB manual (p =0.001), 20.82 *N·m* (SE= 7.92 *N·m*) for HAB fixed (p =0.016), 25.77 *N·m* (SE= 5.63 *N·m*) for HAD manual (p <0.001), and 31.47 *N·m* (SE= 11.81 *N·m*) for HAD fixed (p =0.015) (Table 6.14).
- Distal location of stabilisation: the increase was 22.43 *N·m* (SE= 6.37 *N·m*) for HAB manual (p =0.002), 12.09 *N·m* (SE= 4.05 *N·m*) for HAB fixed (p =0.007), 23.35 *N·m* (SE= 9.58 *N·m*) for HAD manual (p =0.024), and 21.87 *N·m* (SE= 8.95 *N·m*) for HAD fixed (p =0.024) (Table 6.14). The change in scores was the smallest when the fixed method was applied distally for both muscle groups (Table 6.14).

Sex also had a significant effect in explaining the variation. The female participants scored less than the males in the second MicroFET2 session, which was supported by negative changes in the scores of all measurements:

- Abduction: the biggest difference in scores between males and females of -46.67 *N·m* (SE= 11.43 *N·m*) was observed when the fixed method was applied distally (p =0.001) (Table 6.14).
- Adduction: the biggest difference in scores between males and females of -90.69 *N·m* (SE= 17.32 *N·m*) and -67.48 *N·m* (SE= 11.93 *N·m*) were observed when the fixed method was applied proximally and distally (in consecutive order, both p <0.001) (Table 6.14).

Level 1 (values of the two sessions, within-participants) and level 2 (between-participants) variances were not consistent. Overall, the level 1 variances varied within-participants significantly in all measurements, indicating considerable variation between the scores of the two MicroFET2 sessions. On the other hand, there were no significant differences in level 2, except in HAB testing (fixed method) on a distal location. The insignificant variance of scores indicates that the scores between the participants were very similar (in general). A summary of the sensitivity analysis results can be found in Table 6.13, while Table 6.14 shows detailed regression coefficient results of model 3, the intercept or grand mean (β_{oj}) and the changes from Model 1.

Table 6.13 - Summary of the effects of factors on abduction and adduction strength measurements of the second MicroFET2 session in phase one (based on Model 3).

	Abduction Manual		**Abduction Fixed**		**Adduction Manual**		**Adduction Fixed**	
	Prox.	Dist.	Prox.	Dist.	Prox.	Dist.	Prox.	Dist.
Second session [time]	↑	↑	↑	↑	↑	↑	↑	↑
Females [sex]	↓	↓	↓	↓	↓	↓	↓	↓

Prox., proximal; Dist., distal.

Key Results

- Having two MicroFET2 testing sessions (number of days between the sessions) and sex significantly explained the variation in the measurements obtained proximally and distally for both methods of stabilisation.
- The smallest change in scores was noticed when the fixed method was used distally.
- The female participants scored less than males in the second session. Also, applying the fixed method distally in HAB and HAD testing, and proximally in HAD testing only significantly increased the difference in scores between males and females.

Table 6.14 – Results of the sensitivity analysis using multilevel modelling based on Model 3 showing the regression coefficients (and standard error) to estimate trends in means of measurement scores between the two dynamometer sessions.

	Abduction Manual (*N·m*)		Abduction Fixed (*N·m*)		Adduction Manual (*N·m*)		Adduction Fixed (*N·m*)	
	Proximal (SE)	**Distal (SE)**	**Proximal (SE)**	**Distal (SE)**	**Proximal (SE)**	**Distal (SE)**	**Proximal (SE)**	**Distal (SE)**
Model 3 Parameters								
df (-2LL) of Model 3	6 (355.5)	6 (366.9)	6 (375.7)	6 (358.8)	6 (360.3)	6 (395.8)	6 (418.3)	6 (392.2)
df and χ^2 Change from Model 1	2 (23.7)*	2 (24.1)*	2 (20.3)*	2 (19.5)*	2 (27.7)*	2 (20)*	2 (23.3)*	2 (24.4)*
Fixed Part (Regression model)								
β_{oj} (intercept)	87.04 (14.11)***	138.61 (16.66)***	101.4 (18.09)***	161.4 (18.55)***	104.9 (15.57)***	177.7 (23.57)***	201.18 (31.93)***	185.81 (22.68)***
Time [in 2nd session]	22.63 (6.03)**	22.43 (6.37)**	20.82 (7.92)*	12.09 (4.05)**	25.77 (5.63)***	23.35 (9.58)*	31.47 (11.81)**	21.87 (8.95)**
Sex [in females]	-30.29 (7.06)***	-40.89 (8.90)**	-40.96 (8.90)***	-46.47 (11.43)**	-37.12 (8.53)***	-57.05 (12.18)***	-90.69 (17.32)***	-67.48 (11.92)***
Random Part (ANOVA model)								
Within-Var. (ε_{oj})	364.8 (115.4)**	406.2 (128.5)**	627.2 (198.3)**	164.14 (51.91)**	317.23 (100.3)**	918.4 (290.4)**	1398.2 (441.5)**	801.0 (253.3)**
Between-Var. (μ_{oj})	64.51 (97.07)	188.7 (139.6)	78.79 (158.9)	565.0 (206.3)**	201.9 (124.6)	275.2 (273.9)	787.3 (519.0)	303.4 (256.1)

N·m, Newton Metre; SE, standard error; β, intercept representing the grand mean; Time, two sessions; Within-Var. (ε_{oj}) = within participants variance; Between-Var. (μ_{oj}) = between participants variance; significant increase or decrease or significant variation in scores at p <0.05, *; p <0.01, **; p <0.001, ***; Note 1: The squared brackets contain the reference group; Note 2: the intercept and coefficients are based on Model 3.

6.3. *Phase Two Results*

6.3.1. Descriptive Results

Twenty-one participants were tested in the period between the 28th of September 2016 and the 8th of November, 2016. Two participants did not attend their second testing session and their results were excluded. One participant informed the tester that he purposely manipulated his caloric intake and training regime to score higher in the second testing session. Thus, his results were also excluded because he did not follow the instructions.

The remaining eighteen participants were (*mean age* ± *SD*) 19.7 ± 1.7 years old and consisted of eleven males (mean age 20.1 ± 1.9 years) and seven females (mean age 19.0 ± 1.0 years). The male participants were taller (mean stature 1.76 ± 0.07 *m*) and had higher body mass (mean body mass 76.10 ± 8.81 *kg*) compared to the females (mean stature 1.67 ± 0.06 *m*; mean body mass 61.02 ± 10.28 *kg*). The right side was the dominant side in most of the participants (*n*=15) compared to the left side (*n*=3). The average duration between the two testing sessions was 6.9 ± 1.2 days (Table 6.15).

All participants were physically active. Football was the most played sport (*n*=4), followed by athletics (*n*=3), fitness training (*n*=2), rugby (*n*=2) and swimming (*n*=2). Volleyball, boxing, CrossFit, lacrosse, and netball were played by other participants (*n*=1 each) (Table 6.16).

Table 6.15 - Baseline characteristics of the participants in phase two.

Parameters		Males	Females	Overall
Age (years)		20.1 ± 1.9	19.0 ± 1.0	19.7 ± 1.7
Stature (*m*)		1.76 ± 0.07	1.67 ± 0.06	1.73 ± 0.08
Body mass (*kg*)		76.10 ± 8.81	61.02 ± 10.28	70.24 ± 11.84
Body Mass Index (BMI, $kg.m^{-2}$)		24.56 ± 2.87	21.85 ± 3.59	23.51 ± 3.36
Days between sessions		6.5 ± 1.2	7.4 ± 2.1	6.9 ± 1.2
Dominant Side (*n*)	Right	10	5	15
	Left	1	2	3
Total (*n*)		11	7	18

Values are presented as mean ± standard deviation (*M* ± *SD*)

Table 6.16 - Sports played by the participants in phase two.

Sport	*n*
Athletics	3
Fitness training (gym workouts)	2
Football (soccer)	4
Rugby	2
Swimming	2
Volleyball, boxing, CrossFit, lacrosse, and netball	1 each
Total (*n*)	18

6.3.2. Sides Comparison from MicroFET2 Testing (RQs 1 and 2)

The testing results of the dominant (DS) and non-dominant (NS) sides were analysed*. The results of paired *t*-test showed no statistically significant differences between both sides in the first (p =0.944) and second sessions (p =0.873). In the first session, the highest difference of 7.1% was observed in the hip flexion (HFL) muscle group. The hip external rotation (HER) muscle group had the second highest difference with the NS showing 6.5% more torque compared to the DS. In the second testing session, the hip adductor (HAD) and HFL muscle groups had the highest torque deficits of 4.4% and 4.3% consecutively (Figure 6.19 and Figure 6.20).

The hip abductor (HAB) and hip internal rotation (HIR) muscles had the least bilateral torque deficits. The deficit was equal to 1.2% and 0.3% respectively in the first session showing higher values on the NS side. However, the NS demonstrated reduced torques of 1.2% and 0.8% consecutively in the second session (Figure 6.20 and Figure 6.21).

Muscle torque was also analysed based on the plane of motion. In the sagittal plane, knee flexion (KFL) muscles showed the highest torque levels followed by the HFL and hip extension (HEX) muscles. The NS side of the KFL muscles had higher torque than

* Difference or strength deficit= strength on the NS – strength on the DS

the DS side. Opposite findings were observed in the HFL and HEX muscles (Figure 6.19).

In the frontal plane, the HAB muscles had higher torque values than the HAD muscle. The deficit between the DS and the NS sides in the HAB muscles was minimal. However, the HAB demonstrated higher torque deficit of 4.4% in the second session (Figure 6.20). In the horizontal plane, the HIR muscles had higher torque than the HER muscles with reduced deficit (Figure 6.21). The *t*-test results and muscle torque values can be found in Appendix E. 8.

Key Results

- From MicroFET2 testing, there were no statistical differences between the DS and NS.
- HFL muscles testing had the biggest bilateral torque deficits, while the HIR and HAB muscles showed the smallest bilateral torque deficits from MicroFET2 testing.

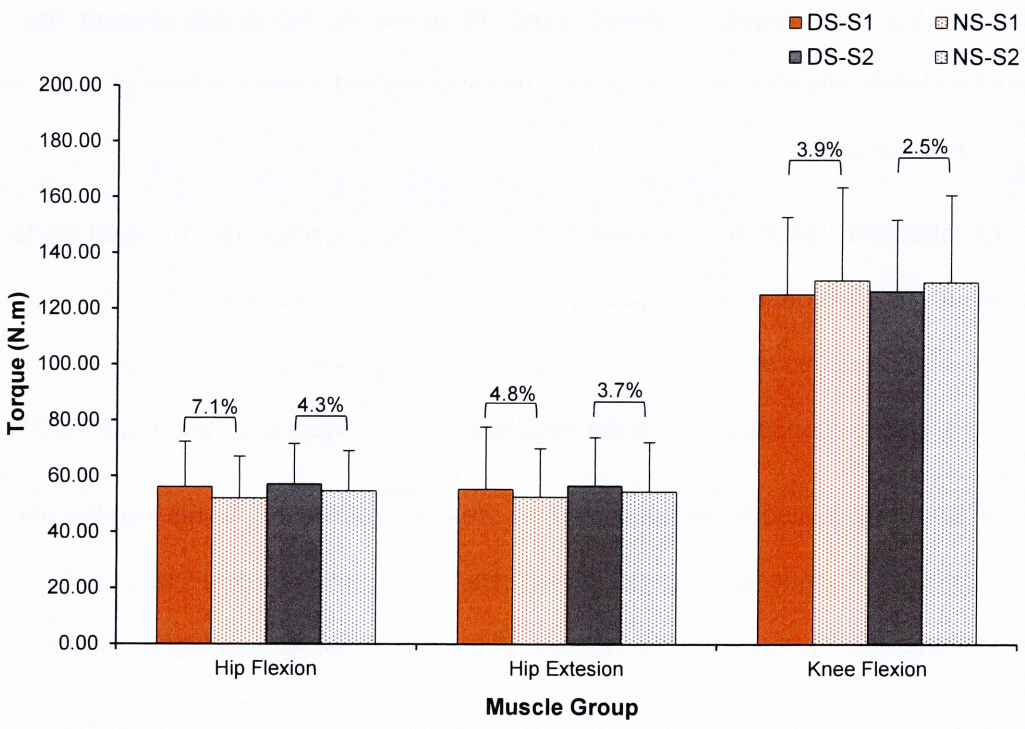


Figure 6.19 – Torque values of the hip extensor, hip and knee flexor muscles in the sagittal plane in phase two from MicroFET2 testing (with percentage of difference and standard deviation bars; DS, dominant side; NS, non-dominant side; S1, S2 first and second sessions).

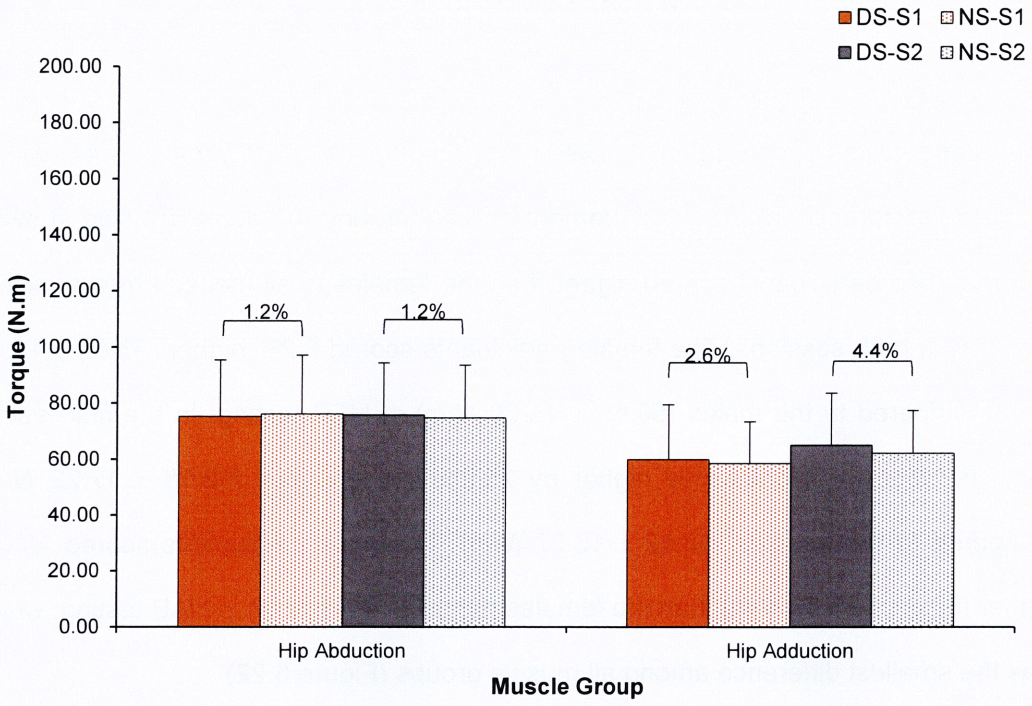


Figure 6.20 - Torque values of the hip abductor and adductor muscles in the frontal plane in phase two from MicroFET2 testing (with percentage of difference and standard deviation bars; DS, dominant side; NS, non-dominant side; S1, S2 first and second sessions).

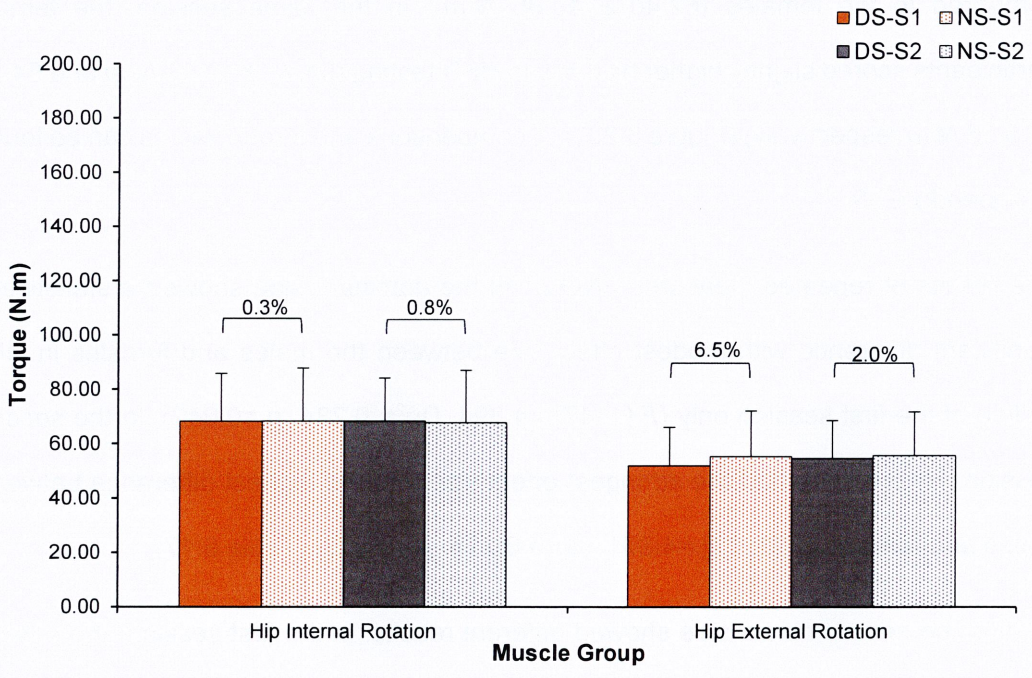


Figure 6.21 - Torque values of the hip internal and external rotator muscles in the horizontal plane in phase two from MicroFET2 testing (with percentage of difference and standard deviation bars; DS, dominant side; NS, non-dominant side; S1, S2 first and second sessions).

6.3.3. *Sex Comparison from MicroFET2 Testing (RQ 2)*

The torque values varied greatly between the males (n=11) and females (n=7) with a significant difference in one measurement only. Analysing the dominant side showed that the male participants scored higher than the females in all measurements except the HER in both sessions. The female participants scored 6.2% higher (53.80 ± 12.67 *N·m*) compared to the males (50.65 ± 15.69 *N·m*) in HER testing. In the first testing session, HFL was significantly higher by 25.2% in the males (62.18 ± 17.22 *N·m)* compared to the females (58.52 ± 15.27 *N·m)*. The male participants scored 10.0% higher (62.31 ± 23.59 *N·m)* than the females (56.10 ± 10.79 *N·m*) in HAB testing, which was the smallest difference among all muscle groups (Figure 6.22).

The variation between males and females was not the same in the second session. The biggest difference was observed in HAB, in which the males (82.40 ± 17.77 *N·m*) scored 21.0% higher than the females (65.10 ± 15.97 *N·m*). HIR testing showed the smallest difference with 13.4% higher values obtained by the males (72.03 ± 18.19 *N·m*) compared to the females (62.40 ± 10.05 *N·m*). In the same session, the female participants scored slightly higher by 0.5% in HER testing (54.79 ± 12.11 *N·m* and 54.54 ± 15.56 *N·m* respectively) (Figure 6.23). All comparisons and torque values can be found in Appendix E. 9.

The results of repeated measures ANOVA of the dominant side showed a statistically significant difference with modest effect size between the males and females in HFL testing of the first session only (F [1, 17] =4.894, ηp^2= 0.234, p =0.042). In the second session, HAB testing had the strongest effect size and the highest difference between males and females (F [1, 17] =4.371, ηp^2= 0.215, p =0.053) (Table 6.17).

Testing the non-dominant side showed different results. In the first session, there were statistically significant differences with modest effect size in HFL (F [1, 17] =7.106, ηp^2= 0.308, p =0.017) and HIR testing (F [1, 17] =5.094, ηp^2= 0.241, p =0.038). In the second

session, similar findings were observed in HFL (F [1, 17] =5.983, ηp^2= 0.272, p =0.026), HEX (F [1, 17] =5.013, ηp^2= 0.239, p =0.040) and HAD testing (F [1, 17] =4.637, ηp^2= 0.225, p =0.047) (Table 6.18).

Key Results

- Males are stronger than females in all measurements except HER in MicroFET2 testing.
- From MicroFET2 testing, the biggest difference between males and females was observed in HFL and HAB testing of the DS. The results of NS testing were inconsistent between the first and second testing session.

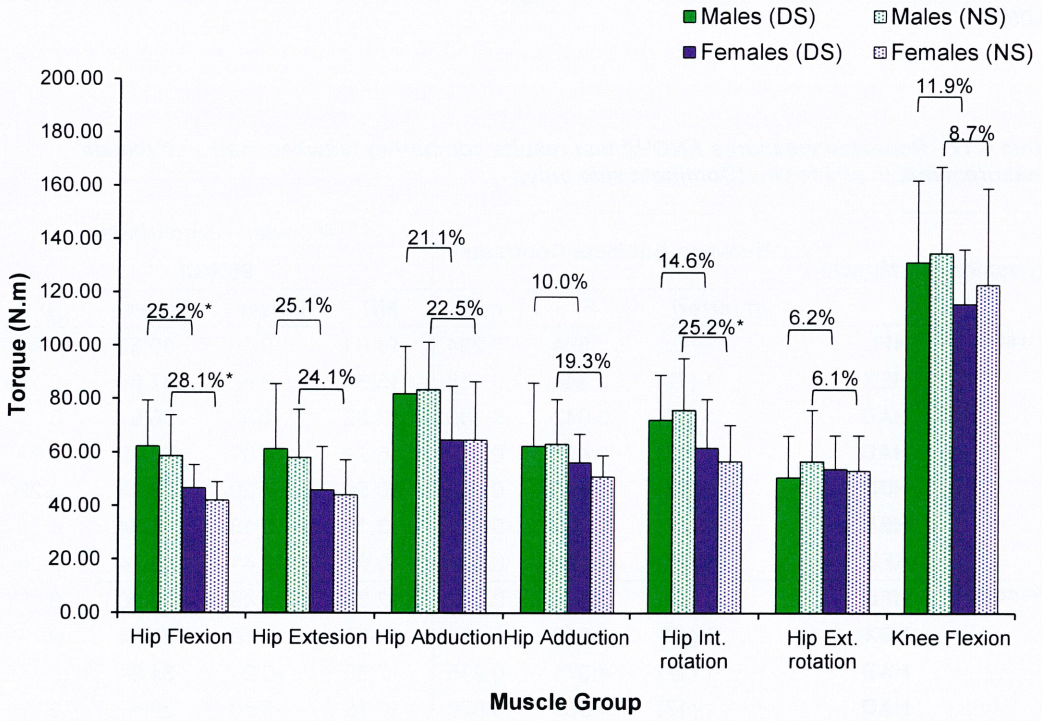


Figure 6.22 - Torque values based on gender and side as measured in the first session in phase two from MicroFET2 testing (with percentage of difference and standard deviation bars; DS, dominant side; NS, non-dominant side; *, significant difference between males and females at 0.05).

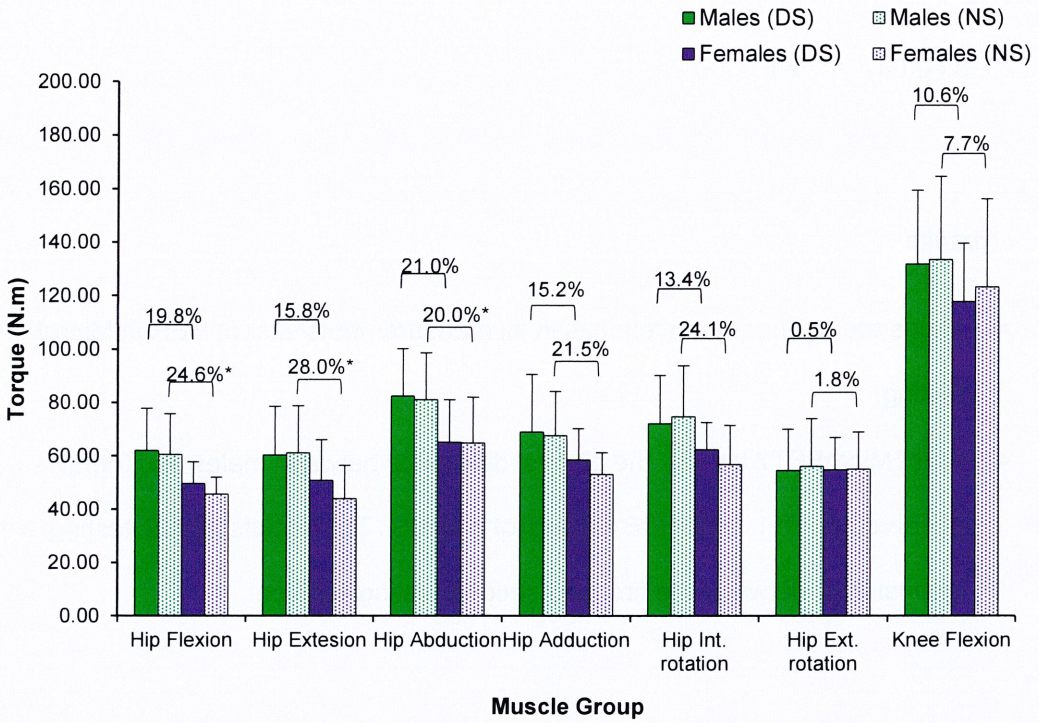


Figure 6.23 - Torque values based on gender and side as measured in the second session in phase two from MicroFET2 testing (with percentage of difference and standard deviation bars; DS, dominant side; NS, non-dominant side; *, significant difference between males and females at 0.05).

Table 6.17 - Repeated measures ANOVA test results comparing between male and female measurements in phase two (Dominant side only).

Session	Muscle	Between-Subjects Contrasts			Pairwise Comparisons 95% CI			
		df (error)	*F*	ηp²	*MD*	Lower	Upper	*p*
First	**HFL**	1 (17)	4.894	0.234	15.64	0.65	30.62	0.042*
	HEX	1 (17)	2.144	0.118	15.34	-6.87	37.56	0.163
	HAB	1 (17)	3.642	0.185	17.32	-1.92	36.57	0.074
	HAD	1 (17)	0.421	0.026	6.21	-14.07	26.49	0.525
	HIR	1 (17)	1.586	0.090	10.54	-7.20	28.28	0.226
	HER	1 (17)	0.199	0.012	-3.15	-18.15	11.84	0.662
	KFL	1 (17)	1.400	0.080	15.67	-12.41	43.74	0.254
Second	**HFL**	1 (17)	3.574	0.183	12.30	-1.49	26.08	0.077
	HEX	1 (17)	1.312	0.076	9.52	-8.10	27.14	0.269
	HAB	1 (17)	4.371	0.215	17.30	-0.24	34.84	0.053
	HAD	1 (17)	1.367	0.079	10.45	-8.50	29.41	0.259
	HIR	1 (17)	1.620	0.092	9.63	-6.41	25.66	0.221
	HER	1 (17)	0.001	0.000	-0.25	-14.97	14.47	0.972
	KFL	1 (17)	1.275	0.074	13.95	-12.23	40.13	0.275

HFL, hip flexion; HEX, hip extension; HAB, hip abduction; HAD, hip adduction; HIR, hip internal rotation; HER, hip external rotation; KFL, knee flexion; *MD, m*ean of difference; df, degrees of freedom; ηp², partial eta squared; *, significant difference between males and females at 0.05.

Table 6.18 - Repeated measures ANOVA test results comparing between male and female measurements in phase two (Non-dominant side only).

Session	Muscle	Between-Subjects Contrasts			Pairwise Comparisons 95% CI			
		df (error)	*F*	ηp^2	*MD*	Lower	Upper	*p*
First	**HFL**	1 (17)	7.106	0.308	16.46	3.37	29.55	0.017*
	HEX	1 (17)	3.124	0.163	13.96	-2.78	30.70	0.096
	HAB	1 (17)	4.015	0.201	18.79	-1.09	38.66	0.062
	HAD	1 (17)	3.189	0.166	12.18	-2.28	26.63	0.093
	HIR	1 (17)	5.094	0.241	19.07	1.16	36.99	0.038*
	HER	1 (17)	0.170	0.011	3.45	-14.31	21.21	0.686
	KFL	1 (17)	0.512	0.031	11.75	-23.06	46.57	0.484
Second	**HFL**	1 (17)	5.983	0.272	14.91	1.99	27.84	0.026*
	HEX	1 (17)	5.013	0.239	17.14	0.91	33.37	0.040*
	HAB	1 (17)	3.712	0.188	16.20	-1.62	34.02	0.072
	HAD	1 (17)	4.637	0.225	14.52	0.23	28.81	0.047*
	HIR	1 (17)	4.466	0.218	17.98	-0.06	36.02	0.051
	HER	1 (17)	0.017	0.001	1.03	-15.85	17.92	0.898
	KFL	1 (17)	0.450	0.027	10.30	-22.25	42.85	0.512

HFL, hip flexion; HEX, hip extension; HAB, hip abduction; HAD, hip adduction; HIR, hip internal rotation; HER, hip external rotation; KFL, knee flexion; *MD, m*ean of difference; *df*, degrees of freedom; ηp^2, partial eta squared; *, significant difference between males and females at 0.05.

6.3.4. Muscle Strength Ratios from MicroFET2 Testing (RQ 5)

The muscle strength ratios of both testing sessions were similar. HEX to HFL ratio was 0.98 ± 0.22 and 0.99 ± 0.16 in the first and second sessions respectively. These two ratios had the lowest coefficients of variation among all muscle groups (Table 6.19).

HAD to HAB ratio increased from 0.82 ± 0.23 to 0.87 ± 0.19, but it was paired with a reduction in the coefficient of variation. HIR to HER ratio changed from 1.36 ± 0.39 to 1.28 ± 0.29 indicating that the internal rotator muscles are stronger than the external rotator muscles (Table 6.19).

The ratio between the KFL muscles to the hip extensors was also calculated. The ratios showed that the KFL (i.e. hamstring muscle) in a flexed position could produce twice as much force as the HEX (i.e. gluteus maximus assisted by the hamstring) in an extended position from side-lying. The ratio slightly reduced from 2.44 ± 0.61 to 2.32 ± 0.47 in the second session (Table 6.19).

Key Results

- Similar strength ratios were obtained in the two MicroFET2 sessions.
- The HIR muscles were stronger than the HER muscles, and the KFL muscles were twice as strong as the HEX muscles in MicroFET2 testing.

Table 6.19 - Ratios of muscle strength obtained by the MicroFET2 in phase two (dominant side only).

Session	Ratio	*M*	*SD*
First	HEX / HFL	0.98	0.22
	HAD / HAB	0.82	0.23
	HIR / HER	1.36	0.39
	KFL / HEX	2.44	0.61
Second	HEX / HFL	0.99	0.16
	HAD / HAB	0.87	0.19
	HIR / HER	1.28	0.29
	KFL / HEX	2.32	0.47

HFL, flexion; HEX, hip extension; HAB, hip abduction; HAD, hip adduction; HIR, hip internal rotation; HER, hip external rotation; KFL, knee flexion; *M, m*ean; *SD,* standard deviation.

6.3.5. Test-retest Results from MicroFET2 Testing (RQ 8)

The limits of agreement tests indicated a minimal increase in the measurements of the second session. The change in muscle torque values ranged between 0.16 and 4.99 *N·m*.

The smallest change of scores was 0.23%, which was observed in HIR muscle testing (*MD*= 0.16 *N·m*, 95% LoA [-14.05, 14.37], *ME*= ±10.59%, *p* =0.929) (Figure 6.24, Table 6.20). HAB muscle testing had the second smallest change of 0.67% (*MD*= 0.50 *N·m*, 95% LoA [-16.75, 117.75], *ME*= ±11.65%, *p* =0.812) (Figure 6.25, Table 6.20). In contrast, HAD muscle testing showed a significant change of 8.34%, which was also the biggest change in scores (*MD*= 4.99 *N·m*, 95% LoA [-5.34, 115.32], *ME*= ±8.74%, *p* =0.001) (Figure 6.26, Table 6.20).

The KFL muscle tests showed the highest measurement error of ±14.56%. Even though, the percentage of change in scores was only 0.86% (*MD*= 1.08 *N·m*, 95% LoA [-34.85, 37.01], *p* =0.806) (Figure 6.27, Table 6.20). All limits of agreement tests can be found in Table 6.20, and all Bland-Altman plots can be found in Appendix E. 10.

Key Results

- In MicroFET2 testing, the HIR and HAB testing showed the smallest change in scores, while HAD testing showed the biggest change in scores.
- The KFL testing had the highest measurement error in MicroFET2 testing.

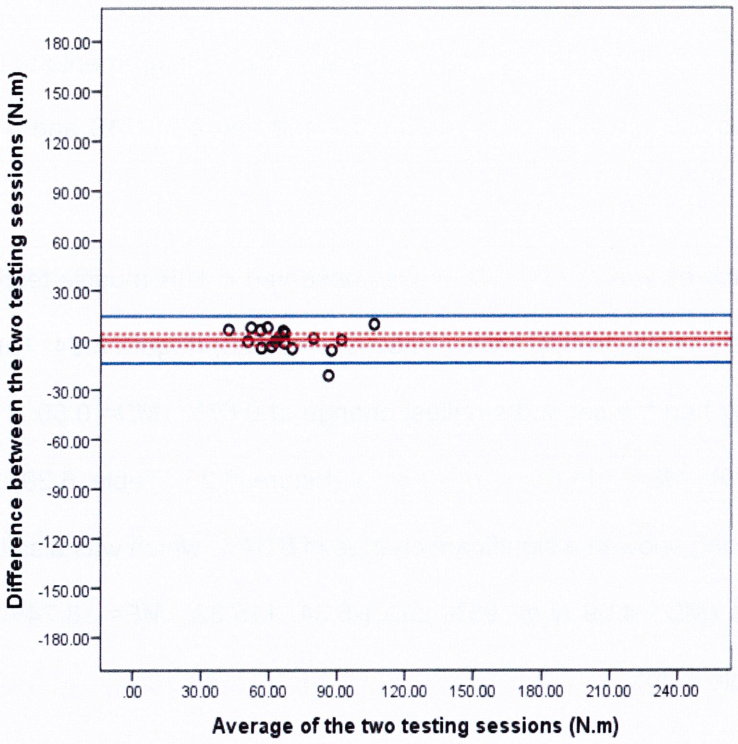


Figure 6.24 - Bland-Altman plot of hip internal rotation testing values obtained in the two MicroFET2 sessions in phase two (pair 5 from Table 6.20; red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits).

Note: Scale of the plot is based on the biggest difference (180 N·m) and highest average of scores (240 N·m) reported in both phases.

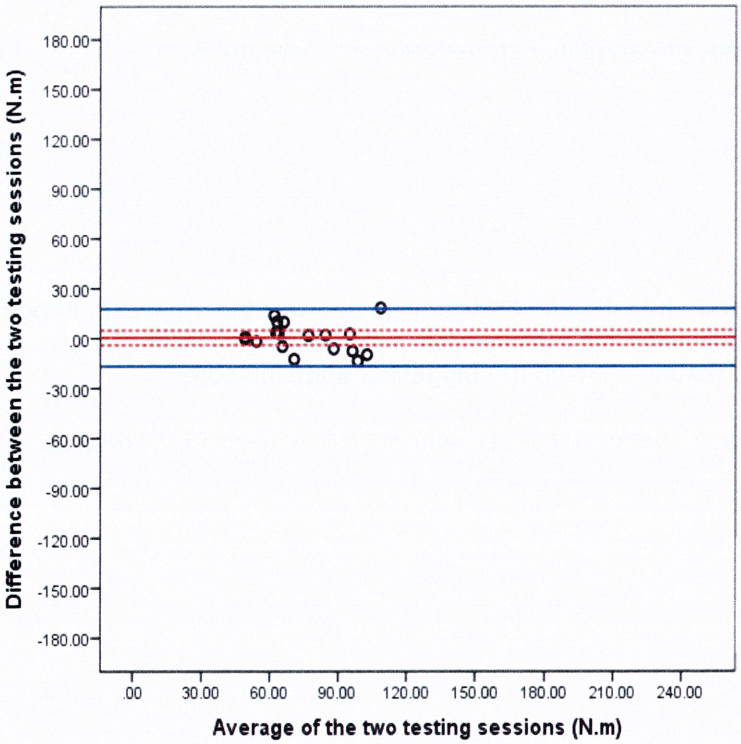


Figure 6.25 - Bland-Altman plot of hip abduction testing values obtained in the two MicroFET2 sessions in phase two (pair 3 from Table 6.20; red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits).

Note: Scale of the plot is based on the biggest difference (180 N·m) and highest average of scores (240 N·m) reported in both phases.

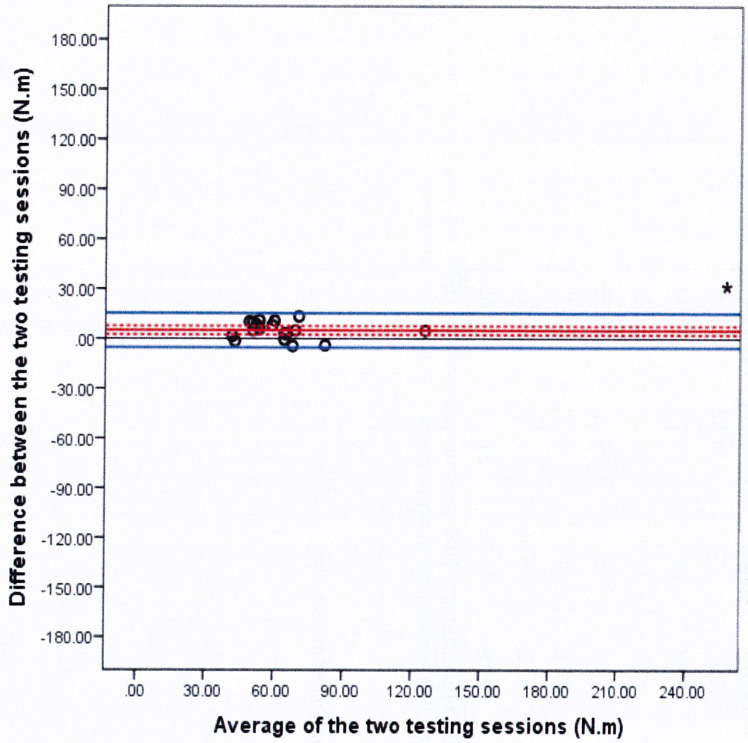


Figure 6.26 - Bland-Altman plot of hip adduction testing values obtained in the two MicroFET2 sessions in phase two (pair 4 from Table 6.20; red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits; *, significant difference at 0.05).

Note: Scale of the plot is based on the biggest difference (180 N·m) and highest average of scores (240 N·m) reported in both phases.

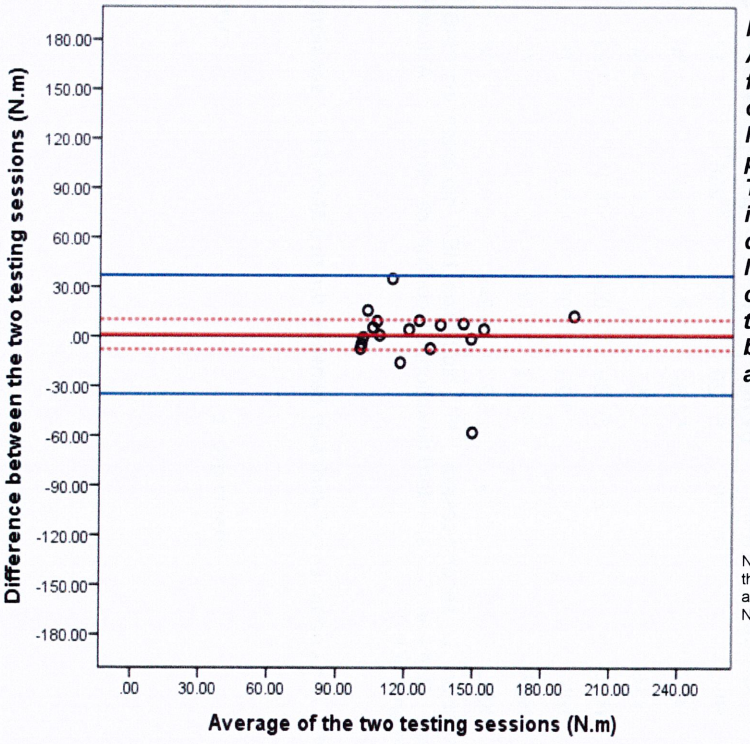


Figure 6.27 - Bland-Altman plot of knee flexor testing values obtained in the two MicroFET2 sessions in phase two (pair 7 from Table 6.20; red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits).

Note: Scale of the plot is based on the biggest difference (180 N·m) and highest average of scores (240 N·m) reported in both phases.

Table 6.20 - Limits of agreement tests between the results of the two MicroFET2 testing sessions of phase two (dominant side only).

Pair	Muscle Group	S1 (*N·m*)	S2 (*N·m*)	Paired *t*-test		Mean of Difference (Bias)							Change (%)	Error (%)	CoR (*N·m*)
						Mean (*N·m*)	*SD* (±)	95% CI		LLA (*N·m*)	ULA (*N·m*)	Limit range			
				t	*p*			Lower	Upper						
1	HFL	56.10	57.20	1.224	0.238	1.10	3.81	-0.80	3.00	-6.37	8.57	14.94	1.96	6.76	7.47
2	HEX	55.23	56.59	0.814	0.427	1.36	7.07	-2.16	4.87	-12.50	15.22	27.71	2.45	12.73	13.86
3	HAB	75.17	75.67	0.241	0.812	0.50	8.80	-3.88	4.88	-16.75	17.75	34.50	0.67	11.65	17.25
4	HAD	59.89	64.89	4.024	0.001*	4.99	5.27	2.38	7.61	-5.34	15.32	20.66	8.34	8.74	10.33
5	HIR	68.13	68.28	0.091	0.929	0.16	7.25	-3.45	3.76	-14.05	14.37	28.42	0.23	10.59	14.21
6	HER	51.87	54.63	2.065	0.055	2.76	5.67	-0.06	5.58	-8.35	13.87	22.23	5.32	10.88	11.11
7	KFL	125.19	126.27	0.249	0.806	1.08	18.33	-8.04	10.19	-34.85	37.01	71.85	0.86	14.56	35.93

S1, first session; S2, second session; HFL, hip flexion; HEX, hip extension; HAB, hip abduction; HAD, hip adduction; HIR, hip internal rotation; HER, hip external rotation; KFL, knee flexion; *M,* mean; *SD,* standard deviation; CI, confidence interval; LLA, lower limit of agreement; ULA, upper limit of agreement; *, significant difference between the two measurements at 0.05.

Table 6.21 – The coefficient of variation between the two MicroFET2 sessions in phase two (dominant side only).

Muscle group	CV%
HFL	9.13%
HEX	9.69%
HAB	7.14%
HAD	9.77%
HIR	10.04%
HER	8.82%
KFL	8.58%

CV, coefficient of variation; HFL, flexion; HEX, hip extension; HAB, hip abduction; HAD, hip adduction; HIR, hip internal rotation; HER, hip external rotation; KFL, knee flexion.

6.4. *Combined Results of Phases One and Two of Testing Using the MicroFET2 in Testing the Hip Abductor and Adductor Muscles (RQ 9)*

The HAB and HAD muscles were tested using the GrAASP in phase one and the STAPH in phase two. The findings of *t*-test showed that the participants of phase one study were significantly older (mean age 22.9 ± 4.4 years) compared to phase two participants (19.7 ± 1.7 years, p =0.005). Additionally, the number of days between the two MicroFET2 sessions was significantly higher in phase one (23.8 ± 9.0 days) in comparison to phase two (6.9 ± 1.2 days, p <0.001). Further analysis showed no significant differences in body measurements, sex or the dominant side between the two phases (p =0.974).

In testing the HAB, the measurements obtained by the STAPH had narrower limits of agreement and smaller error (*MD*= 0.50 *N·m,* 95% LoA [-16.75, 17.75], *ME*= ±11.65%, p =0.812) compared to the measurements obtained by the GrAASP (*MD*= 12.09 *N·m,* 95% LoA [-24.34, 48.52], *ME*= ±16.41%, p =0.009) (Figure 6.28; Table 6.23).

The results of HAD testing were similar. Testing using the STAPH showed reduced differences in scores and measurement error (*MD*= 4.99 *N·m,* 95% LoA [-5.34, 15.32], *ME*= ±8.74%, p =0.001). In contrast, testing using the GrAASP was characterised with higher means of differences and bigger measurement error (*MD*= 13.87 *N·m,* 95% LoA [-26.85, 54.60], *ME*= ±17.32%, p =0.009) (Figure 6.29; Table 6.23).

Key Results

- The participants and testing protocols of the two phases were similar except in age and in the number of days between the two MicroFET2 testing sessions.
- The STAPH is more reliable with reduced measurements error and differences in scores compared to the GrAASP.

Table 6.22 - Baseline characteristics of the participants in phases one and two.

Parameters		**GrAASP**	**STAPH**	***p***
Age (years)		22.9 ± 4.4	19.7 ± 1.7	0.005*
Stature (*m*)		1.73 ± 0.07	1.73 ± 0.08	0.818
Body mass (*kg*)		70.12 ± 11.26	70.24 ± 11.84	0.974
Body Mass Index (BMI, $kg.m^{-2}$)		23.27 ± 2.21	23.51 ± 3.36	0.79
Days between MicroFET2 sessions		23.8 ± 9.0	6.9 ± 1.2	<0.001*
Sample				0.567
	Male	12	11	
	Female	11	7	
Dominant Side (*n*)				0.439
	Right	21	15	
	Left	2	3	
Total (*n*)		23	18	

Values are presented as mean ± standard deviation (*M* ± *SD*); GrAASP, groin abductor-adductor strength platform; STAPH, hip strength assessment platform; *, significant difference at 0.05.

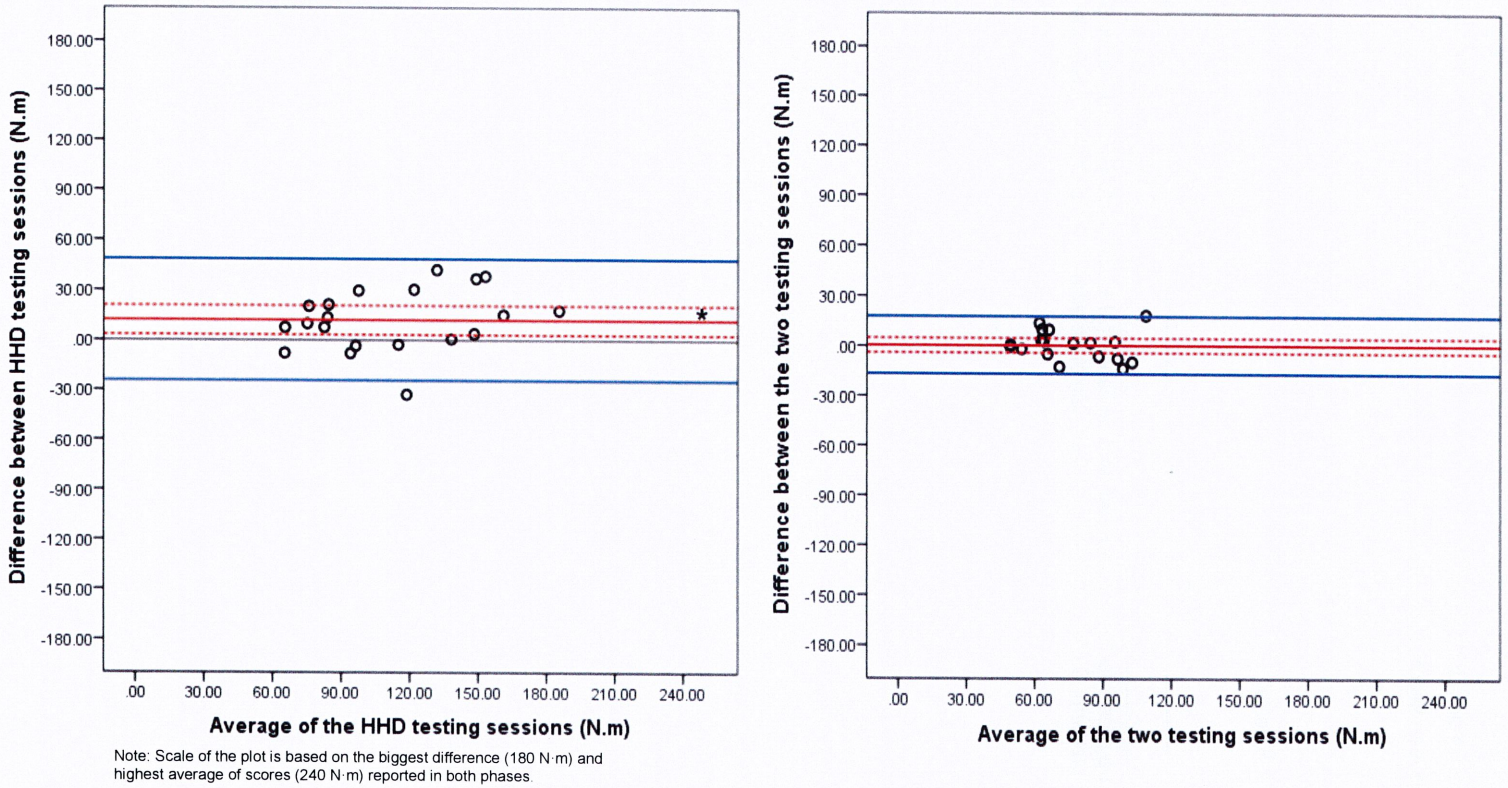


Figure 6.28 - Bland-Altman plot of hip abduction testing using GrAASP (fixed method on a distal location) (left) and using the STAPH (right) (red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits; *, significant difference at 0.05).

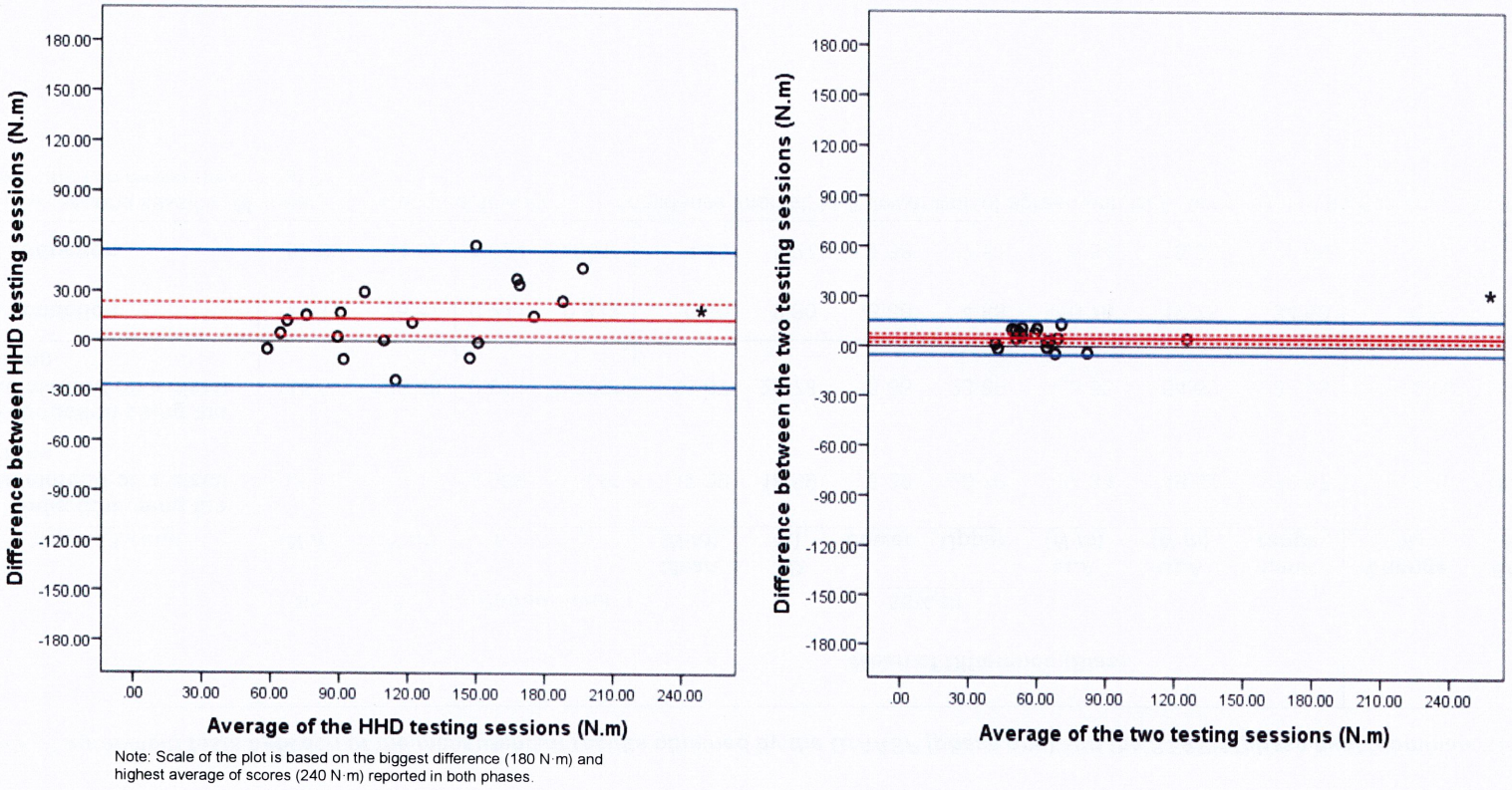


Figure 6.29 - Bland-Altman plot of hip adduction testing using GrAASP (fixed method on a distal location) (left) and using the STAPH (right) (red line illustrates mean of difference; dotted red lines illustrates 95% confidence interval of the mean of difference; blue lines illustrate agreement limits; *, significant difference at 0.05).

Table 6.23 - Limits of agreement tests between of the measurement results obtained by the GrAASP [phase one] and the STAPH [phase two] (dominant side only).

Platform	Muscle Group	S1 (*N·m*)	S2 (*N·m*)	Paired *t*-test *t*	Paired *t*-test *p*	Mean of Difference (Bias) Mean (*N·m*)	*SD* (±)	95% CI Lower	95% CI Upper	LLA (*N·m*)	ULA (*N·m*)	Limit range	Change (%)	Error (%)	CoR (*N·m*)
GrAASP	Hip abduction using the fixed method on a distal location	106.06	118.15	2.909	0.009*	12.09	18.59	3.39	20.79	-24.34	48.52	72.87	11.40	16.41	36.43
	Hip adduction using the fixed method on a distal location	115.61	129.48	2.911	0.009*	13.87	20.78	3.86	23.89	-26.85	54.60	81.45	12.00	17.32	40.72
STAPH	Hip abduction	75.17	75.67	0.241	0.812	0.50	8.80	-3.88	4.88	-16.75	17.75	34.50	0.67	11.65	17.25
	Hip adduction	59.89	64.89	4.024	0.001*	4.99	5.27	2.38	7.61	-5.34	15.32	20.66	8.34	8.74	10.33

S1, first session; S2, second session; *M*, mean; *SD*, standard deviation; CI, confidence interval; LLA, lower limit of agreement; ULA, upper limit of agreement; *, significant difference between the two measurements at 0.05.

Chapter 7 Study Two: Discussion

7.1. *Chapter Overview*

The validity and reliability of the Groin Abductor-Adductor Strength Platform (GrAASP) and the Hip Strength Assessment Platform (STAPH) were assessed in two phases. The results of phase one of testing showed statistically significant differences between the three angles of hip abduction in isometric testing (hip abduction only) and the two speeds of isokinetic testing. Moreover, the results indicated that stabilising the MicroFET2 at a distal location had a stronger agreement with the Biodex and showed less measurement error. Based on this finding, the participants were tested using the STAPH in phase two by fixing the MicroFET2 at a distal location to test seven muscle groups. The findings of both phases of testing revealed no significant differences between the dominant and the non-dominant sides. Also, the male participants were significantly stronger than the female participants. Overall, it is recommended to use the STAPH over the GrAASP in testing hip and thigh muscle strength due to the reduced measurement error and improved stability. In this chapter, the results of study two will be discussed and compared to similar research.

7.2. *Bilateral Torque Differences*

The coordinated contractions of the deep and superficial muscles result in tri-axial movements in the hip joint (Neumann, 2009b). The imbalance and alteration in force production by any of these muscles may affect the quality of movement in the hip and knee joints (Neumann, 2009a). The importance of studying the hip muscles is based on their reciprocal and synchronised activity, which is an essential component of lumbopelvic stability and smooth transfer of movement to the lower limbs (Lee & Kim, 2015).

The bilateral torque variation was minimal in most muscle groups. Moreover, there were statistically insignificant differences between the dominant side (DS) and the non-dominant side (NS) in both phases of testing. The study results also indicated that the muscle torque values on the DS were higher than the NS except for the hip adductor (HAD) muscles in phase one, together with hip external rotation (HER) and the knee flexion (KFL) muscles in phase two. A previous study showed that the DS (81.0 ± 23.7 *N·m*) could be significantly stronger than the NS (76.1 ± 19.9 *N·m*) in isometric peak torque measurements of the hip abductor (HAB) muscles (p =0.02) (Jacobs, et al., 2005).

The findings in phase one testing showed that the NS of the HAD muscle had higher torque values than the DS when tested isometrically at 0° and isokinetically at both speeds of testing. Belhaj et al. (2016) tested twelve asymptomatic football players and found that the NS was stronger than the DS in HAB (DS, 129.2 ± 22.87 *N·m*; NS, 131.8 ± 19.2 *N·m*) and HAD isokinetic testing (DS, 111.3 ± 24.0 *N·m*; NS, 115.5 ± 36.8 *N·m*) at 60°.sec^{-1} (both p >0.05). In a study involving thirty-six male professional football players, the NS was also stronger than the DS in knee extension (KEX) (DS, 208.8 ± 47.8 *N·m*; NS, 209.3 ± 50.5 *N·m*, p =0.311) and KFL testing at 60°.sec^{-1} (DS, 100.1 ± 27.6 *N·m*; NS, 102.3 ± 30.5 *N·m*, p =209) (Daneshjoo, et al., 2013). The significant differences in isometric torque between both sides in the study by Jacobs et al. (2005) are due to the larger sample size (n=42), mode of testing, and testing the participants in a side-lying position compared to a standing position in the current study. The differences in the testing position will be discussed later. Nevertheless, the difference between both sides is about 5 *N·m*, which may be clinically irrelevant. However, the HAD isokinetic side-to-side differences of the present study agree with the findings of Belhaj et al. (2016) and Daneshjoo et al. (2013).

Bilateral torque differences of >10% were reported in patients with various musculoskeletal complaints (Burkett, 1970; Calmels, et al., 1997; Nadler, et al., 2002).

In the current study, side-to-side isometric and isokinetic torque differences were found to be less than 10% when measured using the Biodex, the GrAASP and the STAPH. The participants who were involved in both phases of the study were healthy, pain-free and did not complain of any recent injury to the lower extremities in the past three months. The only exception was the NS of the HAD muscles in isokinetic testing, which showed 20.6% and 16.6% higher torque values than the DS at $60°.sec^{-1}$ and $120°.sec^{-1}$ respectively. The high torque deficit found in HAD isokinetic testing and 4.0% drop in bilateral side torque differences when the speed changed from $60°.sec^{-1}$ to $120°.sec^{-1}$ may be the result of testing error and learning effect, which impacts Biodex measurements (Lund, et al., 2005). Consequently, the author recommends adding more familiarisation sets when isokinetic testing is performed to reduce the learning effect.

Testing using the STAPH revealed that the highest bilateral torque deficits were observed in the hip flexor (HFL) muscles. The rectus femoris aids the iliopsoas muscle in producing HFL (see Appendix A. 1) (Neumann, 2009b). By placing the participants in a neutral hip and an extended knee position, the rectus femoris flexes the hip and extends the knee simultaneously to push against the MicroFET2. An EMG study showed that the rectus femoris, adductor longus and iliacus muscles contracted at the same time before the psoas major muscle* during the active straight leg raising test, which is performed in a supine position with an extended knee. Moreover, the psoas major muscle was active bilaterally during a unilateral test (Hu, et al., 2011). Therefore, this muscle is assisted by the rectus femoris to flex the hip and it is nearly impossible to isolate the iliopsoas from the rectus femoris in this position. The combined action of the psoas major and rectus femoris muscles may be the reason behind the increased bilateral torque differences since both muscles are strong and contract simultaneously. For perspective research, the author recommends testing the hip flexor muscles by

* Part of the iliopsoas muscle group complex.

placing the participants in a seated position with the hip joints flexed to 90° to minimise the effect of rectus femoris muscle action on flexing the hip joint.

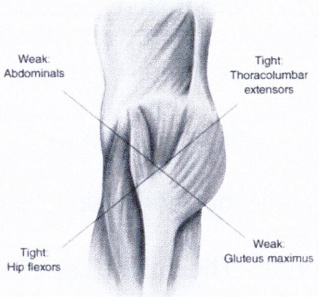


Figure 7.3 - Lower Crossed Syndrome (Janda, 2014).

The hip internal rotation (HIR) and HAB tests showed the smallest bilateral torque deficits. These two movements are produced by different fibres of the gluteus medius and minimus muscles (Neumann, 2009b), and play a major role in hip and knee kinematics during landing (Suzuki, et al., 2015). The bilateral HAB torque difference measured by the STAPH (1.2%, Figure 6.20) was very close to the bilateral deficit measured by the Biodex at 0° in phase one (1.9%, Figure 6.1). Both torque measurements were performed in 0° of HAB. Such observation indicates that testing using the STAPH is closer to a gold standard (Biodex) compared to the GrAASP. Additionally, the proximity in findings may be due to the similarities in the characteristics of the two samples (Table 6.22).

One of the interesting findings of STAPH testing was the recognition of a muscle strength-weakness pattern. The pattern is very much like the Upper (UCS) and Lower Crossed Syndromes (LCS) described by Vladimir Janda in 1987. The UCS is formed by the facilitation of the upper trapezius, sternocleidomastoid and the pectoralis major combined with inhibition of the deep cervical flexor muscles, lower trapezius and serratus anterior. In the pelvic region, the facilitation of the thoracolumbar extensors, iliopsoas and rectus femoris, with inhibition of the transverse abdominis and the gluteal muscles, make the LCS (Figure 7.3) (Janda, 2014).



Figure 7.2 – The pattern of hip external rotator muscles weakness (Circles refer to gluteus maximus; triangles refer to deep external rotator muscles; green and yellow colours refer to stronger and weaker muscles respectively in comparison to the other side).

The pattern observed during STAPH testing showed that the strong hip extensor (HEX) muscle (i.e. the gluteus maximus [GMax]) was coupled with weak HER (i.e. the small deep muscles) on the DS. The opposite is true for the contralateral side (Figure 7.2). In the

horizontal plane, enhanced stability of the hip and external rotation are achieved by contracting the HER muscles and the GMax (see Appendix A. 1) (Neumann, 2009b). In 90° of hip flexion, which is the testing position for the HER muscles, the external rotation moment arm of the GMax reduces dramatically (Delp, et al., 1999). As a result, the role of this muscle in producing HER in a sitting position is minimal, even though it produced high torque when tested in 0° of hip flexion (Neumann, 2010). The reduction in HER torque output in a sitting position indicates that the deep rotator muscles, not the GMax, are inhibited on the DS compared to the NS. If the HER muscles of the same participants were tested in an ideal horizontal plane (i.e. standing with 0° of hip flexion), the GMax might compensate for the inhibited deep rotator muscles because of the increased moment arm. Therefore, it is advised to test the HER muscles from a position that minimises the work of the GMax such as a seated position to obtain true strength values. As a result, testing using the STAPH can be considered a quick and easy method of detecting the presence of muscle weakness.

The pattern of hip muscle weakness found in the study can be considered normal. All participants were healthy, active and free of injury, which explains the reduced side-to-side isometric (0.4% to 2.1%) and isokinetic (0.1% to 4.9% [except the HAD muscles]) torque differences as measured by the Biodex. Therefore, it is safe to assume that their bodies have adapted to the muscular imbalances by recruiting accessory muscles to maintain stability. If the side-to-side strength asymmetry exceeds 10% to 15%, lower extremity ligamentous and soft tissue injuries such as hamstring strains and lower back pain can develop (Burkett, 1970; Calmels, et al., 1997; Nadler, et al., 2002).

Previous studies showed that increasing the strength of the HER and HAB to minimise bilateral differences has been associated with improved hip and knee dynamics, function and pain (Fukuda, et al., 2010; Khayambashi, et al., 2012). As a recommendation, once a clinician identifies a bilateral torque difference of more than 15%, further testing should be performed if required and a strengthening exercise program should be prescribed to reduce such this deficit.

Key Points

- No significant difference between both sides in isometric and isokinetic strength of the muscles measured in both phases of the study.
- Weakness of the HER muscles can be masked by strong HEX muscles. A thorough examination should be performed to detect this pattern.
- Bilateral torque differences of >15% should be further investigated because it may result in future complications.

7.3. ***Sex Differences***

Sex-based analyses indicated that the males were significantly stronger than the females in Biodex, GrAASP and STAPH testing. The differences in torque output between males and females were observed based on the ANOVA test results in both phases and multilevel modelling (MLM). When sex was considered as a factor in MLM, the female participants showed a statistically significant reduction in scores using both methods of stabilisation at both placement locations (Table 6.14). Thus, the differences between males and females were observed in both phases of the present study using the Biodex, GrAASP and STAPH.

The differences between sexes ranged from 18.0% to 49.5% in Biodex testing and from 1.8% to 28.1% in STAPH testing. Unlike the male participants, the female participants had reduced overall bilateral torque differences. Moreover, the DS of the male participants was always higher than the NS, and higher than both sides of the female participants. The greatest sex differences were observed in HAD isometric testing at 0° on the DS (49.5%) using the Biodex in phase one.

When the participants were tested using the STAPH in phase two, the females had higher torque values on the DS in all measurements except KFL testing. The results are consistent with the findings of Suzuki et al. (2015), who noted that female basketball

players had significantly weaker HAB, and weaker HER in a seated position compared to the male players.

The muscle strength differences between males and females are linked to changes in hip and knee kinematics. Female basketball players show an increased knee flexion, knee valgus, and hip internal rotation angles at initial contact compared to the male players (Suzuki, et al., 2015). When the lower limb mechanics at initial contact were analysed, the female runners presented with greater hip adduction, and increased knee valgus in comparison to the male runners (Willson, et al., 2012) (see section 2.2).

The variation in force production between the two sexes can be due to several factors. Dissimilarities in the pelvic dimensions, muscle morphology, hormonal production and limb measurements are possible reasons to interpret the findings (see section 2.2). However, the pattern of sex-torque discrepancies in both phases of the study cannot be explained because they were not based on a specific mode of testing, muscle group or plane of motion. To detect a clear pattern in torque differences, a sample size of more than forty participants is recommended for this type of study (Lu, et al., 2011).

The results of the current study together with the findings of Willson et al. (2012) and Suzuki et al. (2015) demonstrate significant differences in hip muscle strength between males and females. The reduction in strength and altered biomechanical constraints put the female players at a higher risk of developing non-contact ACL injuries and chronic overuse syndromes (Taylor-Haas, et al., 2014). Physiotherapists and athletic trainers are advised to focus more on strengthening these muscle groups to reduce the likelihood of internal and external joint pain and complex injuries.

Key Points

- Torque output varies greatly between males and females. The measured strength values should be considered separately based on sex, mode of testing and the muscle group.

- Well-structured and customised strength training programs are needed to achieve enough hip muscle strength and to minimise the anatomical and hormonal differences between males and females.
- A larger sample size is needed to detect a clear pattern in torque differences between males and females based on the muscle group or plane of motion.

7.4. *Angles and Speeds Differences*

Comparing HAB and HAD torque values in 0°, 15°, and 30° of hip abduction revealed unique findings for each muscle group. Isometric HAB torques showed no statistically significant changes as the angle increased from 0° to 30°. The results of the current study regarding the inverse relationship between the angle of hip abduction and the measured HAB torque oppose the findings of previous studies (Figure 7.3) (Olson, et al., 1972; Neumann, et al., 1988).

Neumann et al. (1988) argued that the hip angle had a statistically significant effect on the HAB torque (p <0.0001) (Figure 7.3). Neumann et al. (1988) also reported lower values at 0° (Rt = 96.55 ± 28.75 $N{\cdot}m$; left = 88.53 ± 24.90 $N{\cdot}m$) and 30° (Rt = 52.82 ± 18.94 $N{\cdot}m$; left = 50.04 ± 15.85 $N{\cdot}m$) compared to the current study. The differences in the measurement device and position may explain the variation in findings. The participants in the study by Neumann et al. (1988) were assessed in a supine position using a strain gauge compared to a standing position and using the Biodex in the current study.

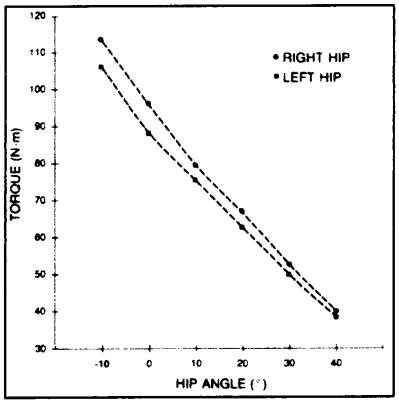


Figure 7.3 - Hip abductor muscle torque-angle data for both sides (Neumann et al., 1988)

Isometric HAD testing showed a marked increase in torque values between 0° and 30°. The increase in torque values is due to the increase in the length of the muscle fibres resulting in higher force production (Martini & Nath, 2008). The published research on

hip abductor and adductor muscles isometric testing using the Biodex is very limited. Also, due to the anatomical and biomechanical variations between these two muscle groups and other muscles (e.g. the quadriceps), further comparisons between the findings of this study and other similar studies cannot be achieved at this stage. Based on the findings of isometric HAB and HAD testing using the Biodex, the researcher recommends testing the HAB muscles from 0° of hip abduction since there were no significant difference between the three angles. On the other hand, the HAD muscles can be tested from 30° of abduction to record the highest values in muscle torque, which is the aim of muscle strength testing.

The results of isometric and isokinetic testing using the Biodex and isometric testing using the STAPH showed that the HAB muscle group produced higher scores than the HAD muscle group. The HAD muscles' lines of force approach the hip joint in all three planes of motion. As a result, the HAD muscles are active as primary and secondary muscles in the frontal, sagittal and transverse planes. For example, the adductor magnus can act as a hip extensor and as a hip flexor based on the position of the hip in the sagittal plane (Neumann, 2009b). Even though the HAD muscle group consists of multiple large muscles, aligning the hip in a single plane of motion during testing can limit their function. Thus, much of the strength and functional abilities are lost resulting in lower torque values.

The findings of HAB isokinetic testing revealed statistically significant differences between $60°.sec^{-1}$ and $120°.sec^{-1}$, and slightly higher values at $120°.sec^{-1}$ than $60°.sec^{-1}$, which was also reported by Calmels et al. (1997). On the other hand, isokinetic HAD torque values at $120°.sec^{-1}$ were significantly lower than the values at $60°.sec^{-1}$. Compared to the results of Lourencin et al. (2012), the reported values in the current study were lower by 35% and 61% at $60°.sec^{-1}$, and 20% and 63% at $120°.sec^{-1}$ for the HAB and HAD muscle groups respectively. The participants in the study by Lourencin et al. (2012) consisted of 20 males who were assessed in a side-lying position, while the current study consisted of 11 males and 9 females assessed from a standing position.

Aligning the participants in a side-lying position stabilised the contralateral side resulting in higher reported values. A side-lying position can increase HAB values by 10% compared to the standing position (Widler, et al., 2009). On the contrary, other research has shown that the standing position resulted in a 32% increase in the HAB and HAD torques at $30°.sec^{-1}$ and a 28% increase at $90°.sec^{-1}$ (Abdel-aziem, et al., 2013). The inconsistent results of isokinetic testing in these studies are due to the variation in the participants' characteristics, position, protocol and speed. Therefore, reaching a definite conclusion about the effect of testing position on torque values may not be possible with the available information.

The reduction in torque as the speed increased from $60°.sec^{-1}$ to $120°.sec^{-1}$ was expected. It is based on the inverse relationship between angular velocity and force development, which states if the velocity of muscular shortening is high, the developed force is low (Lourencin, et al., 2012). Gerodimos et al. (2015) also reported a reduction in the HAB and HAD peak torques as the speed changed from $30°.sec^{-1}$ to $90°.sec^{-1}$ in which the HAB muscles showed higher concentric peak values than the HADs. Another study showed 48% reduction in load range* in repeated testing as the speed of testing gradually increased from $60°.sec^{-1}$ to $360°.sec^{-1}$. More specifically, the load range reduced by 6.7% for the HAB and 7% for the HAD from $60°.sec^{-1}$ to $120°.sec^{-1}$ (Gautrey, et al., 2013). Another explanation could be muscular fatigue since isokinetic testing was performed after isometric testing in the present study. To obtain more accurate isokinetic strength values, the author recommends that the participants should be tested in two sessions with sufficient rest time in between.

* The range of motion during which an external resistance is applied.

Key Points

- The findings suggest a recommendation to isometrically test the HAB muscles at 0° and the HAD muscles at 30°.
- Testing concentric muscle strength at moderate speed should be set at 60°.sec^{-1} or lower for both muscle groups.

7.5. ***Muscle Strength Ratios***

Isometric HAD to HAB strength ratios obtained by the Biodex differed as the angle of hip abduction and speed changed. The ratios are directly related to the change in HAD muscle strength, showing higher strength values as the angle increased towards abduction (Table 6.7). This phenomenon is explained by the length-tension relationship, which states if muscle fibres are either shorter or longer in relation to the mid-range, the tension developed by the muscle will be decreased. However, when the muscle fibres are at optimum length, in a state of filament overlap, they produce the highest amount of force (Martini & Nath, 2008).

The effect of changing the muscle fibre length was less evident in the HAB muscle group. A study showed that repositioning the femur in the frontal plane from a neutral position to 45° of abduction increases the moment arm of the anterior (6.0 cm to 7.4 cm respectively), middle (5.4 cm to 7.2 cm respectively), and posterior fibres (4.2 cm to 6.2 cm respectively) of the gluteus medius muscle (Henderson, et al., 2011). Thus, the gluteus medius muscle can produce higher torque and work more efficiently as the angle increases. However, HAB muscle torque did not change as the angle of hip abduction increased.

The results of the current study are not very different from the findings of a study that involved six healthy competitive tennis players. The HAD and HAB strength were measured in a side lying position at 45° of abduction, and the ratio was approximately 1.16. Moreover, it was concluded that an imbalance of 16% to 36% in the DS and 14%

to 36% in the NS were present between the HAD and HAB with the second muscle being stronger (Zuša, et al., 2015). Isokinetic testing in the present study showed a reduction in the strength ratio as the speed increased from 60°.sec^{-1} to 120°.sec^{-1}. The drop in ratio is related to the inverse relationship between angular velocity and strength output (explained in section 7.4). In the study by Lourencin et al. (2011), the HAD-HAB strength ratio of twenty male volunteers reduced from 1.03 at 60°.sec^{-1} to 0.98 at 120°.sec^{-1}.

The differences between the strength ratios reported in the present study to the studies by Zuša et al. (2015) and Lourencin et al. (2011) emphasise on the importance of considering such ratios carefully. The calculated ratios apply to the mode of testing, position and sample characteristics, and should not be generalised. Other studies looking into similar ratios in the isokinetic testing of the HAB and HAD were not found.

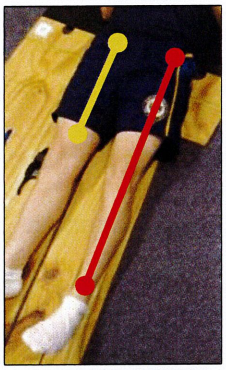


Figure 7.4 – The short (yellow) and long (red) lever arms.

The ratios associated with the fixed method were overall higher than the manual method in phase one of testing. The ratios can be compared to the study by Thorborg et al. (2010a) due to the similarities in the design, inclusion criteria, contraction, and rest periods. In the current study, the ratios of HAD to HAB in the supine position using the manual method of stabilisation distally were 1.19 ± 0.40 in the first session and 1.17 ± 0.21 in the second session when the best of three repetitions was considered. However, the ratios reported by Thorborg et al. (2010a) were equivalent to 0.96 ± 0.16 when the best of four repetitions was considered. The variation in the results between the present study and the study by Thorborg et al. (2010a) is due to several differences in the sample and testing protocol. In the latter study, the sample consisted of five male (mean age 27 ± 5 years) and four female (mean 25 ± 4 years) participants who were tested on an examination table using a different HHD (Powertrack II commander) and could stabilise themselves. In contrast, the present study consisted of eleven males and nine females,

who were tested on the floor using the MicroFET2 and were not allowed to stabilise themselves.

A pattern was noticed in the results when the methods of stabilisation were compared. Proximally, the strength ratios using manual stabilisation were always less than the fixed stabilisation. Testing error, in which the examiner was unable to resist the HAD of stronger participants, may have resulted in the reduced recorded values and ratios. On the other hand, the higher ratios obtained proximally using the fixed method could have resulted from higher force application against the MicroFET2, which was more stable than the examiner's hand (Fenter, et al., 2003b; Kollock Jr, et al., 2013). During testing, several participants reported that the lack of stabilisation with the manual method negatively affected their performance.

When comparing proximal to distal placement, the reduced ratios obtained distally could be due to the longer lever arm from the point of application to the axis of motion, making it harder to push against (Figure 7.4). A biomechanical model of the hip joint was proposed and discussed in Appendix F to explain this finding. The researcher is not aware of other studies looking at the HAD-HAB strength ratios at two different locations and using two methods of stabilisation. Thus, a further comparison with existing literature is not possible.

The strength ratios obtained in phase two of the study were different than the ratios from phase one. Surprisingly, the ratio between the HEX and HFL was close to 1 in both MicroFET sessions. Such a value for this ratio indicates that the strength of the HEXs and HFLs in the tested sample is almost equal. Other ratios indicated that the HIR muscles were almost 36% stronger than the HER muscles, which is normal due to the anatomical properties and size of the HIR muscles compared to the HER muscles (excluding the GMax in a sitting position) (Neumann, 2009b).

The average HAD to HAB muscle strength ratio from phase two were slightly lower than those in phase one. The ratio obtained at a distal location using a fixed method is

considered in comparing between the two phases (Table 6.7 and Table 6.19). The participants involved in both phases were similar in stature, body mass, sex and dominant side. Therefore, the ratios are more likely related to the accuracy of testing using the GrAASP and STAPH. A direct comparison between the two platforms is further discussed in section 7.9.

The hamstring muscle is a primary knee flexor and contributes to producing hip extension because it crosses two joints. The results have shown that the KFL muscle (i.e. hamstring muscle) to HEX (i.e. gluteus maximus and hamstring muscles combined) strength ratio is very high (Table 6.22). In other words, the hamstring muscles can produce about 130% to 140% higher torque values than the HEX when the hip and knee joints are placed in 90° of flexion.

The dramatic increase in torque production between sitting (for KFL) and side lying (for HEX) positions in the current study is due to the change in the muscle's biomechanical properties. Sixteen participants were recruited in a study to test the relationship between the hamstring muscle's length, activity and torque. The results of testing using the Cybex showed a significant increase in the hamstring muscle's torque from 38.7 $N{\cdot}m$ at 0° to 57.3 $N{\cdot}m$ at 90° of hip flexion (p <0.05). The increase in torque was coupled with a drop in the biceps femoris'* muscle EMG activity from 1952 mV.sec^{-1} at 0° to 1547 mV.sec^{-1} at 90° (p >0.05) indicating that a fewer motor units are recruited. The findings also point that the hamstring muscles produced higher torque in the lengthened position (90° in HFL) compared to a shortened position (0° in HFL) (Lunnen, et al., 1981). Additionally, it is thought that placing the muscle fibres in a lengthened position elongates the elastic components of the muscle, which act as a recoil to take the fibres back to normal length (Lunnen, et al., 1981). The tension-length relationship could have also played a role. Nevertheless, placing the knee in 90° of flexion can significantly lower the hip extension component of the hamstring muscle from 73.95 ± 26.05 to 36.05 ± 15.75 (EMG activity

* Part of the hamstring muscle group.

normalised to %MVC) (p <0.05) (Kwon & Lee, 2013). Thus, the functional properties of the hamstring muscle can change based on the hip and knee positions and the desired motion.

Muscle strength ratios have an important clinical use in the diagnosis, progression and prognosis of cases. Tyler et al. (2001) found a reduction in the HAD-HAB strength ratio from 95% in uninjured players to 78% in the injured players with reduced HAD strength (p =0.011). Also, the players who had a preseason history of HAD strain and injury reoccurrence showed similar results with a ratio of 86% on the uninjured side compared to 70% on the injured side. Further analysis showed that the HAD-HAB strength ratio was the best predictor of developing HAD strain in the future. A study by Belhaj et al. (2016) was performed on male professional football players with adductor-related groin pain at $60°.sec^{-1}$. The HAD to HAB strength ratio in those players was approximately 0.50, and 0.86 in asymptomatic football players. However, when healthy volunteers were tested at the same speed, the ratio increased to roughly 1.27. The findings of Belhaj et al. (2016) show that the HAD-HAB strength ratio can markedly reduce in football players with adductor-related groin pain compared to asymptomatic players or healthy volunteers. Thus, the obtained ratios of the current study are clinically beneficial to detect any change in muscle strength. If such ratios were considered, addressing the muscular weakness in a timely fashion by a proper strength training program could reduce the likelihood of developing muscle strains (Jensen, et al., 2014).

Key Points

- The ratios apply to the mode of testing, position and sample characteristics, and should not be generalised.
- The ratios associated with the fixed method were overall higher than the manual method and length of the lever arm affects the ratios.

- The HIR muscles were 36% stronger than the HER muscles. The hamstring muscles can produce about 130% to 140% higher torque values than the HEX in 90° of hip flexion.
- Muscle strength ratios have an important clinical use in the diagnosis, progression and prognosis of cases.

7.6. *Agreement Between the Biodex and the MicroFET2 in Phase one*

The limits of agreement tests in phase one revealed that the MicroFET2 results were not equivalent to the scores obtained on the Biodex. The measurements of HAB torque by the MicroFET2 were less than those measured by the Biodex, and the opposite is true for the HAD torque measurements. Martin et al. (2006a) found that using a HHD in assessment can underestimate the quadriceps' muscle strength by a mean of 14.5 $N{\cdot}m$, 95% CI [8.5, 20.6]. The strength values of the same muscle were also underestimated by 33.7 $N{\cdot}m$ on the right and 35.6 $N{\cdot}m$ on the left in another study (Bohannon, et al., 2011).

The Bland-Altman plots showed that the variation in scores differed as the strength of the participants changed. For instance, the difference between the two devices got wider with stronger participants, which was more evident in HAD measurements. Martin et al. (2006a) also reported that the measurement differences were more noticeable in the stronger participants. The authors recommended that the strength of the tester relative to the strength of the participant should be taken into consideration when larger muscle groups are being tested (Martin, et al., 2006a).

The limits of agreement tests in HAB and HAD isometric testing were related to the angle of hip abduction. In HAB, the smallest limit range and ME were observed between isometric testing at 0º and when the fixed method was applied distally (Table 6.8). On the other hand, the smallest MDs and MEs in HAD testing were between isometric testing at 30º and the fixed method when applied at both locations (Table 6.9). The

results indicate that manual stabilisation of the MicroFET2 at a proximal location led to higher differences in scores and an increase in the ME compared to the Biodex. Thus, this finding supports testing the hip muscles using a fixed method of stabilisation at a distal location.

The agreement between the two devices was different in isokinetic testing. The smallest MDs and MEs were reported between testing at the speed of $60°.sec^{-1}$ and the distal (fixed) location in HAB testing and the proximal (manual) location in HAD testing. As mentioned earlier, the HAB torque values were significantly less when tested using the MicroFET2 compared to the Biodex. Meanwhile, the torque values of the HAD muscles using the MicroFET2 were close to those obtained using the Biodex. The disagreement between the lower speed of testing and the location of placement in both muscle groups is based on the torque values only. From a biomechanical view, the isometric muscle strength of the participants was tested using the MicroFET2. Despite this, their isometric strength values were compared against isokinetic measurements using the Biodex. Therefore, the researcher does not recommend comparing isometric to isokinetic values. The reader should be aware of these differences when he or she comes across a similar comparison.

The correlation between the Biodex and MicroFET2 also varied depending on the muscle group, testing location and stabilisation method. In isometric HAB muscle testing, the highest correlation was found between Biodex testing at 0° and MicroFET2 testing using the fixed method distally (r =0.757, p <0.001) (Table 6.10). In contrast, the highest correlation in isometric HAD testing was between Biodex testing at 30° and using the fixed method proximally (r =0.778, p <0.001). In isokinetic mode, Biodex testing at $60°.sec^{-1}$ was significantly correlated with HAB testing using the manual method distally (r =0.755, p <0.001) but insignificantly correlated in HAD testing using similar settings (r = -0.241, p =0.294). The reasons behind this dissimilarity could be due to testing error in the participants' positioning and not following or understanding the commands properly.

The correlation between the HHD and isokinetic devices in testing lower limb muscle groups has been investigated by several researchers. Bohannon et al. (2011) found a strong correlation between a belt-stabilised HHD and the Biodex in KEX testing of the right (r =0.884, p <0.001) and left sides (r =0.864, p <0.001) (Table 2.15). Also in KEX strength testing, Kim et al. (2014) concluded that a fixed or a stabilised HHD correlated with the Biodex (Rt, r =0.806; Lt, r =0.524, both p <0.05) more than the non-fixed HHD (Rt, r =0.185, p >0.05; Lt, r =0.384, p <0.05) in a supine position (Table 2.16). Martin et al. (2006a) showed that the HHD has a strong correlation with the Biodex (r =0.91, p <0.0001) in measuring the quadriceps' strength of elderly participants (Table 2.11). In testing other hip muscles, Kawaguchi and Babock (2010) found strong correlation between the two devices in HEX (r =0.419) and HAB testing (r =0.834) (Table 2.11). Katoh et al. (2011) observed a moderate correlation between a belt-stabilised dynamometer and the Biodex in HAB testing (r =0.65) compared to KFL and KEX (r =0.88 and r =0.75 respectively, all p <0.05). The lower reported correlation values are due to the testing position, in which the knee muscles were tested in a seated position while the hip muscles were tested in a side-lying position. Overall, the findings of the present study that using a fixed HHD is more reliable and has a stronger correlation with the Biodex come in agreement with these studies.

Key Points

- The Bland-Altman plots showed that the variation in scores differed as the strength of the participants changed.
- Using a fixed HHD at a distal location is more reliable and has a stronger agreement with the Biodex.
- It is not recommended to compare isometric to isokinetic strength values.

7.7. *Agreement Between the MicroFET2 Sessions in Phase One*

The limits of agreement tests have been used by many researchers in dynamometry testing (Beshay, et al., 2011; Katoh, 2015; Lu, et al., 2011; Savva, et al., 2014). In this research, the measurements of the second MicroFET2 session were significantly higher than the first session. Between the two MicroFET2 testing sessions, HAB muscle testing using the fixed method at a distal location had a lower mean of difference and percentage of error in comparison to the manual method at a proximal location (Pair 41 and 44, Table 6.11). The Bland-Altman plot of HAB testing using the manual method proximally showed that the difference between the two sessions increases when the average score of the participants exceeded 85 *N*.m. However, it increased to 100 *N*.m approximately using the fixed method distally (Figure 6.15 and Figure 6.16).

HAD testing showed similar results using the fixed method distally compared to the manual method proximally (Pair 45 and 48, Table 6.11). The Bland-Altman plots of HAD measurements indicated that the difference between the two sessions increased when the average score exceeded 75 *N*.m approximately in using the manual method proximally. In contrast, the threshold was set at 150 *N*.m approximately when the fixed method was used distally (Figure 6.17 and Figure 6.18).

The placement location of the MicroFET2 can affect the torque measurements. Placing the MicroFET2 distally showed considerably different results to the proximal placement location. For example, testing the HAD muscles had higher means of differences proximally compared to distally. Krause et al. (2007) demonstrated that torque measurements using a long lever arm in HAB testing was significantly greater and had a higher intra-tester reliability than the short lever arm (short, ICC =0.91; long, ICC =0.93, p <0.001). Additionally, a statistically significant interaction between the lever arm and bench stabilisation of the opposite side was found in HAD testing (short, ICC =0.62; long, ICC =0.82, p <0.001) (Table 2.12). Lu et al. (2011) found that a high MD and high CoR in testing the KEX were associated with and lower reliability. On the other hand, lower MD, lower CoR in HAB tests were linked with higher reliability (Table 2.10).

The difference in measurements between the two sessions was also affected by the method of stabilisation. Applying the fixed method on a distal location showed the smallest change, measurement error and CoR in HAB and HAD tests (Table 6.11). Moreover, the smallest variation in scores between the two sessions was also associated with the distal location (Table 6.12). In this phase of the study, the MicroFET2 was stabilised using a plastic box attached to a metal upright. Other researchers have used various methods of stabilisation, which served the same purpose (Table 2.15). Kramer et al. (1991) also noticed that using a belt-stabilised dynamometer was more reliable in testing the young (Rt, ICC =0.92; Lt, ICC =0.92) and elderly patients (Rt, ICC =0.98; Lt, ICC =0.98). Katoh and Yamasaki (2009) reported excellent reliability of using a belt-stabilised dynamometer (ICC range 0.97 to 0.99) compared to manual stabilisation (ICC range 0.41 to 0.83) in hip muscle strength testing. Both, Kramer et al. (1991) and Katoh and Yamasaki (2009) observed higher torque values when the belt was used as a method of stabilisation (Table 2.13). For further details, refer to section 2.7.

The differences in scores between the two MicroFET2 sessions of phase one were also explored using MLM. The findings of MLM indicated that the measurements of the second session showed a statistically significant increases in both methods of stabilisation at both placement locations (Table 6.14). The possible reasons behind increase in measurements of the second MicroFET2 session are discussed in section 7.9.

Key Points

- Applying the stabilised method reduced measurement error, produced a better coefficient of repeatability and a narrower range of agreement limits.
- Similarly, the distal location of fixation has shown to be better than the proximal due to the higher consistency in results.

- The findings support the use of the GrAASP to stabilise the MicroFET2 at a distal location to improve repeatability of the scores.

7.8. *Agreement Between the MicroFET2 Sessions in Phase Two*

The second MicroFET2 testing session of phase two showed that the change in muscle torque was less than 5 *N*.m. The gluteus medius and minimus, which are the primary muscles for HAB and HIR, demonstrated the highest coefficient of variation (CV) values of 10.04% among all muscles. The variation indicates that the DS and NS scores differed between the two sessions. Even though, when the means of scores were considered, they showed the smallest mean of difference and percentage of change between the two sessions (Table 6.20). An important point to consider is that the CV is a measure of relative variability and precision of the collected data not the measurement device, and it is more applicable to the individual's variation in scores (Hopkins, 2000). The sample consisted of males and females with varying strength levels, therefore, a high CV is expected. On the other hand, the mean of difference between the two sessions reflects the change in scores, which is more important in this study.

The test-retest agreement in scores of hip muscle testing using the STAPH is not affected by the tested muscle group or plane of motion. The HFL and HAD muscle groups had the narrowest limit range, testing error and CoR. These two muscle groups function in two different planes of motion but showed the smallest testing error. The HFL muscles were tested in a side-lying position while the HAD were tested in a supine position. Moreover, the hip joint was aligned differently in each position. Similarly, HAB and HIR, which are controlled by the same muscles, showed the highest variation in scores even though they were tested in two positions.

Testing the remaining muscle groups revealed different results. KFL testing had the highest testing error, limit range and CoR. The increased testing error of KFL testing could be caused by the anatomical nature of the hamstring muscle. As mentioned

earlier, the hamstring muscle can produce higher torque in a seated position. High torque output combined with the large standard deviation have resulted in an inflated measurement error and CoR (Figure 6.19). Furthermore, individual variation in scores supported by the widest limit of agreement range is the main reason behind these findings (Figure 6.27). Therefore, the stability and design features of the platform are less likely to be the cause.

The higher variability in scores of the HIR (CV =8.82%) and HEX (CV =9.69%) may have occurred from the inability to isolate the primary muscles of these movements from the accessory muscles. Another possible explanation is performing HIR is awkward for some participants. Therefore, they tried to perform with maximal efforts (Kelln, et al., 2008). Since the two sessions were performed within seven days, the variability in scores may have resulted from the inconsistent effort from the participants but not from an actual gain in muscle strength (Aufsesser, et al., 1996). The fact that the CV results of the current study are less than those reported by Agre et al. (1987) and close to the values obtained by Nadler et al. (2000), Tourville et al. (2013) and Traina et al. (2010) support the use of the STAPH in assessing hip muscle strength.

Key Points

- The change in scores between the two MicroFET2 testing session of phase two was less than 5 *N.m.*
- The test-retest agreement in scores of hip muscle testing using the STAPH is not affected by the tested muscle group or plane of motion.
- Reducing the number of days between the two MicroFET2 sessions to seven days may have reduced the test-retest variability in scores.
- Resolving the issues encountered in testing using the GrAASP by developing the STAPH resulted in reduced measurement differences and error between the two consecutive MicroFET2 sessions.

7.9. *Comparison Between the Two Assessment Platforms*

The stabilising platforms were designed to improve the reliability of testing using the hand-held dynamometer. Kramer et al. (1991) showed that applying a belt-resisted HHD in testing young and elderly patients was more reliable compared to a manually-resisted dynamometer. Kim et al. (2014) also found that using a fixed HHD was more reliable than a non-fixed HHD (Table 2.15 and Table 2.16).

The GrAASP and STAPH are hip strength assessment platforms that complement the work of other researchers (Figure 7.5). The platform designed by Nadler et al. (2000) is highly reliable in assessing the HAB and HEX muscle strength. Later, Widler et al. (2009) and Traina et al. (2010) came up with simpler platforms to measure HAB muscle strength. Tourville et al. (2013) introduced a multipurpose platform to measure HAB, HAD, HFL and HEX muscle strength, which formed the designing principles of the GrAASP and STAPH. More recently, Jackson et al. (2016) utilised a PVC pipe to stabilise a HHD in testing the HAB, HAD, HER, KEX and APF. So far, the platform by Jackson et al. (2016) is the simplest and most mobile stabilising platform of all (refer to section 2.7 and Table 2.16 for details). However, none of these authors improved the platforms designs or included them in further validation and reliability studies on individuals with medical conditions.

The researcher designed the GrAASP as an objective, easy and portable measurement tool to improve clinical testing and the decision-making process. Compared to the platforms that were introduced before, the GrAASP is considered big and heavy with varying levels of testing error depending on the location of placement and method of stabilisation. Moreover, phase one of testing using the GrAASP can be classified as a pilot study to determine the best location and method of stabilisation. The STAPH was later presented with enhanced stability and adaptability to the main hip muscle groups. The validity and reliability of using the STAPH in measuring hip muscle strength of patients with hip and groin related conditions will be assessed in the future.

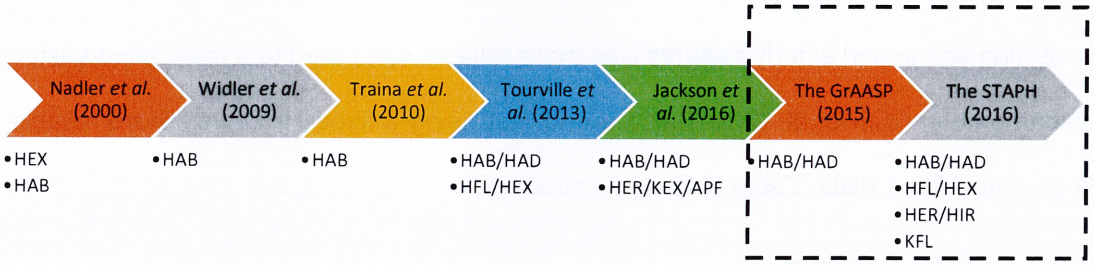


Figure 7.5 – The updated timeline of the dynamometer stabilising platforms according to the authors and the muscle groups of interest showing the designed platforms by the author in the dashed box (HFL, hip flexion; HEX, hip extension; HAB, hip abduction; HAD, hip adduction; HER, hip external rotation; HIR; hip internal rotation; KEX, knee extension; KFL, knee flexion; APF, ankle planter flexion).

The participants recruited in both phases of the study were similar in body measurements, sex and the dominant side. Also, the same recruitment and warm-up protocols were followed in both phases. Testing the hip muscles using the STAPH improved the test-retest agreement in scores in comparison to using the GrAASP. The results of STAPH testing of the HAB and HAD muscles had narrower limits of agreement and smaller error compared to the measurements obtained by the GrAASP.

The position of the participants differed in the second phase. The HAD muscles were tested in 30° of abduction in the second phase compared to 10° degrees in the first phase. The dissimilarities in the starting position may have played a role in changing the torque output of the HAD muscles (Table 6.23). Even though the findings of Biodex testing suggest that increasing the hip angle should result in higher torque, the torque values dropped in the second phase. The researcher believes that the enhanced stability of the STAPH compared to GrAASP may have minimised the undesired movements of the MicroFET2. Therefore, there were no recorded spikes during testing.

The CV values of testing using the STAPH are close to the results of Tourville et al. (2013) and Nadler et al. (2000) (Table 6.21). Testing the lower extremity muscle can have higher variability in scores (CV range 11.3% to 17.8%) compared to upper

extremity muscle testing (CV range 5.1% to 8.3%) using a non-stabilised HHD (Agre, et al., 1987). Nadler et al. (2000), who is one of the pioneers in the field of using a stabilised dynamometer, tested the HAB and HEX muscles using a wooden platform. The findings of his study indicate that testing the HEX muscles can have higher variability in scores (CV =8.06%, range: 1.84% to 27.7%) compared to the HAB muscles (CV =4.77%, range: 0.70% to 13.98%). Tourville et al. (2013) reported similar variability in scores between the sessions for the HAB (CV =8.86%), HAD (CV =8.53%), HFL (CV =8.66%) and HEX muscles (CV =21.75%). However, in a study using a portable frame to test the HAB muscles, Traina et al. (2010) reported slightly lower values for the first (CV =4.93%) and second testers (CV =3.88%).

Key Points

- Bilateral torque differences were based on the muscle, position, and mode of testing.
- Testing the hip muscles using the STAPH improved the test-retest agreement in scores in comparison to the GrAASP.
- The results of HAB and HAD testing using the STAPH had narrower limits of agreement and smaller error compared to the measurements obtained by the GrAASP.
- Both testing platforms are valid and reliable in testing, but the STAPH showed less testing error and narrower limits of agreement compared to the GrAASP.

7.10. *Study Limitations*

The results of HHD testing in both phases of the study may have been affected by several factors related to the platforms designs, testing protocols, intra-tester, intra- and inter-participants variations. In addition, there are a few limitations to the results and findings of this study in the following domains:

Platform Design

- *Stability:* The GrAASP was designed to test HAB and HAD muscles. However, when the stronger participants were tested, the dynamometer holder was unstable. Since the back surface of the MicroFET2 is curved, it occasionally wobbled during testing and both factors could have affected the measurements (Kelln, et al., 2008). Such design flaws and issues were addressed when the STAPH was constructed.
- *Design:* The STAPH was more stable and more comfortable for the participants and the tester compared to the GrAASP. Overall, some of the variability between the results of the two phases and reduced measurement error in phase two was attributed to the improved platform design and enhanced stability.
- *Uses:* Using the platforms as a method of fixation can improve the reliability of testing and reduce the measurement error. However, the clinical applications of the platforms introduced in the current study are limited to the hip and thigh area only.
- *Testing error:* The testing error associated with each platform should be considered when they are used in the clinic or future research. The author did not come across and did not specify a cut-off point (regarding testing error percentage) to what is considered a reliable measurement tool (Thorborg, et al., 2010a). As a result, the results of testing using both platforms should be considered along with the percentage of error.

Biodex Testing Protocol

- *Testing Protocol:* The high cost, size and space requirements hinder the use of isokinetic devices in clinics (Jackson, et al., 2016). The complicated protocol and position set-up are other factors that discourage the clinicians from using it (Kollock Jr, et al., 2013). During testing, it was noticed that the Biodex was very sensitive in detecting the changes in body position or movement leading to testing error alerts and repeating the procedure. After an extensive search of the literature, the researcher recognised a lack of published protocols in hip joint testing that describe the optimal testing position, joint angle and the number of repetitions.
- *Testing Position:* Testing the participants in the standing position using the Biodex may have caused early fatigue on the contralateral hip abductor muscles. Also, trying to maintain an erect position and avoiding compensatory movement could have resulted in some distraction and reduction in performance output. These issues have been reported by Kawagucki and Babcock (2010) to increase the variation of scores. Moreover, Testing the participants using the Biodex in side-lying or standing positions is not like their supine position on the GrAASP or to the activities of daily living, which may have affected the agreement of scores with the MicroFET2 in phase one of testing.
- *Muscle Isolation:* The muscle groups involved in testing were isolated by asking the participants to maintain their hips in a neutral position with their feet pointing forward. However, some participants unintentionally placed the hip joint in an externally rotated position while performing abduction, which was previously reported to increase the activity of the gluteus medius muscle (Neumann, et al., 1988). The activities of daily living such as walking up the stairs and standing from sitting are produced by the combined action of several muscles in multiple planes of motion (Neumann, 2009b). Isolating muscle groups to perform a single-plane motion in Biodex testing is not practical and may give false results.

It is agreed that knowledge about the strength of specific muscle groups is required, but real life application should also be considered in testing. It is questionable if the Biodex or isokinetic machines, in general, can still be considered a gold standard when it comes to hip isometric testing. A more straightforward, cost-effective and less complicated device needs to be produced to replace isokinetic machines in a clinical setting.

- *Gravity-Correction:* The Biodex automatically adjusts the obtained values based on the weight of the limb, which is measured during the process of setting-up the machine (Biodex, 2015). Therefore, the measured values increase if the movement was against gravity and vice-versa. Theoretically, the corrected values allow more accurate estimation of the muscle torque values (Keating & Matyas, 1996). However, the values obtained using the MicroFET2 were performed in a gravity-reduced position (Hislop, et al., 2013). Gravity-correction may partially explain the narrower agreement limits and inflated testing error between the Biodex and the MicroFET2 in HAB testing compared to HAD testing (Table 6.8 and Table 6.9).

MicroFET2 Testing Protocols

- *Inter-sessions duration:* The bigger differences in scores that were reported in the first phase may be related to the period between the sessions. The period was significantly longer in the first phase (23.8 ± 9.0 days) than the second phase (6.9 ± 1.2 days, p <0.001). The wide gap between the two MicroFET2 sessions may have resulted in motor learning, which increased the differences in scores, and reduced the reliability of measurements (Coldwells, et al., 1994).
- *Learning effect:* Unawareness of the expected task in the first MicroFET2 session in both phases may have been a contributing factor in measurement variation. Even though the testing sequence was explained using a chart and the

participants were allowed to perform a submaximal familiarisation set, it was observed that the participants followed the instructions and performed better in the second MicroFET2 session. The learning effect resembled in the familiarity with the positioning and testing protocol in the second session, can greatly influence the measurements (Coldwells, et al., 1994; Hickey, et al., 1992; Lu, et al., 2011).

- *Lever arm length:* In Biodex testing, the attachment pad of the Biodex was placed at a proximal location (above the knee), which is a logical set-up to compare against the proximal placement of the MicroFET2. However, the MicroFET2 was also placed distally and had to be compared with the Biodex measurements in the set-up mentioned above, resulting in differences in the lever arm. Krause et al. (2007) showed that a long lever arm could give significantly different results compared to the short lever arm in HAB and HAD strength testing (Table 2.12). The variation in lever arm length between the Biodex and the MicroFET2 had an impact on the calculated torque values. Thus, it may explain the large measurement error seen in Table 6.8 and Table 6.9.

Intra-Tester Variation:

- *Tester's strength:* In certain occasions, more than two participants were tested in a row. As a result, the tester may not have provided consistent resistance throughout the testing session due to fatigue. A drawback of HHD testing is that it measures peak torque at a certain point of the ROM (Sisto & Dyson-Hudson, 2007). When the tester does not sufficiently counter-resist the force produced by the participant, the HHD can underestimate muscle strength resulting in a human testing error (Lu, et al., 2011; Krause, et al., 2014). Also, the participants might have perceived that they should not push harder if they overpowered the tester (Kelln, et al., 2008; Krause, et al., 2014). Thus, suitable stabilisation should be provided by the examiner (Sisto & Dyson-Hudson, 2007).

- *Manual stabilisation:* Given the fact that force is a vector consisting of magnitude and direction, changes in the alignment of the MicroFET2 relative to the line of push could have influenced the transmission of forces to the HHD (Denton, et al., 2014). The tester had to stabilise himself with one hand and to stabilise the MicroFET2 with the other hand, which may not have been sufficient. The effect of the human testing error in this situation was reduced by using the stabilised method of testing, resulting in higher reported torques compared to the manual method (Lu, et al., 2011; Keating & Matyas, 1996).
- *Learning effect:* The necessary training in HHD testing was provided before conducting the study. However, the researcher was not familiar with this technique before. Therefore, a learning effect could have existed at least with the first three participants. It may have also had an impact on the amount of testing error during Biodex testing and the first phase of MicroFET2 testing (Aufsesser, et al., 1996).
- *Feedback:* During Biodex testing, the participants were given visual feedback on the screen about their performance and were verbally encouraged to perform better. On the other hand, they were only verbally encouraged during phase one of MicroFET2 testing, followed by informing them about the highest score after each repetition. The sound level and type of verbal feedback could have varied between the testing sessions. Verbal encouragement was avoided during phase two of MicroFET2 testing because some participants compensated and over-exhorted to achieve higher scores. The type and level of feedback have been previously found to affect the measurements (Jung & Hallbeck, 2004; Keating & Matyas, 1996).

Intra- and Inter-participants Variations:

- *Sample characteristics:* The two samples were heterogeneous in strength levels and the sport of interest. The participants played a variety of sports including general fitness training at the gym, cycling, martial arts and triathlon. Even though all the participants were university students at the time of testing, their skill levels varied. Some of the participants took part in university teams or competed on a national level, while others were merely playing sports to remain active and healthy. The strength ratios, isokinetic strength and sprint speed, can vary (Cometti, et al., 2001), and the physiological responses and techniques can fluctuate among athletes of mixed skill levels and football players (Dellal, et al., 2011; Keating & Matyas, 1996). Several participants of the current study played more than one sport, which may have also helped in developing and improving muscular force output and increasing the variations between participants.
- *Standardisation of measurement methods:* The male and female participants were both tested based on the same protocol. However, the male participants had higher stature and body mass compared to the females. Body mass has been previously found to affect HHD measurements proportionally, and the variability between the participants increased when the range of their body mass increased (Keating & Matyas, 1996). Keating & Matyas (1996) reported that the dynamometry measurements of male participants were higher than their female peers even if they were matched in age and athletic background. Thus, the differences in body measurements of the current study could have played a role in the significant differences observed in muscle strength between males and females (Table 6.5). The generalisation that females are weaker than males is supported by the sex-based results in both phases, the sensitivity analysis and previous research findings (Keating & Matyas, 1996; Andrews, et al., 1996; Neumann, et al., 1988). Since muscle strength output is directly proportional to

body mass, the obtained measurements should have been normalised to body mass to reduce between-participants variation.

- *Starting position:* While the testing protocols aimed to unify the starting position between all participants, some variation existed. The starting position in the second MicroFET2 testing session using both platforms was close to that of the first session. Even though, it was not possible to match the exact starting position, which may have affected the measurement values. The most logical explanation is that the change in position could have affected the muscles' lever arm lengths, resulting in different values and influencing the agreement in scores (Keating & Matyas, 1996).
- *Stabilisation:* The participants could stabilise themselves by holding onto the sides of the GrAASP. However, stabilisation was not allowed during STAPH testing. It was noticed that the participants scored much higher when they could stabilise the upper body and contralateral limb during testing. Thus, the author believes that stabilisation also played a role in intra-participant variation (Keating & Matyas, 1996).
- *Location:* The location of placing the MicroFET2 was set for all participants. However, due to individual variation in the contours and anatomical features of the femoral lateral condyle or the lateral malleolus, the position of the MicroFET2 had to be adjusted slightly to meet the participant's comfort. Even though the MicroFET2 is padded, it does not conform well to body parts, which leads to tenderness over the tested area (Kelln, et al., 2008). These small changes in the lever arm could have also affected the measurements (Arnold, et al., 2010).
- *Training intensity:* The participants were not asked to hold nor to reduce their training intensity due to commitment with the university sports teams or for personal reasons. The development of muscle strength and fatigue are also possible sources of variation in scores. An increase of >34 *N* (>3.5 *kg*) in strength is more likely to reflect an actual change in strength (Sisto & Dyson-Hudson,

2007). However, the sensitivity analysis of phase one showed that the increase in muscle strength was less than 34 *N* in all measurements, which is probably due to testing error or a combination of several factors.

- *Motivation:* The participants were encouraged to perform at their best level to achieve the highest force output and to obtain consistent results (Aufsesser, et al., 1996). However, since motivation is internally initiated and encouraged by external stimuli, the researcher was not able to control or modify it in every situation.
- *Outliers:* Exceptionally weak or strong muscles may influence the conclusion drawn from the analysis. Katoh et al. (2011) found a weak effect and no significant correlation between a belt-stabilised dynamometer and an IKD device in hip abduction testing (n=24, r =0.34, p >0.05). However, when the highest three measured values were excluded from the analysis, the correlation became significantly strong (n=21, r =0.65, p <0.05). It indicates that the participants with very high scores do affect the findings and interpretation of the results (Katoh, et al., 2011). In the current study, the data were analysed as a whole, and all the values were included in the analysis. If the lowest and highest measured values were excluded, the agreement scores could have differed. Such approach may be considered in future research.

Sample and Study Design

- *Sample size:* The recruitment approach applied in this study aimed at reaching the biggest number of individuals who may meet the inclusion criteria. It is recommended that a sample size of >40 should be considered for reliability studies (Lu, et al., 2011). However, more participants could not be recruited because the data collection had to be completed within a specified period and the testing labs were reserved for other research or teaching purposes.

- *Sample characteristics:* The participants involved in this study were healthy and active within a particular age group. Thus, the muscle strength reference values obtained in this study may be higher than the values of a sedentary population (Keating & Matyas, 1996). Therefore, the results of testing cannot be generalised to include a broad spectrum of the population.
- *Statistical approach:* Only the relative reliability, represented by the agreement tests, was considered in the current study. It is related to the degree to which a participant's position is maintained with repeated measures. On the other hand, absolute reliability is concerned with the extent of variation in the repeated measures of an individual (e.g. ICC). Therefore, the latter approach gives an insight into individual variations in scores and not for the whole sample (Atkinson & Nevill, 1998). Moreover, the findings of relative reliability tests make it difficult for comparison with the other studies that are based on absolute reliability tests (Thorborg, et al., 2010a).
- *Design:* The study is lab based on a sample of healthy participants under semi-controlled circumstances. It is an important point to consider if the results of this study were compared to a clinically based study, where many variables could play a role.

Key Points

- Several possible sources of testing error were observed in the study.
- The results of the study are limited to the population with similar characteristics, measurement devices, type and location of placement, and platforms.

Chapter 8 Conclusion

8.1. *Overview and Prospective Research*

Internal hip pain can be caused from a lot of pathologies such as labral tears and femoroacetabular impingement (FAI) (Samora, et al., 2011). Such conditions affect hip muscle strength, which eventually reflects on the quality of life (QoL) and activities of daily living (ADL) (Kierkegaard, et al., 2017a; Mendis, et al., 2014; Kierkegaard, et al., 2017b). Most of the pathologies and sources of pain could be treated with hip arthroscopy (HA) if physiotherapy did not show significant improvements (Quinn, 2010; Wall, et al., 2012). As a result, having the HA procedure improves pain, muscle strength, range of motion, and function to varying levels depending on the patients' characteristics and the performed procedure (Brunner, et al., 2009; Byrd & Jones, 2000; Dippmann, et al., 2014; Kemp, et al., 2014b; Nepple, et al., 2015). Longitudinal studies to examine these changes have been executed before, but they are specific to certain pathologies, groups of patients or athletes (Byrd & Jones, 2015; Daivajna, et al., 2015; Domb, et al., 2015a; Philippon, et al., 2012a).

Muscle strength changes are one of the leading aspects of HA patients' follow-up (Casartelli, et al., 2014). However, only a handful of studies looked at the muscle strength changes in the past decade (Table 2.3) (Charlton, et al., 2016; Kierkegaard, et al., 2017b; Mayne, et al., 2017). Moreover, the use of hand-held dynamometry (HHD) in the hip muscle strength assessment is biased towards total hip (THA) and total knee arthroplasty (TKA) (Alnahdi, et al., 2014; Friesenbichler, et al., 2017; Harikesavan, et al., 2017; Rasch, et al., 2010). Muscle strength testing results in the THA or TKA patients cannot be generalised to include HA patients due to the differences in the procedure, mechanics, postoperative rehabilitation and prognosis (Madsen, et al., 2004; Masonis & Bourne, 2002; Trudelle-Jackson, et al., 2002). Therefore, there is a need for more longitudinal studies examining the postoperative changes in function and range motion in general, and muscle strength in specific.

The first study of this thesis was performed to analyse the secondary data of 309 hip arthroscopy patients at 2, 8 and 24 weeks postoperatively. The main findings suggest that muscle strength, function and pain significantly improved showing the biggest changes in the 24th week postoperatively compared to preoperative measurements (Table 3.5 and Table 3.6). However, the hip range of motion (ROM) did not show statistically significant changes post-arthroscopy, which contradict previous findings that HA can improve the hip motion and mechanics (Brunner, et al., 2009; Nho, et al., 2011; Rylander, et al., 2011). The data analysis was supposed to include the 52 weeks of follow-up, but the high percentage of missing values did not permit an accurate analysis, and the data may be mistakenly interpreted. The researcher recommends that the future short- to mid-term studies should consist of more time points, and more effort should be placed in following up with the patients to maximise the amount of collected data.

The results of the first study also revealed some interesting findings. Age, sex, and level of athletic activity were found to impact postoperative changes in varying degrees. These factors can be incorporated into perspective research to study their effects on the type of internal hip pathology, preoperative status and postoperative duration of recovery. By doing so, the clinicians can have a better understanding of the inter-patient variation in recovery and prognosis, which will aid in the design of better rehabilitation protocols.

The reduction in the MHHS scores of athletes may not necessarily reflect lower function. The MHHS questionnaire lacks a sports subscale and can be affected by a ceiling effect, which may impact its uses in the athletic population (Kemp, et al., 2013; Schilders, et al., 2011). Moreover, the MHHS is limited in its capabilities as an outcome measure in predicting the patients' satisfaction (Aprato, et al., 2012). Future research should focus on developing a new outcome measure questionnaire for the athletic population who had the hip arthroscopy.

The size of the same sample can be increased in prospective research by adding the information of more patients. Also, future comparisons could be based on the level of

physical activity, sex, or duration of symptoms. A more important aspect to consider is to follow up with these patients for longer periods to better understand the postoperative changes. Fulfilling these recommendations can be challenging due to financial reasons on a patients' and research levels.

In the first study of this thesis, the pre- and post-operative hip muscle strength of the patients were assessed using a hand-held MicroFET2 dynamometer. However, the location of placing the MicroFET2 on the legs could have varied during testing. Moreover, the patients were assessed by more than two examiners with possibly ranging body builds and strengths. Consequently, the variations in testing circumstances may have affected the measurements (Krause, et al., 2014; Krause, et al., 2007).

Based on the clinical notes mentioned above and the research findings by Krause et al. (2007) and Krause et al. (2014), the second study was performed to investigate if the method and location of stabilising the MicroFET2 could have influenced the measurements. Therefore, the Groin Abductor-Adductor Strength Platform (GrAASP) was designed to stabilise the MicroFET2. The results of testing using a manually held and fixed MicroFET2 using the GrAASP were compared with the Biodex. It was established that applying a fixed method of stabilisation at a distal location results in muscle strength scores closer to those obtained on the Biodex with less measurement error. As a result, the Hip Strength Assessment Platform (STAPH) was designed as an enhanced version of the STAPH to test seven muscle groups using a fixed method of stabilisation at a distal location. The STAPH resulted in less measurement error and differences in scores compared with the GrAASP and can be used to test the muscle strength of healthy and active individuals.

The second study findings suggest a recommendation to test the hip abductor muscles (HAB) from 0° of abduction and to test the hip adductor muscles (HAD) from 30° of abduction to obtain the highest torque values. Another aspect of the study showed that stabilising the HHD at a distal location gives less testing error and variation in scores

compared to manual stabilisation at a proximal location. Moreover, the GrAASP was found to be a valid and a reliable stabilising platform with specific measurement errors that need to be considered. Lastly, the STAPH provided better testing reliability with lower measurement error and stronger agreement in scores compared to the GrAASP. The main outcomes of this study and the clinical applications are further mentioned in sections 8.2 and 8.4.

The findings of the second study will set the grounds for future research related to HHD testing. The validity and reliability of the platforms were assessed in a group of a healthy and active sample of participants, and the results cannot be generalised to include patients. Therefore, the researcher is aiming to evaluate the validity and reliability of the STAPH on hip arthroscopy patients to look at if the platform can be utilised in a clinical setting on a group of patients or not. Furthermore, a similar study can be performed on patients with groin pain and knee injuries. Completing these two studies will increase the clinical uses of the STAPH and enhance the objectivity of clinical testing.

Prospective Research

The MHHS along with other hip outcome measure questionnaires are widely used in the Western countries with patients undergoing hip arthroscopy and THA. However, they are rarely applied in the clinical settings of Middle Eastern countries. As a result, another future research project of the author is based on translating and cross-validating the hip and lower extremity related questionnaires such as the MHHS, HOS and HAGOS into Arabic. The study will first be performed on a group of hip arthroscopy and THA patients, who speak Arabic and English very well since English is the second spoken language in most Middle Eastern countries. The results will vastly improve the evidence-based practice of the Arabic-speaking therapists and empower the use of objective measurement tools in various clinical settings.

The use of objective outcome measures such as the HHD and questionnaires will have a direct impact on the clinical practice of the Arab and Middle Eastern therapist (Abrams, et al., 2006). Another study should be performed to study the impact of applying these measures in clinical practice. Once the systematic assessment and follow-up protocols are established and implemented, a database of hip arthroscopy procedures and postoperative outcomes based on a Middle Eastern population will be created. Once completed, the data of these patients will be analysed to look at the role of race, physical activity awareness and the environment on the presence of internal hip pathologies and prognosis. The results may help us understand the prevalence of femoroacetabular impingement and the effects of these factors on a Middle Eastern sample of patients, which was not done before.

8.2. *Novel Contribution and Main Outcomes*

Study One

The first study of this thesis was performed to analyse the secondary data of 309 hip arthroscopy patients. It complements the work of Casartelli et al. (2014), Philippon et al. (2009) and Kemp et al. (2014b). Various aspects of the postoperative changes in the physical and functional status of these patients were assessed at 2, 8 and 24 weeks postoperatively. The novel contribution and main outcomes of this study in relation to the research questions (RQ) can be summarised as:

1. The pattern of clinical changes differs from a group of patients to another based on multiple factors (McCarthy, et al., 2011; McCarthy & McMillan, 2013). Also, the currently available studies are limited regarding the sample size, characteristics, and outcome measures (Daivajna, et al., 2015; Kemp, et al., 2014b). The current study is one of a few studies to include a sample of heterogeneous patients regarding age groups, the level of athletic activity and duration of symptoms.

2. It is also one of the few studies to consider the use of the flexion-abduction-external rotation (FABER) test distance as an outcome measure for hip arthroscopy patients.
3. It is the first study to use multi-level modelling to examine within- and between-patients variation in scores and the factors that affect the postoperative outcome.
4. RQ 1: The reference values of this sample of heterogeneous patients were reported. Previous research studies were either based on a specific category of hip arthroscopy patients such as older patients in the study by Kemp et al. (2014b), or a specific diagnosis such as femoroacetabular impingement in the study by Casartelli et al. (2014), or a case study similar to the work by Philippon et al. (2009).
5. RQ 2: Currently, there is a lack of short- to medium-term longitudinal studies. Most available studies have followed patients for over a year (Brunner, et al., 2009; Charlton, et al., 2016). The present study is one of the few, recent, medium-term studies that looked at postoperative changes in muscle strength, range of motion (ROM) and function from preoperative or baseline measurements to a period of 24 weeks postoperatively. The findings suggest that muscle strength, function and pain require a minimum of 8 weeks postoperatively to show significant improvements compared to baseline measurements, which was not discussed before. In addition, the most prominent postoperative changes were observed in the 24th week postoperatively (Table 3.5 and Table 3.6). However, the ROM of the patients did not show statistically significant changes, which contradict previous findings that HA can improve the hip ROM and mechanics (Brunner, et al., 2009; Nho, et al., 2011; Rylander, et al., 2011).
6. RQ 3 and 4: The variation within- and between-patients, and the factors that influence postoperative changes were also reported. Older patients showed reduced hip ROM and flexibility compared to the other age groups. Therefore, significant ROM improvements should not be expected from patients aged more

60 years old. Regarding sex, the females were characterised by reduced strength and external rotation ROM but increased internal rotation ROM. These findings are not surprising since females with various lower extremity pathologies were found to have reduced hip muscle strength and altered kinematics compared to males (Baldon, et al., 2009; Fredericson, et al., 2000; Fukuda, et al., 2010).

Another aspect of the factors analysis showed that athletes possessed stronger muscles and scored lower on the MHHS compared to amateur and sedentary patients. It has been previously found that athletes of different sports are stronger and respond differently to exercises than nonathletes (Ahtiainen, et al., 2003; Siqueira, et al., 2002). In this study, the athletes also demonstrated that they respond much better to postoperative rehabilitation than the other groups.

Study Two

The second study combined multiple aspects of HHD testing of the hip and thigh muscles. The novel contribution and main outcomes of this study, which are linked to the research questions (RQ), can be summarised as:

1. Testing the hip abductor and adductor muscles strength is underrated in the literature. It is the first study to simultaneously test the hip abductor and adductor muscles isometrically based on three angles of hip abduction and isokinetically using two speeds of testing.
2. The use of HHD in muscle strength testing is affected by the examiner's strength and method of stabilisation (Krause, et al., 2014; Krause, et al., 2007). It is the first study to include the evaluation of two new dynamometer stabilising platforms designed to test hip and thigh muscle strength.
3. Janda (2014) described the upper and lower cross syndromes. In this research, a pattern of hip extensor-external rotator muscles weakness was identified,

which is similar to the upper and lower cross syndromes but was not described before in the literature.

4. RQ 1: The reference values of hip muscle strength or torque lack in the literature. The reference torque values of the two samples of healthy and active individuals were provided using the Biodex, two methods of stabilisation, two locations of placements and two stabilising platforms.
5. RQ 2: The number of studies looking into bilateral torque differences is limited (Daneshjoo, et al., 2013; Jacobs, et al., 2005). In this study, there were no statistically significant differences between the torques of the dominant and non-dominant sides when measured on the Biodex.
6. RQ 3: The starting position of the hip can affect the measurements. The only study that investigated this issue was described more than two decades ago by Neumann et al. (1988). The results from Biodex testing suggest a recommendation to test the hip abductor and hip adductor muscle groups from a position of 0° and 30° of abduction respectively.
7. RQ 4: It is not recommended to test the hip abductors isokinetically due to the inflated testing error and variation in scores.
8. RQ 5: The study is one of the few studies to describe the hip adduction to abduction strength ratios in healthy active individuals using the Biodex. Moreover, four hip and thigh muscle strength ratios were obtained using the STAPH. The strength ratios from Biodex testing were dependent on the angle of abduction and speed of testing, while the ratios obtained from GrAASP testing were dependent on the fixation method and location of placement. It is the first research to consider the changes in hip adduction-abduction and other muscles strength ratios based on these factors.
9. RQ 6 and 7: The findings agree with and support the results of Krause et al. (2007) and Krause et al. (2014) in the effects of stabilisation and location of placement on the outcome of HHD testing. One of the main findings of the study

supports the use of a fixed stabilisation method at a distal location to obtain higher consistency in the measurement of muscle strength.

10. RQ 8 and 9: Two new hip muscle strength assessment platforms were introduced to assist the clinicians in improving the objectivity of testing compared to manual muscle testing and HHD testing. Both platforms were valid and reliable, but testing using the STAPH has reduced measurement error and differences in scores compared to the GrAASP.

8.3. *Overall Thesis Limitations*

The methodology applied in both studies aimed at collecting as much data as possible within the scope of the thesis. The analysis was performed using statistical procedures to examine the data in multiple ways. However, the overall limitations of this thesis can be summarised as:

- *Sample size:* In study one, the sample size of the hip arthroscopy patients reduced during the analysis due to missing values in the dataset. In the second study, the sample size also reduced after excluding the results of the participants who dropped out. However, it was not possible to include the data of more patients (in study one) or to recruit more participants (in study two) because of time constraints and lab availability.
- *Sample characteristics:* The samples of both studies were heterogeneous on different levels. The patients of study one consisted of a wide range of age groups and may have had various pathologies. The diagnostic and surgical details were not available for analysis. Similarly, the participants of the second study played different sports. Therefore, the reader should be aware of these issues when the results are considered clinically.
- *Outcome measures:* The Modified Harris Hip Score may not be the best outcome measure for the professional athletes due to the ceiling effect in

measuring the changes in physical activity and function. In study two, the Biodex was very sensitive to the slightest changes in the participant's position, which often resulted in an error, repeating the test and increasing the test duration. These two outcome measures should be included within a battery of tests.

- *Literature:* There is a lack of published Biodex protocols to test the hip muscles isometrically. Also, comparing the results of Biodex testing to previous research was limited due to the reduced number of available and accessible studies related to this area of testing.
- *Data collection and testing error:* In study one, manually transferring the patient's data from a computer software to an excel sheet may not have been entirely accurate. In study two, the strength differences between the examiner and the participants along with the stabilisation issues of the GrAASP may be the reason behind the inflated measurement error.
- *Standardisation:* The obtained values of both studies could have been standardised to body mass and stature to reduce the variation between participants, but such information was not provided with the hip arthroscopy dataset and were not applied in study two.
- *Absolute reliability:* Study two of this thesis involved the statistical tests of relative reliability only. Absolute reliability tests such as ICC was not included.
- *Generalisability:* The findings of study one can be generalised to a wide range of hip arthroscopy patients. On the other hand, the reliability results of study two are limited to healthy and active participants within a particular age group.

8.4. *Recommendations and Clinical Applications*

Study One

The two studies included in this thesis differ in nature. Study one is considered more clinical, while study two was lab-based and involved a controlled sample of healthy participants. The findings of the first study suggest a recommendation that the rehabilitation programs post hip arthroscopy should be designed to accommodate the time it takes for the muscles to show a noticeable improvement, function to improve and pain to reduce. Moreover, the variation between patients of different sexes, age group and level of athletic activity should be considered.

Older age was found to affect the postoperative range of motion while being a female can affect the postoperative muscle strength, the range of motion, function and pain measurements. Additionally, a professional level of athletic activity results in higher muscle strength measurements. Lastly, the preoperative duration of symptoms did not influence the postoperative measurement. Although some of these factors were analysed before, it is the first study to analyse the effects of these factors using within- and between-patients' variation in scores via multilevel modelling.

The postoperative changes can take a minimum of 8 weeks to show noticeable improvement. Based on the sample of patients that was analysed, hip external and internal rotation range of motion did not statistically change. However, gains in range of motion have been reported clinically (Brunner, et al., 2009). The wide range of age groups, which includes older patients, could have affected these results (Table 3.3). The hip joints of older patients may have been undergoing arthritic changes resulting in reduced internal and external range of motion. To better understand the reflection of postoperative clinical changes on the range of motion, the data can be re-analysed by classifying the patients into age groups. It will allow direct comparison between these groups to understand if the range of motion changes are based on the age group or not.

Also, the degree of arthritic changes can be included in a sensitivity analysis to determine its impact on the outcome.

From a rehabilitation perspective, the exercises prescribed to females should aim to improve muscle strength and control the excessive internal rotation motion as a precautionary measure to avoid developing further musculoskeletal pathologies (Cometti, et al., 2001; Willson, et al., 2012). Also, it is advisable that the female patients be given strengthening and mobility control exercise before the arthroscopy. It will be beneficial to train them on the correct movement patterns ahead of time.

The differences in outcome based on the level of athletic activity is vital to consider when the rehabilitation protocols for athletes are designed. Once this type of patients meets the minimal criteria to progress in the rehabilitation, they can be pushed harder (within the clinical and surgical recommendations), and return to sport quicker (Bolia, et al., 2017; Voight, et al., 2010). Therefore, their rehabilitation programs and exercises should be designed separately than the others. The treating physiotherapist should not expect equal strength development and functional changes from patients with different sexes or physical activity levels.

The researcher believes that a combination of multiple questionnaires should be considered to get a clear picture on how athletes, amateur players and inactive patients progress. Moreover, adding functional balance, agility and strength tests to pre- and postoperative assessment is vital to determine participants' eligibility to go back to sport.

Study Two

The findings of the second study can be clinically applied in multiple ways. First, the MicroFET2 or other HHDs should be placed at a distal location when testing the hip and thigh muscles. Second, it is preferable to stabilise the MicroFET2 to overcome the strength differences between the clinician and the patient, and to improve the objectivity of testing. Third, the validity and reliability of using the STAPH in testing the hip and

thigh muscle strength of hip arthroscopy patients should be assessed before the results of testing using the STAPH are considered in this population.

The results of this research indicate no significant differences between the dominant and non-dominant sides of two hip muscle groups when tested using the Biodex and six out of seven muscle groups when tested using the STAPH. Orthopaedic surgeons, physiotherapist, athletic trainers and other related professionals can quickly scan one side to determine the baseline level of muscle strength. The findings also suggest that the starting position can have a crucial role in changing the strength output. Therefore, the hip abductor muscles should be assessed at 0° of abduction, and there were no statistically significant differences between the three angles in hip abduction. In contrast, hip adductor muscles should be tested at 30° of abduction due to the increased torque production in this angle and position of the muscle fibres. These can be considered the best positions for the muscle fibres.

The female participants in both phases of the second study had a statistically significant reduction in muscle strength values compared to their male peers. Such findings agree with previous studies (Saadat, et al., 2014; Suzuki, et al., 2015; Willson, et al., 2012). The reduced strength of females put them at risk of developing patellofemoral syndrome, iliotibial band syndrome and other lower extremity pathologies. These findings reiterate the importance of assessing the hip muscle strength of female patients complaining of knee pain.

The torque reference values of the hip abductor and adductor muscle groups and the remaining muscles collected in this research are of relevance to the age and level of athletic activity group studied. Moreover, providing such values in three different angles of isometric testing, two speeds of isokinetic testing, two modes of stabilisation and locations allow for comparison with other studies and provide a reasonable basis for future research.

The muscle strength ratios obtained in this research can assist clinicians in determining if the tested muscles suffer from weakness or not. A potential weakness in the hip adductors, for example, can predispose long-distance runners to developing groin injuries (Tyler, et al., 2010). Even though the ratios should be considered, a full screening of the hip and thigh muscles is needed to detect other hidden issues.

Many still consider isokinetic machines to be a gold standard in muscle strength assessment. However, the high costs, immobility and complexity of these devices will hinder their clinical use. Furthermore, contractions under isokinetic conditions do not reflect true muscular activity in real-world contexts. It is questionable if they should still be considered a gold standard when new techniques or tools are developed. Such a conclusion does not necessarily mean that the Biodex should not be used in testing at all. However, a cheaper, portable and reliable method of testing such as the GrAASP and STAPH can be used as an alternative to isokinetic machines.

Based on the results of this research and comparative studies, clinicians are advised to use a fixed method of stabilisation, which supports the use of the platform to measure HAB and HAD muscle strength using the MicroFET2. Although this research was based on a robust fixed method, a belt or other objects could also be used provided that they have sufficient stability. In addition to the stabilisation method, a distal location of placement should be used in testing. It will lead to more consistent results, less testing error and a smaller difference in scores. Also, it will give the assessor mechanical advantage over the tested individuals.

The GrAASP and STAPH were both found to be valid and reliable in testing hip muscle strength. Based on the summarised comparison in Table 6.23, the researcher recommends using the STAPH more than the GrAASP due to the reduced testing error, higher consistency in the results, and smaller differences between two consecutive sessions. The use of these platforms in testing hip muscle strength is recommended while taking into consideration the different issues involved in testing.

Appendix A Chapter 2 Supporting Material

Appendix A. 1 - Hip Muscular Anatomy

The major muscle groups include the hip abductor, adductor, extensor, flexor, and rotator muscles. In the thigh, the quadriceps and the hamstring muscles, which are primary knee joint muscles, can also play a role in the hip movement. The hip abductor muscle consists of the gluteus medius and gluteus minimus muscles forming about 60% and 20% respectively of the entire abductor cross section. The tensor fasciae latae is the smallest muscle of the three (Neumann, 2009b). These muscles originate from the lateral aspect of the iliac bone and insert into the lateral side of the greater trochanter of the femur. The tensor fasciae latae crosses the hip and knee joints and inserts in the iliotibial tract of the tibia (Appendix A. 2) (Moore, et al., 2006). The anterior fibres of gluteus medius and minimus also work as internal rotators in the horizontal plane especially at 90° of hip flexion (Neumann, 2009b).

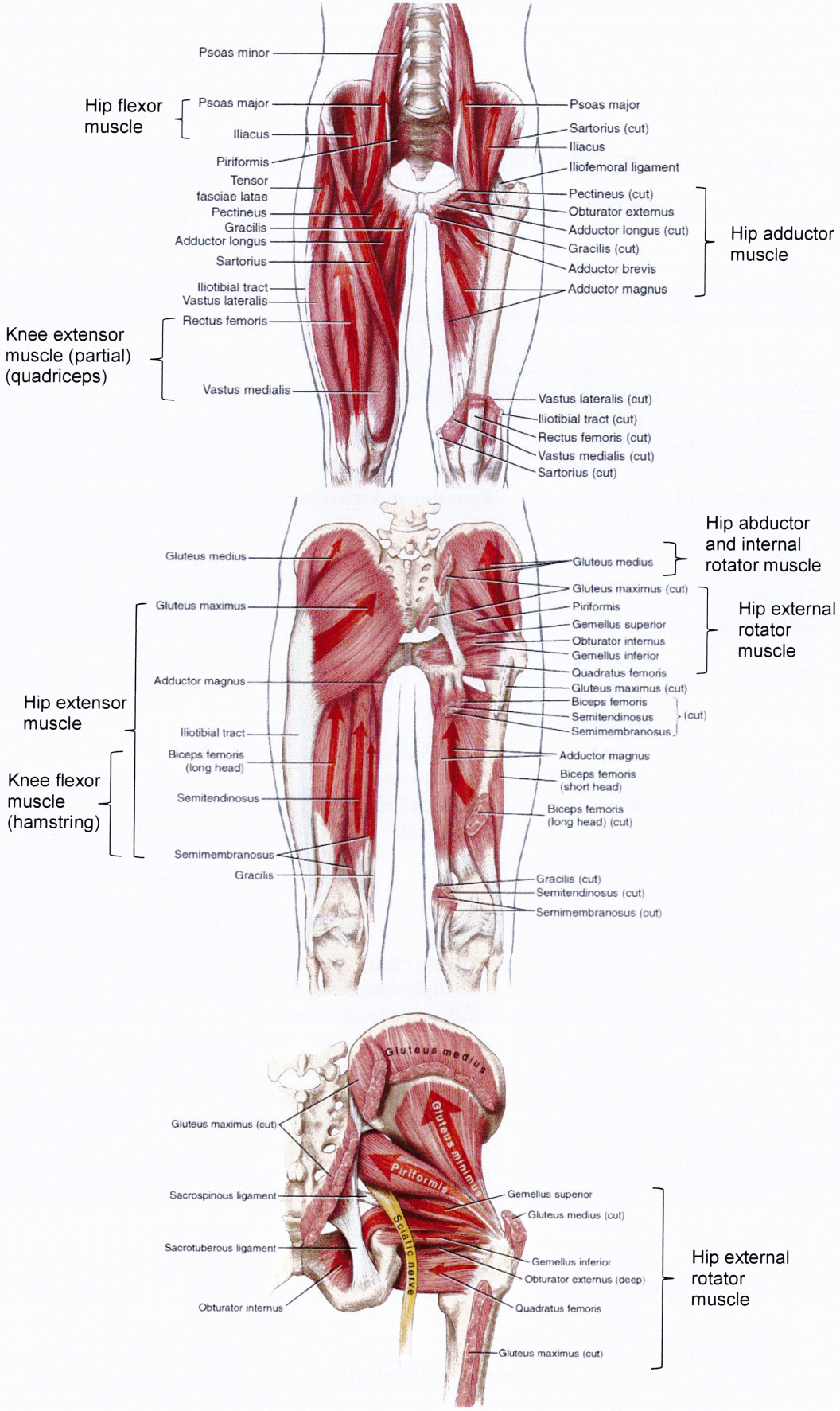
The adductor muscles group consist of the adductor magnus, adductor longus, adductor brevis, gracilis, and pectineus muscles. They originate from the superior and inferior rami of the pubic bone and insert into the linea aspera of the femur except the gracilis muscle, which inserts medially into the proximal part of the tibial shaft (Appendix A. 2) (Moore, et al., 2006).

The hip flexor muscles lay deep within the abdomen and pelvic region. The hip flexor muscles consist of the iliopsoas, sartorius, tensor fasciae latae, rectus femoris, adductor longus, and pectineus. The iliopsoas muscle consists of the iliacus and the psoas major muscles. The iliacus muscle originates from the iliac fossa, whereas the psoas major originates from the transverse processes of lower thoracic and lumbar vertebrae. The fibres of both muscles partially blend and insert medially into the lesser trochanter of the femur (Neumann, 2009b). The sartorius muscle originates from the anterior superior

iliac spine (ASIS) and extends downward towards the leg inserting into the medial surface of the proximal tibia (Appendix A. 2) (Neumann, 2009b).

The gluteus maximus, the hamstrings (i.e. semitendinosus, semimembranosus and long head of the biceps femoris) and the posterior head of the adductor magnus are considered the primary hip extensors. The gluteus maximus, which is also considered the primary external rotator of the hip, originates from multiple posterior locations in the ilium, sacrum, coccyx, and the posterior sacroiliac ligaments. It inserts in the iliotibial band and the gluteal tuberosity of the femur. It has a crucial role in sacral, pelvic and hip stability. The hamstring muscle, which comprises of three muscles, originates from the ischial tuberosity and inserts into the proximal parts of the tibia and fibula. Therefore, it acts to extend the hip and flex the knee joints (Appendix A. 2) (Neumann, 2009b).

The anterior fibres of the gluteus medius and minimus serve as the hip internal rotators. The hip external rotator muscles include the gluteus maximus with six short and deep muscles, which are the piriformis, obturator internus, obturator externus, gemellus superior, gemellus inferior and quadratus femoris. The smaller muscles are aligned horizontally to compress and stabilise the hip along with producing movement. The piriformis originates from the sacrum and inserts superiorly into the greater trochanter. On the other hand, the obturator internus, which is a fan-like muscle, starts from the obturator membrane and surrounding bones of the obturator foramen and passes on top of the lesser sciatic notch inserting into the medial surface of the greater trochanter. The twin muscle, the obturator externus, starts from the external side of the obturator membrane and the ilium close-by and ends in the trochanteric fossa of the femur. The gemellus superior and inferior (in *Latin*, meaning twins) start on either side of the lesser sciatic notch and merge with the obturator internus for a common insertion. Just below these muscles, the quadratus femoris muscle originates from the ischial tuberosity and inserts into the femur proximally (Appendix A. 2) (Neumann, 2009c). Further description of the hip and thigh anatomy is described elsewhere (Drake, et al., 2005; Martini & Nath, 2008; Neumann, 2009b).



Appendix A. 2 – Anatomy of the anterior hip muscles and thigh (top), posterior hip and thigh muscles (middle), and the deep short muscles (bottom) (Neumann, 2009b).

Name_____ Date_____

MODIFIED HARRIS HIP SCORE

Please answer the following questions as they pertain to your hip:

Pain:

☐None/Able to ignore it
☐Slight, occasional, no compromise in activity
☐Mild, no effect on ordinary activity, pain after usual activity, use aspirin/ibuprofen/Tylenol
☐Moderate, tolerable, makes concessions, occasional narcotic
☐Marked, serious limitations
☐Totally disabled

Function: Gait

Limp
☐None
☐Slight
☐Moderate
☐Severe
☐Unable to walk

Support
☐None
☐Cane for long walks
☐Cane all the time
☐Crutch
☐2 canes
☐2 crutches
☐Unable to walk

Distance Walked
☐Unlimited
☐6 blocks
☐2-3 blocks
☐Indoors only
☐Bed and chair

Functional Activities

Stairs
☐Can go up/down normally
☐Can go up/down normally with banister
☐Any method
☐Unable

Socks/Shoes
☐With ease
☐With difficulty
☐Unable

Sitting
☐Any chair, 1 hour
☐High chair, ½ hour
☐Unable to sit, ½ hour, any chair

Public Transportation
☐Able to enter public transportation
☐Unable to use public transportation (such as bus, or airport transportation)

For Internal Use:

Score_____
Initials_____

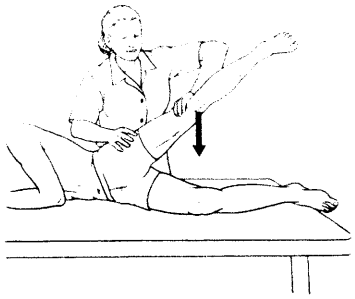
Appendix A. 3 – The Modified Harris Hip Score (MHHS) (ISHA, 2017)

Appendix A. 4 - Manual Muscle Testing

Manual muscle testing (MMT) is one of the easiest and least time-consuming assessment methods. Most physiotherapists and orthopaedic consultants apply it on a daily basis. The application of MMT is not limited to orthopaedic, neurology, physiotherapy, occupational therapy, or sports clinics. The findings of MMT can be used to obtain a general sense of the integrity of the neuromuscular systems and to reach a differential diagnosis. Additionally, it is used to design specific rehabilitation programs, monitor the progress of patients, and determine the discharge status (Frese, et al., 1987; Wadsworth, et al., 1987).

- *Grading System Criteria*

Three grading systems of muscle strength were developed. The grading system started with a numerical score from zero to five (no activity to strongest) that is given to individual muscles (Appendix A. 5). However, Daniels and Worthingham introduced the grading criteria to score all the muscles contributing to a certain motion (Appendix A. 6) (Hislop & Montgomery, 2007).



Appendix A. 5 - Manual muscle testing of the hip abductors (Grade 5) (Hislop, et al., 2013).

Later, Florence Kendall developed a percentage system to grade muscle strength from 0% to 100% (weakest to strongest) (Kendall, 1993). Details of MMT application for all the muscles can be found in the Muscle Testing book by Hislop et al. (2013).

Appendix A. 6 - Muscle Grading Systems and Definitions (Frese, et al., 1987; Hislop & Montgomery, 2007)

Numerical system	Daniels and Worthingham	Kendall and McCreary	Definitions
5	Normal (N)	100%	Able to hold the test against gravity and maximum resistance or move the part into the test position and hold against gravity and maximum pressure
4	Good (G)	80%	Same as normal but able to hold against moderate resistance
3+	Fair+ (F+)	-	Able to hold the position against gravity, or to move the part into the test position and hold against gravity and slight resistance
3	Fair (F)	50%	Able to hold the position against gravity, or to move the part into the test position and hold against gravity
2+	Poor+ (*P*+)	-	Able to move the part through full range with gravity eliminated, but against slight resistance
2	Poor (*P*)	20%	Able to move the part through full range with gravity eliminated
2-	Poor- (*P*-)	-	Able to move the part through partial range with gravity eliminated
1	Trace (T)	5%	Muscle contraction can be palpated
0	0	0%	No contraction can be elicited

- Reliability of Manual Muscle Testing

The number of studies looking into the reliability of MMT is limited. The subjectivity and personal judgment in scoring may have discouraged the use of it in research (Wadsworth, et al., 1987). In 1954, a group of researchers noticed that the muscle score grades of 0 to 5 obtained by 12 to 39 examiners in four different attempts were within the same grade (Lilienfeld, et al., 1954; Frese, et al., 1987). A few years later, another group of researchers reached a comparable conclusion when ten examiners had similar ratings in 89% to 92% of the trials based on each muscle (Appendix A. 7) (Iddings, et al., 1961; Frese, et al., 1987).

A study was performed on thirteen participants with chronic orthopaedic and neuromuscular problems. The results showed that MMT has good intra-rater reliability for shoulder abduction (r =0.98), hip flexion (r =0.74), and knee flexion (r =0.63) (Appendix A. 7) (Wadsworth, et al., 1987). The inter-rater (inter-examiner) reliability of

middle trapezius and gluteus medius muscle testing was investigated in another study. Patients (n=100) with various neuromusculoskeletal conditions were examined by eleven physiotherapists who graduated from three different schools. Each therapist assessed ten patients followed by testing ten other patients that were assessed by another therapist (Total=20). The results showed that 28% to 45% of the patients were rated the same score by two examiners when both muscles on both sides were assessed. When the weighted Cohen's Kappa scores (κ) were calculated, the results were κ=0.58 for the right middle trapezius, κ=0.29 for left middle trapezius, κ=0.25 for the right gluteus medius and κ=0.11 for left gluteus medius, indicating poor reliability for testing these two muscles (Appendix A. 7) (Frese, et al., 1987).

In a related study, the results of testing thirty-seven consecutive patients showed reduced consistency in repeating the test for the upper extremity (r =0.55-0.63), lower extremity (r =0.36-0.88), and bilateral upper and lower extremities (r =0.01-0.88) (Appendix A. 7) (Bohannon, 1997b). In the studies by Frese et al. (1987), Wadsworth et al. (1987) and Bohannon (1997b), the patients suffered from various medical conditions. The reduced correlation in scores could be due to the inability to reproduce the same amount of muscular contraction in the consecutive testing sessions.

The effect of experience on the inter-tester reliability of MMT was also investigated. The results of deltoid and psoas major MMT by a fifth-year chiropractic student were compared to the results of an experienced practitioner. The analysis showed strong inter-tester reliability with an agreement of 82.08% (κ=0.62) of the time when the deltoid muscle was tested and an agreement of 84.9% (κ=0.67) for the psoas major muscle. These findings indicate that both the experienced and novice practitioners have good inter-tester agreement in scoring (Appendix A. 7) (Pollard, et al., 2005). Therefore, experience does not have a significant impact on scoring muscle strength. The main criticism of this study concerns the grading criteria. The examiners did not specify the exact grade of muscle strength (e.g. 3, 3+, 4 or 5) and relied on classifying the muscles

as strong or weak. There would have been more variation in scoring if the typical scoring method was followed.

- *Validity of Manual Muscle Testing*

Lovett and Martin (1916) described the use of MMT to test the muscles of infantile polio patients. Their results gave an indication of the status of the anterior horn of the spinal cord, which gives rise to the motor branch of the spinal nerve. The anterior horn is affected by the virus in polio patients (Martin & Lovett, 1915; Lovett & Martin, 1916).

The use of MMT in diagnostic procedures was later expanded to include orthopaedic patients. Looking at the discriminant validity of MMT and HHD in identifying the presence of knee extensor muscles weakness, Bohannon (1997a) measured knee extension strength of thirty-eight consecutive home-care patients complaining from various internal, neurological and musculoskeletal systems disorders. The sensitivity of MMT to detect the presence of side-to-side difference was found to be 42.9% only compared to the HHD, which was used as a reference. Regarding specificity, MMT was able to detect the absence of weakness in 82.3% of the cases. The agreement between MMT and HHD in detecting side-to-side differences was κ =0.24. Both methods of testing were able to detect the presence of weakness. However, the MMT was inferior compared to the HHD (Appendix A. 7) (Bohannon, 1997a).

The concurrent validity* of MMT demonstrated that MMT could overestimate what is considered a normal muscle strength (Bohannon, 1986). Forty-one consecutive patients were evaluated by two blinded examiners to examine the inter-test reliability and the convergent validity of upper limb muscle MMT. They rated the strength of muscles of symptomatic and asymptomatic sides as either normal or reduced. The range of kappa coefficient was 0.25 to 0.60 for the shoulder muscles, 0.57 to 0.72 for the arm muscles,

* Measures how well an assessment tool agrees and correlates with another tool that has been validated before (Cuthbert & Goodheart, 2007).

and 0.33 to 0.69 for the forearm muscles. The odd ratio to detect actual weakness in the symptomatic side was 4.0, 95% CI [2.5, 7.7]. The results indicate better consistency in testing the arm muscles compared to the shoulder and forearm muscle groups. Moreover, MMT can be used in the diagnostic process to detect weakness (Appendix A. 7) (Jepsen, et al., 2004).

The validity and reliability of MMT have not been investigated much. The latest literature review reported correlation coefficients ranging from 0.63 to 0.98 for individual muscle groups and from 0.57 to 1.0 for total MMT scores. The review also showed that MMT has content, construct, and discriminant validity concerning assessing the muscle strength of various conditions (Cuthbert & Goodheart, 2007). A summary of the reviewed studies can be found in Appendix A. 7.

- *Issues Related to MMT*

The application of manual muscle testing can be influenced by its subjectivity and reduced sensitivity. In clinical diagnosis, sensitivity refers to the ability of a test to correctly detect the presence of a disease (Lalkhen & McCluskey, 2008). Since it was introduced as a clinical assessment technique, MMT was critiqued by many as having a high degree of subjectivity and variation in scoring between practitioners. Testing the muscles graded as Fair+ (3+) and above relies largely on the estimation of force applied by the examiner. Therefore, classifying muscle strength as Fair+, Good or Normal is based on the relative strength of the examiner and patient, together with the patient's age and size (Frese, et al., 1987; Wadsworth, et al., 1987). It was reported that strength scoring differences of 20%-25% were observed when MMT was used. Moreover, the children who were considered to have normal strength using MMT had just 50% of the strength of what is seen as normal when objective strength measurement was used (Bohannon, 2002).

Several factors can influence grading muscle strength. First, the testing method can vary from one examiner to another based on their educational background and experience. Other factors that may play a role are: (1) location of force application; (2) magnitude, speed and direction of force application; (3) the period of sustained contraction; (4) amount of cooperation from the patient; (5) amount of interaction between the examiner and patient; (6) various distractions; (7) instructions from the examiner and tone of voice; (8) experience of therapist; (9) fatigue; and (10) amount of stabilisation and substitution by other muscular activity (Frese, et al., 1987; Pollard, et al., 2005; Sapega, 1990; Wadsworth, et al., 1987).

The adequacy of MMT as a screening test to detect changes in strength is also questioned. A comparison was made between MMT and hand-held dynamometry testing of hospital patients who were admitted for acute rehabilitation (n=107; age: 62.1 ± 16.4) to look at the presence of weakness and side-to-side differences. The diagnostic accuracy of MMT compared to the HHD ranged from 72% to 77.6%. Furthermore, the sensitivity of MMT to detect the presence of weakness compared to the normal values ranged from 60.9% to 70.3%, which increased when the HHD showed more difference. It was also found that the accuracy of MMT (70.3% to 73.6%) increased when 25% to 30% reduction in weakness was detected using the HHD (Bohannon, 2005). Lincoln, et al. (2003) recommended that a test should have a sensitivity of more than 80% to be considered a good diagnostic measure or a sensitivity of more than 90% if it was a screening measure (Bohannon, 2005). However, MMT does not meet these requirements. Strain gauges and dynamometers were developed and used as part of the clinical muscle strength testing procedures to overcome the issues faced in MMT.

Appendix A. 7 - Overview of the studies investigating the validity and reliability of mainly lower extremities muscle strength testing using MMT.

Study (first author, year)	Study Details			Sample Details		Results	Main Findings and Comments
	Muscles tested	Validity assessment	Reliability assessment	Number and sex	Age in (M±SD) or range in years		
(Bohannon, 1997a)	Various upper and lower extremity muscles.	Yes (with a HHD)	No	8 male and 30 female patients with chronic orthopaedic, neurological and medical conditions (n=38).	75.6±10.4	Agreement between both methods (Kappa coefficient) = 0.240. A disagreement of 31.6% between both methods in detecting bilateral difference. MMT had a sensitivity of 42.9% and specificity of 82.3%.	MMT is less discriminate compared to HHD.
(Bohannon, 1997b)	SAB, EFL, HFL, KEX and ADF.	No	Yes (inter-tester)	15 male and 22 female patients with chronic orthopaedic, neurological and medical conditions (n=37).	36 to 94	Inter-tester reliability (Spearman r_s): Upper extremity muscles=0.55-0.63. Lower extremity muscles=0.36-0.88. Bilateral upper and lower extremities=0.01-0.88.	Inconsistent results in re-testing due to heterogeneity of the sample.
(Florence, et al., 1992)	KEX, HFL, SER, HAB, HEX, SAB, EEX, NFL, NEX, EFL, ADF, KFL, AEV, AIN, APF, TAB, WEX and WFL.	No	Yes (intra-tester)	Males with Duchenne's muscular dystrophy (n=102)	9.1±2.7	Intra-tester reliability (Weighted Kappa coefficient): KEX=0.93, HFL=0.90, SER=0.89, HAB=0.89, HEX=0.88, SAB=0.87, EEX=0.86, NFL=0.85, NEX=0.84, EFL=0.82, ADF=0.81, KFL=0.79, AEV=0.73, AIN=0.72, APF=0.71, TAB=0.71, WEX=0.69 and WFL=0.65.	MMT on proximal muscles in more reliable than distal muscles in patients with Duchenne's muscular dystrophy.
(Frese, et al., 1987)	MT and HAB (both sides)	No	Yes (inter-tester)	60 male and 50 female patients with chronic orthopaedic, neurological and medical conditions (n=110).	41±15	Inter-tester reliability (Weighted Kappa coefficient): MTT=0.29-0.58 and HAB=0.11-0.25. Agreement in rating was 28%-45% for the four muscles and 89%-92% for each muscle.	Poor reliability of using MMT in testing the MDT and HEX muscles.

(Jepsen, et al., 2004)	Various upper extremity muscles (details NA).	No	Yes (inter-tester)	Males with upper extremity conditions (*n*=22).	25 to 61	Inter-tester reliability (Unweighted Kappa coefficient): Shoulder muscles=0.25 to 0.60, arm muscles=0.57 to 0.72 and forearm muscles=0.33 to 0.69.	High variability in MMT scoring.
(Pollard, et al., 2005)	SFL and HFL.	No	Yes (inter-tester)	50 males and 56 females (*n*=106).	19 to 63	Inter-tester reliability (Unweighted Kappa coefficient and agreement) SFL=0.62 (82.08%) and HFL=0.67 (84.9%).	Experienced and novice examiners have good agreement in scores. The muscles were not graded based on the MMT grading criteria. Therefore, the results are questionable.
(Wadsworth, et al., 1987)	SAB, EEX, WEX, HFL and KFL.	No	Yes (intra-tester for MMT and HHD).	Patients with chronic orthopaedic and neurological conditions (*n*=13).	21 to 96	Intra-tester reliability (Pearson *r*) for MMT: SAB=0.98, EEX=NA, WEX=NA, HFL=0.74 and KFL=0.63. Intra-tester reliability (Pearson *r*) for HHD: SAB=0.69, EEX=0.90, WEX=0.88, HFL=0.72 and KFL=0.75.	MMT was more reliable in certain tests but it was less sensitive in detecting change in strength compared to the HHD.

M, mean; SD, standard deviation; NA, not available; HHD, hand-held dynamometer; NFL, neck flexors; NEX, neck extensors; MT, middle trapezius; SAB, shoulder abductors; SER, shoulder external rotators; EFL, elbow flexors; EEX, elbow extensors; WEX, wrist extensors; WFL, wrist flexors; TAB, thumb abductors; HFL, hip flexors, HEX, hip extensors; HAB, hip abductor; HAD, hip adductor; HER, hip external rotator; HEX, hip extensor; KEX, knee extensor; KFL, knee flexor; APF, ankle planter flexor; ADF, ankle dorsi flexor; AEV, ankle eversion; AIN, ankle inversion.

Appendix A. 8 – Isokinetic Devices

Isokinetic devices (IKDs) are considered a gold standard when it comes to muscle strength testing. The main aim behind introducing isokinetic devices was the need for a quantitative and accurate measurement of muscle strength rather than only assessing the integrity of the neuromuscular system. This method of testing is not exclusive to evaluating the strength of patients only. It was also used as a pre-screening test for players before competitions to determine their readiness, and it is being used in lawsuits and medical cases to determine the eligibility for compensation or classification as disabled (Amundsen, et al., 1990).

- *Reliability and Validity of Isokinetic Devices*

In a study looking at the reliability of isokinetic testing of the hip flexors, extensors, adductors and abductors, twenty-nine healthy boys (age range: 6 to 10 years) had their muscles tested at $30°.sec^{-1}$ and $90°.sec^{-1}$. The data analysis showed higher reliability for flexion (ICC=0.63 and ICC=0.75) and extension testing (ICC=0.68 and ICC=0.84) compared to abduction (ICC=0.59 and ICC=0.59) and adduction testing (ICC=0.55 and ICC=0.49) at $30°.sec^{-1}$ and $90°.sec^{-1}$ respectively. Further statistical results were not provided. The variability in the results could be due to the age of participants, the absence of a pre-set end point for the range of motion, inadequate familiarisation repetitions, and having two testers involved in the study (Burnett, et al., 1990).

The studies performed on adult participants reported higher reliability values. The Biodex was used in a study to test concentric knee extension and flexion peak torque at $60°.sec^{-1}$, $180°.sec^{-1}$, $240°.sec^{-1}$, and $300°.sec^{-1}$. Nineteen healthy and active participants (age range: 20-35 years) volunteered for the study. The results based on ICC for peak torque were in the range of 0.95 to 0.97 for the knee extensors, and 0.82 to 0.98 for the knee flexors (Feiring, et al., 1990).

In a similar study on 18 healthy participants (age: 22 ± 3 years), the Kin-Com 500H (Isokinetic International, TN, USA) was used to assess the knee extensor and flexor muscles isokinetic torque at 60°.sec^{-1} over two days. The reliability values were high for concentric extension (ICC =0.95, 95% CI [0.85, 0.98]) and flexion contractions (ICC =0.94, 95% CI [0.83, 0.98]). Eccentric testing showed slightly less reliability in extension (ICC =0.94, 95% CI [0.85, 0.98]) and flexion (ICC =0.92, 95% CI [0.79, 0.97]). The Bland-Altman plots for the agreement in scores between days 1 and 2 showed higher concentric extensor and flexor torque values on day 2 (Sole, et al., 2007). Nevertheless, the isokinetic device showed high reliability in testing the knee flexors and extensors.

The reliability of isokinetic testing was also assessed on the hip muscles. The results of a study showed that adduction eccentric testing was more reliable (Rt, ICC =0.85, 95% CI [0.68, 0.93]; Left: ICC =0.84, 95% CI [0.67, 0.92]) compared to concentric testing (Rt, ICC =0.67, 95% CI [0.41, 0.83]; Left: ICC =0.60, 95% CI [0.32, 0.79]). The reliability of flexor muscle testing was considerably lower in eccentric (Rt, ICC =0.28, 95% CI [0.09, 0.66]; Left: ICC =0.35, 95% CI [0.04, 0.62]) and concentric testing (Rt, ICC =0.37, 95% CI [0.06, 0.64]; Left: ICC =0.04, 95% CI [0, 0.35]) (Emery, et al., 1999). The repeated eccentric-concentric contraction cycles may have been difficult to accomplish by the participants. Isolating such muscles and performing eccentric contractions is not a common task. Combining these two contractions should be avoided in testing.

The previous models of the Biodex (before 2003) were mechanically reliable and valid. Later, the Biodex System 3 model was introduced as a modern dynamometer machine based on an electronically controlled motor with better control of acceleration rates and velocity (Drouin, et al., 2004). In a study to assess the validity and reliability of the newer version, a 72.5 cm lever arm set at 90° to the ground, was attached to the dynamometer. A calibrated 4.55 kg weight was placed at the end of the lever arm, and testing was performed at speeds ranging from 30°.sec^{-1} to 500°.sec^{-1}. A hand-held inclinometer was set as the criterion measure to assess the validity. The results showed very high trial-to-trial reliability for the position (*ICC* =0.99 to 1.00, *SEM* =0.00-0.60°), torque (*ICC*

=0.99 to 1.00, *SEM* =0.00-0.39 *N·m)*, and velocity (*ICC* =0.99 to 1.00, *SEM* =0.00-12.89°.sec^{-1}). Additionally, the day-to-day reliability was also high (*ICC* =0.99). Regarding validity, the coefficient of variation (CV) between the two measurement devices was 0%-3% for the position, 0%-2% for the torque and 1%-4% for the velocity (Drouin, et al., 2004). These findings indicate that the Biodex System 3 isokinetic dynamometer is a mechanically reliable instrument for testing angular position, isometric torque and slow to moderately high velocities (< 300°.sec^{-1}). A summary of the reviewed studies discussing the reliability and validity of IKD can be found in Appendix A. 9.

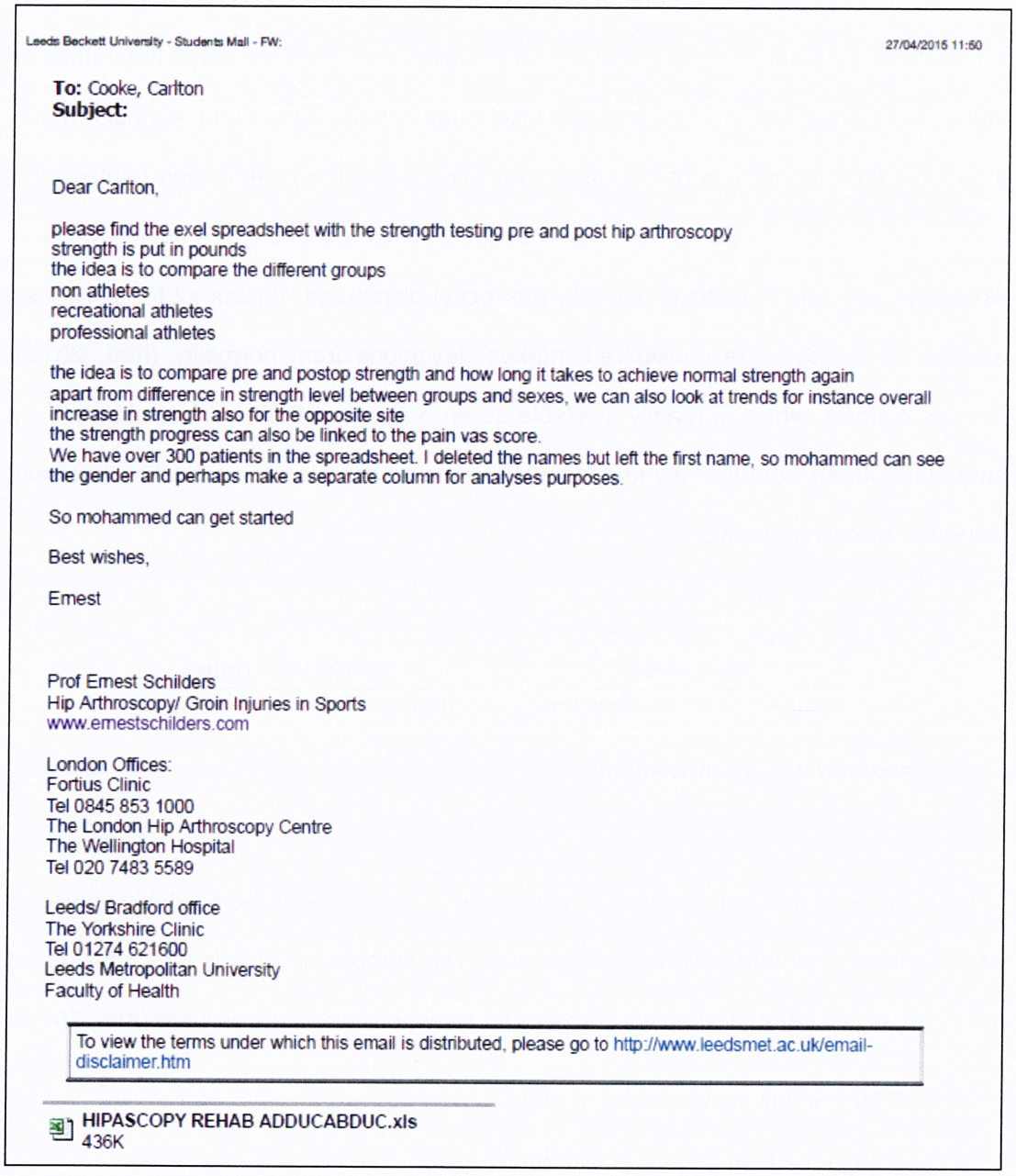
The fact that IKDs are considered a gold standard of muscle strength testing may have resulted in the reduced number of validity and reliability studies on healthy individuals. Many of the available and accessible studies are typically old and were performed on older models of IKDs. In contrast, the recent studies focus on the application of isokinetic testing in very specific groups such as those with knee ligaments laxity, stroke, or chronic obstructive pulmonary disease. The primary purpose of reviewing IKDs in this section was to the shed the light on their high reliability and why they are classified as a gold standard. Therefore, it is out of the scope of this review to go beyond the studies that were performed on hip and thigh muscles. Further details can be found in the literature (Zapparoli & Riberto, 2016).

Appendix A. 9 - Overview of the studies investigating the reliability of lower extremities muscle strength testing using IKDs.

Study (first author, year)	Study Details			Sample Details		Results	Main Findings and Comments
	IKD Model	Muscles tested	Reliability assessment	Number and sex	Age in (M±SD) or range in years		
(Burnett, et al., 1990)	NA	HFL, HEX, HAD and HAB.	Yes (inter-session)	Males (*n*=9).	6 to 10	Testing at 30°.sec^{-1} (ICC): HFL=0.63, HEX=0.68, HAB=0.59 and HADD=0.55. Testing at 90°.sec^{-1} (ICC): HFL=0.75, HEX=0.84, HAB=0.59 and HADD=0.49.	Testing at a higher speed is more reliable. High variability due to age of the participants and absence of a familiarisation set.
(Drouin, et al., 2004)	Biodex System 3	NA	Yes, trial-to-trial and day-to-day (and validity)	NA	Na	Reliability (ICC): Using a hand-held inclinometer was set as the criterion measure to assess validity. Trial-to-trial reliability for the position =0.99 to 1.00 (SEM=0.00-0.60°), torque=0.99 to 1.00 (SEM=0.00-0.39 N·m), and velocity=0.99 to 1.00 (SEM =0.00-12.89°.sec^{-1}). Day-to-day reliability was also high (ICC = 0.99). Validity (CV): Position= 0%-3%, torque= 0%-2% and velocity= 1%-4%.	Biodex System 3 is a very reliable and valid isokinetic device.
(Emery, et al., 1999)	Cybex Norm	HFL and HAD	Yes	19 males.	21 to 43	[Reliability based on ICC, 95% CI] Eccentric: HAD Rt= 0.85 [0.68, 0.93], HAD Lt= 0.84 [0.67, 0.92], HFL Rt= 0.28 [0.09, 0.66] and HFL Lt= 0.35 [0.04, 0.62]. Concentric: HAD Rt= 0.67 [0.41, 0.83], HAD Lt= 0.60 [0.32, 0.79], HFL Rt=0.37 [0.06, 0.64] and HFL Lt= 0.04 [0, 0.35].	Hip adduction and flexion testing using an IKD is reliable.
(Sole, et al., 2007)	Kin-Com 500H	KEX and KFL.	Yes (inter-session)	11 males and 7 females (*n*=18).	males=20±1, females=22±3	Concentric (ICC [95% CI]): KEX=0.95 (0.85-0.98) and KFL=0.94 (0.83-0.98). Eccentric (ICC [95% CI]): KEX=0.94 (0.85-0.98) and KFL=0.92 (0.79-0.97).	Kin-Com is a highly reliable isokinetic device to measure knee extension and flexion peak torque.

M, mean; SD, standard deviation; NA, not available; IKD, isokinetic dynamometer; Rt, right; Lt, left; HAB, hip abductors; HAD, hip adductors; HEX, hip extensors; HFL, hip flexors; KEX, knee extensor; KFL, knee flexor; ICC, intra-class correlation.

Appendix B Chapter 3 Supporting Material



Leeds Beckett University - Students Mail - FW: 27/04/2015 11:50

To: Cooke, Carlton
Subject:

Dear Carlton,

please find the exel spreadsheet with the strength testing pre and post hip arthroscopy
strength is put in pounds
the idea is to compare the different groups
non athletes
recreational athletes
professional athletes

the idea is to compare pre and postop strength and how long it takes to achieve normal strength again
apart from difference in strength level between groups and sexes, we can also look at trends for instance overall increase in strength also for the opposite site
the strength progress can also be linked to the pain vas score.
We have over 300 patients in the spreadsheet. I deleted the names but left the first name, so mohammed can see the gender and perhaps make a separate column for analyses purposes.

So mohammed can get started

Best wishes,

Ernest

Prof Ernest Schilders
Hip Arthroscopy/ Groin Injuries in Sports
www.ernestschilders.com

London Offices:
Fortius Clinic
Tel 0845 853 1000
The London Hip Arthroscopy Centre
The Wellington Hospital
Tel 020 7483 5589

Leeds/ Bradford office
The Yorkshire Clinic
Tel 01274 621600
Leeds Metropolitan University
Faculty of Health

To view the terms under which this email is distributed, please go to http://www.leedsmet.ac.uk/email-disclaimer.htm

HIPASCOPY REHAB ADDUCABDUC.xls
436K

Appendix B. 1 - Approval from Prof Schilders to analyse the secondary hip arthroscopy dataset.

Appendix B. 2 - Normality Assessment

The *z*-scores were calculated from skewness and kurtosis values to objectively quantify the magnitude of deviation from a normal distribution. Skewness, which measures how symmetrical both sides of the curve are, is positive when the right tail of the values in the curve is longer than the left side and vice-versa. On the other hand, kurtosis, which measures the peakedness of the curve, is positive when the peak is high (leptokurtic) and negative when the peak of the curve is relatively flat (platykurtic). A curve with zero skewness and zero kurtosis is normally (perfectly) distributed. Values >2 for skewness and >7 for kurtosis are considered marked deviations from normality (Kim, 2013). Excess kurtosis, which is readily available as an output in SPSS, can be obtained by subtracting three from the raw kurtosis value. The *z*-scores were calculated manually using the following equation:

$$z = \frac{Skew\ value}{Standard\ Error\ of\ Skewness}, \quad z = \frac{Excess\ Kurtosis\ value}{Standard\ Error\ of\ Kurtosis}$$

Appendix B. 3 - z-scores equation

The sample size was determined as the number of patients with the available values in each variable and not the whole sample size. The sample or variable was considered non-normally distributed when the *z*-scores for skewness and kurtosis were (Kim, 2013):

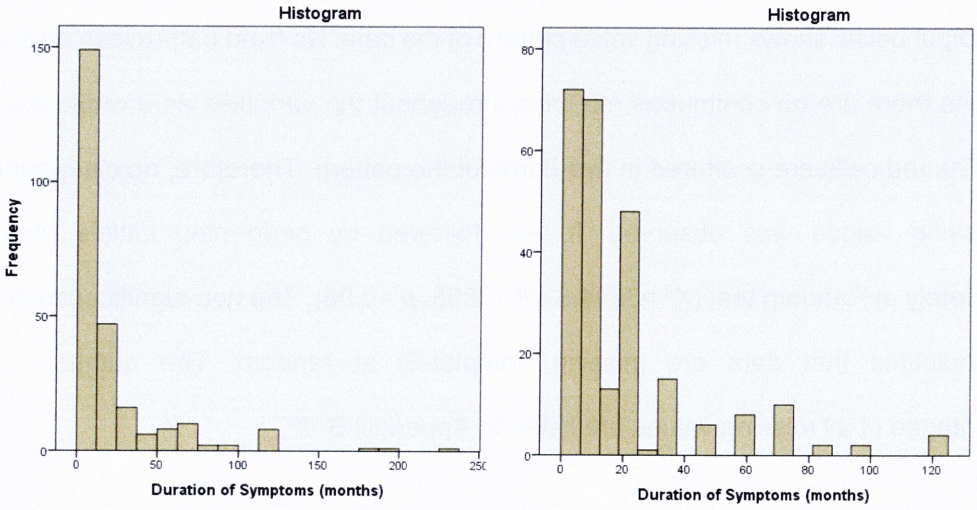
- >1.96 in small sample sizes (n <50)
- >3.29 in medium sample sizes (50< n <300)
- >2 for skewness or >7 for kurtosis are considered marked deviation from normality in large sample sizes (n >300).

The criteria above was used to reach a conclusive decision about the normality of distribution. Therefore, the conclusion was based on subjective (visual), and objective (Shapiro-Wilk and *z*-score) tests to avoid bias.

Appendix B. 4 - Outliers

Exploring outliers was performed statistically and visually. The 5% trimmed mean was calculated and compared to the mean. Any value less than 90% was further explored. When the outliers were identified, they were replaced by the average value.

The only variable that required outliers trimming was the duration of symptoms (23.95 ± 32.95 months). It was decided to change the values of 13 patients (5% of 251) to 24 months (the average). The new mean of the duration of symptoms was reduced to 20.97 ± 23.18 months (Appendix B. 5).



Appendix B. 5 - Histogram of the duration of symptoms (left) and duration of symptoms after 5% mean trimming (right).

Appendix B. 6 - Missing Data

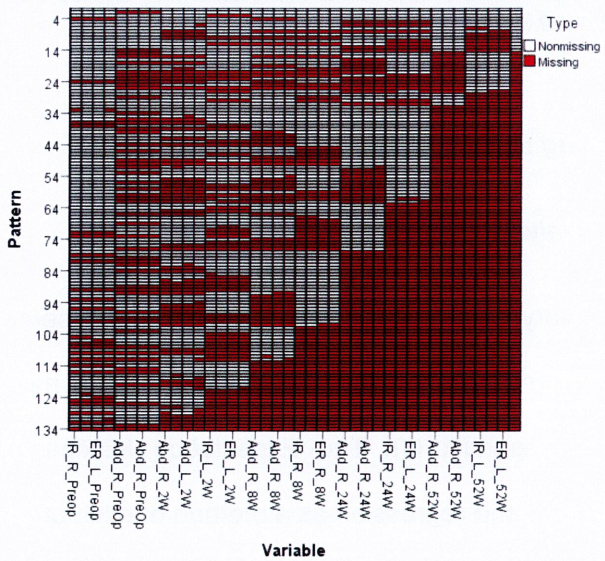
Missing Data Analysis

The first step of secondary data analysis was looking into the percentages of missing values and pattern. The null hypothesis was that values are missing completely at random (explained in Appendix G). The Expectation Maximisation was selected between listwise, pairwise, and regression as a method of estimation.

Percentages of Missing Values

The analysis showed that the highest missing values were found in abduction strength on the left side at the 52nd week (92.4%, *n*=292) and external rotation range of motion on the left side at the 52nd week (91.5%, *n*=289). Regarding the patients' characteristics, 1.3% (*n*=4) of the cases had missing age and 20.6% (*n*=65) in the duration of symptoms. For the categorical variables, there were no missing values. It was decided that variables with missing value percentages exceeding 90% will be excluded from the analysis. Excluding these values means that the analysis will include all time points except the 52nd week.

The output below shows missing value pattern of the data. No fixed pattern can be found because there are no continuous red lines throughout the variables as a whole except one. The red cells are scattered in two-thirds of the pattern. Therefore, no clear pattern in missing values was observed. It was followed by performing Little's Missing Completely at Random test (X^2 =3204.8, *df* =3208, *p* >0.05). The non-significance of the test indicates that data are missing completely at random. The numbers and percentages of all missing values are listed in Appendix B. 8.



Appendix B. 7 - Pattern of Missing Values (labelled in red).

Appendix B. 8 - Missing values analysis of the hip arthroscopy data.

	Cases					
	Valid		Missing		Total	
	n	Percent	n	Percent	n	Percent
Gender	316	100.0%	0	0.0%	316	100.0%
Level of Activity	316	100.0%	0	0.0%	316	100.0%
Age (years)	312	98.7%	4	1.3%	316	100.0%
Duration of Symptoms (months)	251	79.4%	65	20.6%	316	100.0%
Operation Side	316	100.0%	0	0.0%	316	100.0%
Dominant Side	316	100.0%	0	0.0%	316	100.0%
Adduction (R) Preoperative	222	70.3%	94	29.7%	316	100.0%
Adduction (L) Preoperative	222	70.3%	94	29.7%	316	100.0%
Adduction (R) at 2 Weeks	201	63.6%	115	36.4%	316	100.0%
Adduction (L) at 2 Weeks	201	63.6%	115	36.4%	316	100.0%
Adduction (R) at 8 Weeks	163	51.6%	153	48.4%	316	100.0%
Adduction (L) at 8 Weeks	162	51.3%	154	48.7%	316	100.0%
Adduction (R) at 24 Weeks	99	31.3%	217	68.7%	316	100.0%
Adduction (L) at 24 Weeks	99	31.3%	217	68.7%	316	100.0%
Adduction (R) at 52 Weeks	28	8.9%	288	91.1%	316	100.0%
Adduction (L) at 52 Weeks	28	8.9%	288	91.1%	316	100.0%
Abduction (R) Preoperative	222	70.3%	94	29.7%	316	100.0%
Abduction (L) Preoperative	222	70.3%	94	29.7%	316	100.0%
Abduction (R) at 2 Weeks	202	63.9%	114	36.1%	316	100.0%
Abduction (L) at 2 Weeks	195	61.7%	121	38.3%	316	100.0%
Abduction (R) at 8 Weeks	162	51.3%	154	48.7%	316	100.0%
Abduction (L) at 8 Weeks	161	50.9%	155	49.1%	316	100.0%
Abduction (R) at 24 Weeks	99	31.3%	217	68.7%	316	100.0%
Abduction (L) at 24 Weeks	98	31.0%	218	69.0%	316	100.0%
Abduction (R) at 52 Weeks	28	8.9%	288	91.1%	316	100.0%
Abduction (L) at 52 Weeks	24	7.6%	292	92.4%	316	100.0%
Internal Rotation (R) Preoperative	271	85.8%	45	14.2%	316	100.0%
Internal Rotation (L) Preoperative	271	85.8%	45	14.2%	316	100.0%
Internal Rotation (R) at 2 Weeks	194	61.4%	122	38.6%	316	100.0%
Internal Rotation (L) at 2 Weeks	195	61.7%	121	38.3%	316	100.0%
Internal Rotation (R) at 8 Weeks	155	49.1%	161	50.9%	316	100.0%
Internal Rotation (L) at 8 Weeks	155	49.1%	161	50.9%	316	100.0%
Internal Rotation (R) at 24 Weeks	83	26.3%	233	73.7%	316	100.0%

	Cases					
	Valid		Missing		Total	
	n	Percent	n	Percent	n	Percent
Internal Rotation (L) at 24 Weeks	82	25.9%	234	74.1%	316	100.0%
Internal Rotation (R) at 52 Weeks	28	8.9%	288	91.1%	316	100.0%
Internal Rotation (L) at 52 Weeks	28	8.9%	288	91.1%	316	100.0%
External Rotation (R) Preoperative	268	84.8%	48	15.2%	316	100.0%
External Rotation (L) Preoperative	270	85.4%	46	14.6%	316	100.0%
External Rotation (R) at 2 Weeks	195	61.7%	121	38.3%	316	100.0%
External Rotation (L) at 2 Weeks	195	61.7%	121	38.3%	316	100.0%
External Rotation (R) at 8 Weeks	155	49.1%	161	50.9%	316	100.0%
External Rotation (L) at 8 Weeks	155	49.1%	161	50.9%	316	100.0%
External Rotation (R) at 24 Weeks	83	26.3%	233	73.7%	316	100.0%
External Rotation (L) at 24 Weeks	83	26.3%	233	73.7%	316	100.0%
External Rotation (R) at 52 Weeks	27	8.5%	289	91.5%	316	100.0%
External Rotation (L) at 52 Weeks	27	8.5%	289	91.5%	316	100.0%
Faber Test (R) Preoperative	265	83.9%	51	16.1%	316	100.0%
Faber Test (L) Preoperative	265	83.9%	51	16.1%	316	100.0%
Faber Test (R) at 2 Weeks	202	63.9%	114	36.1%	316	100.0%
Faber Test (L) at 2 Weeks	202	63.9%	114	36.1%	316	100.0%
Faber Test (R) at 8 Weeks	167	52.8%	149	47.2%	316	100.0%
Faber Test (L) at 8 Weeks	169	53.5%	147	46.5%	316	100.0%
Faber Test (R) at 24 Weeks	88	27.8%	228	72.2%	316	100.0%
Faber Test (L) at 24 Weeks	86	27.2%	230	72.8%	316	100.0%
Faber Test (R) at 52 Weeks	27	8.5%	289	91.5%	316	100.0%
Faber Test (L) at 52 Weeks	27	8.5%	289	91.5%	316	100.0%
VAS Preoperative	245	77.5%	71	22.5%	316	100.0%
VAS at 2 Weeks	200	63.3%	116	36.7%	316	100.0%
VAS at 8 Weeks	150	47.5%	166	52.5%	316	100.0%
VAS at 24 Weeks	82	25.9%	234	74.1%	316	100.0%
VAS at 52 Weeks	29	9.2%	287	90.8%	316	100.0%
mHHS Preoperative	132	41.8%	184	58.2%	316	100.0%
mHHS at 8 Weeks	91	28.8%	225	71.2%	316	100.0%
mHHS at 24 Weeks	43	13.6%	273	86.4%	316	100.0%
mHHS at 52 Weeks	18	5.7%	298	94.3%	316	100.0%
Distraction Test	316	100.0%	0	0.0%	316	100.0%

Appendix B. 9 - Normality tests results of the hip arthroscopy data.

	Shapiro-Wilk df	Shapiro-Wilk Sig.	Histogram	Q-Q Plot	Skewness	Skewness SE	Skewness z-scores	Kurtosis	Kurtosis SE	Kurtosis z-scores	Conclusion	%	Test
Adduction (R) Preoperative	222	< .001	NT	N	0.625	0.163	3.825	-0.250	0.325	-0.768	Not Normal	20% normally distributed	Non-Parametric
Adduction (L) Preoperative	222	< .001	NT	N	0.561	0.163	3.436	0.114	0.325	0.350	Not Normal		
Adduction (R) at 2 Weeks	201	< .001	NT	NT	0.893	0.172	5.207	1.186	0.341	3.475	Not Normal		
Adduction (L) at 2 Weeks	201	.001	NT	N	0.630	0.172	3.676	0.701	0.341	2.054	Not Normal		
Adduction (R) at 8 Weeks	163	.001	N	N	0.640	0.190	3.365	0.393	0.378	1.040	Not Normal		
Adduction (L) at 8 Weeks	162	< .001	NT	N	0.835	0.191	4.378	1.077	0.379	2.839	Not Normal		
Adduction (R) at 24 Weeks	99	< .001	NT	N	1.060	0.243	4.369	1.909	0.481	3.973	Not Normal		
Adduction (L) at 24 Weeks	99	.001	N	N	0.687	0.243	2.831	0.009	0.481	0.019	Normal		
Adduction (R) at 52 Weeks	28	> .05	-	-	0.372	0.441	0.844	-0.464	0.858	-0.540	Normal		
Adduction (L) at 52 Weeks	28	.002	NT	NT	1.070	0.441	2.428	0.221	0.858	0.258	Not Normal		
Abduction (R) Preoperative	222	.001	N	N	0.447	0.163	2.736	-0.269	0.325	-0.826	Normal	60% normally distributed	Parametric
Abduction (L) Preoperative	222	< .001	NT	N	0.567	0.163	3.471	-0.080	0.325	-0.247	Not Normal		
Abduction (R) at 2 Weeks	202	< .001	N	N	0.653	0.171	3.814	0.597	0.341	1.754	Not Normal		
Abduction (L) at 2 Weeks	195	< .001	NT	N	0.790	0.174	4.539	0.714	0.346	2.060	Not Normal		
Abduction (R) at 8 Weeks	162	.010	N	N	0.541	0.191	2.837	0.411	0.379	1.085	Normal		
Abduction (L) at 8 Weeks	161	.017	N	N	0.494	0.191	2.585	0.270	0.380	0.709	Normal		
Abduction (R) at 24 Weeks	99	.002	NT	NT	0.834	0.243	3.437	0.922	0.481	1.919	Not Normal		
Abduction (L) at 24 Weeks	98	.016	NT	N	0.713	0.244	2.926	0.908	0.483	1.879	Normal		
Abduction (R) at 52 Weeks	28	> .05	-	-	0.062	0.441	0.142	-0.486	0.858	-0.566	Normal		
Abduction (L) at 52 Weeks	24	> .05	-	-	0.371	0.472	0.786	-1.108	0.918	-1.207	Normal		
Internal Rotation (R) Preoperative	271	.031	N	N	0.305	0.148	2.059	0.362	0.295	1.226	Normal	90% normally distributed	Parametric
Internal Rotation (L) Preoperative	271	.005	N	N	0.405	0.148	2.740	0.670	0.295	2.271	Normal		
Internal Rotation (R) at 2 Weeks	194	> .05	-	-	0.154	0.175	0.880	-0.246	0.347	-0.707	Normal		
Internal Rotation (L) at 2 Weeks	195	> .05	-	-	0.126	0.174	0.724	-0.137	0.346	-0.394	Normal		
Internal Rotation (R) at 8 Weeks	155	> .05	-	-	0.267	0.195	1.373	-0.184	0.387	-0.475	Normal		
Internal Rotation (L) at 8 Weeks	155	> .05	-	-	0.473	0.195	2.426	0.780	0.387	2.013	Normal		
Internal Rotation (R) at 24 Weeks	83	.016	NT	N	0.789	0.264	2.985	2.326	0.523	4.451	Normal		
Internal Rotation (L) at 24 Weeks	82	< .001	NT	N	2.024	0.266	7.617	10.961	0.526	20.853	Not Normal		
Internal Rotation (R) at 52 Weeks	28	> .05	-	-	-0.625	0.441	-1.418	0.517	0.858	0.602	Normal		
Internal Rotation (L) at 52 Weeks	28	> .05	-	-	0.003	0.441	0.006	0.367	0.858	0.427	Normal		
External Rotation (R) Preoperative	268	> .05	-	-	0.251	0.149	1.690	0.674	0.297	2.272	Normal	100% normally distributed	Parametric
External Rotation (L) Preoperative	270	> .05	-	-	0.032	0.148	0.213	0.021	0.295	0.073	Normal		
External Rotation (R) at 2 Weeks	195	> .05	-	-	0.229	0.174	1.316	0.165	0.346	0.475	Normal		
External Rotation (L) at 2 Weeks	195	> .05	-	-	0.101	0.174	0.583	0.375	0.346	1.081	Normal		
External Rotation (R) at 8 Weeks	155	> .05	-	-	0.233	0.195	1.197	0.583	0.387	1.506	Normal		
External Rotation (L) at 8 Weeks	155	> .05	-	-	0.068	0.195	0.347	0.200	0.387	0.516	Normal		
External Rotation (R) at 24 Weeks	83	.004	N	NT	0.292	0.264	1.105	1.300	0.523	2.488	Normal		
External Rotation (L) at 24 Weeks	83	> .05	-	-	-0.388	0.264	-1.468	0.687	0.523	1.315	Normal		
External Rotation (R) at 52 Weeks	27	> .05	-	-	0.759	0.448	1.695	1.266	0.872	1.452	Normal		
External Rotation (L) at 52 Weeks	27	> .05	-	-	0.414	0.448	0.925	1.178	0.872	1.350	Normal		
Faber Test (R) Preoperative	265	> .05	-	-	-	-	-	-	-	-	Normal	100% normally distributed	Parametric
Faber Test (L) Preoperative	265	.011	N	N	0.286	0.150	1.907	-0.363	0.298	-1.218	Normal		
Faber Test (R) at 2 Weeks	202	.031	N	N	0.298	0.171	1.743	0.980	0.341	2.874	Normal		
Faber Test (L) at 2 Weeks	202	> .05	-	-	-	-	-	-	-	-	Normal		
Faber Test (R) at 8 Weeks	167	> .05	-	-	-	-	-	-	-	-	Normal		
Faber Test (L) at 8 Weeks	169	> .05	-	-	-	-	-	-	-	-	Normal		
Faber Test (R) at 24 Weeks	88	> .05	-	-	-	-	-	-	-	-	Normal		
Faber Test (L) at 24 Weeks	86	> .05	-	-	-	-	-	-	-	-	Normal		
Faber Test (R) at 52 Weeks	27	> .05	-	-	-	-	-	-	-	-	Normal		
Faber Test (L) at 52 Weeks	27	> .05	-	-	-	-	-	-	-	-	Normal		
VAS Preoperative	245	< .001	NT	N	-0.804	0.156	-5.154	0.308	0.310	0.994	Not Normal	20% normally distributed	Non-Parametric
VAS at 2 Weeks	200	< .001	NT	N	0.501	0.172	2.913	-0.395	0.342	-1.155	Normal		
VAS at 8 Weeks	150	< .001	NT	NT	0.990	0.198	5.000	0.607	0.394	1.541	Not Normal		
VAS at 24 Weeks	82	< .001	NT	NT	0.950	0.266	3.571	-0.014	0.526	-0.027	Not Normal		
VAS at 52 Weeks	29	.001	NT	NT	1.161	0.434	2.675	0.740	0.845	0.876	Not Normal		
mHHS Preoperative	132	.004	N	N	-0.630	0.211	-2.986	1.578	0.419	3.766	Normal	50% normally distributed	Parametric
mHHS at 8 Weeks	91	< .001	NT	NT	-0.906	0.253	-3.581	0.996	0.500	1.992	Not Normal		
mHHS at 24 Weeks	43	.002	NT	NT	-0.756	0.361	-2.094	-0.305	0.709	-0.430	Not Normal		
mHHS at 52 Weeks	18	> .05	-	-	-	-	-	-	-	-	Normal		

*. This is a lower bound of the true significance / a. Lilliefors Significance Correction

SE = Standard Error, N = Normal distribution, NT = Not normally distributed, VAS = Visual Analogue Scale, mHHS = Modified Harris Hip Score

Note: Where the Shapiro-Wilk test was non significant, no further analysis was performed

Appendix B. 10 - Chi-Square Critical Values Table

Chi-Square (χ^2) Distribution

Degrees of Freedom	Area to the Right of Critical Value							
	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01
1	—	0.001	0.004	0.016	2.706	3.841	5.024	6.635
2	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210
3	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345
4	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277
5	0.554	0.831	1.145	1.610	9.236	11.071	12.833	15.086
6	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812
7	1.239	1.690	2.167	2.833	12.017	14.067	16.013	18.475
8	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090
9	2.088	2.700	3.325	4.168	14.684	16.919	19.023	21.666
10	2.558	3.247	3.940	4.865	15.987	18.307	20.483	23.209
11	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725
12	3.571	4.404	5.226	6.304	18.549	21.026	23.337	26.217
13	4.107	5.009	5.892	7.042	19.812	22.362	24.736	27.688
14	4.660	5.629	6.571	7.790	21.064	23.685	26.119	29.141
15	5.229	6.262	7.261	8.547	22.307	24.996	27.488	30.578
16	5.812	6.908	7.962	9.312	23.542	26.296	28.845	32.000
17	6.408	7.564	8.672	10.085	24.769	27.587	30.191	33.409
18	7.015	8.231	9.390	10.865	25.989	28.869	31.526	34.805
19	7.633	8.907	10.117	11.651	27.204	30.144	32.852	36.191
20	8.260	9.591	10.851	12.443	28.412	31.410	34.170	37.566
21	8.897	10.283	11.591	13.240	29.615	32.671	35.479	38.932
22	9.542	10.982	12.338	14.042	30.813	33.924	36.781	40.289
23	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638
24	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980
25	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314
26	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642
27	12.879	14.573	16.151	18.114	36.741	40.113	43.194	46.963
28	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278
29	14.257	16.047	17.708	19.768	39.087	42.557	45.722	49.588
30	14.954	16.791	18.493	20.599	40.256	43.773	46.979	50.892

Appendix B. 11 - Strength and range of motion scores obtained from the RM-ANOVA test.

Measurement		Right		Left	
		M	*SD*	*M*	*SD*
Abduction (*N*) (*n*=54)	**Preoperative**	85.42	29.60	84.90	31.03
	2nd week	78.14	27.26	77.00	26.27
	8th week	94.44	26.18	92.89	27.00
	24th week	103.73	34.52	99.73	33.16
Adduction (*N*) (*n*=53)	**Preoperative**	79.69	25.27	82.82	25.56
	2nd week	76.63	27.67	76.54	27.67
	8th week	92.32	35.45	92.32	29.91
	24th week	101.20	35.45	97.88	29.91
External Rotation (degrees) (*n*=51)	**Preoperative**	30.63	7.84	31.26	10.31
	2nd week	30.88	8.51	31.22	10.09
	8th week	31.12	6.57	31.35	6.94
	24th week	33.55	9.04	33.04	10.43
Internal Rotation (degrees) (*n*=51)	**Preoperative**	33.61	10.91	33.80	11.63
	2nd week	32.90	10.53	33.08	10.39
	8th week	35.39	12.25	35.47	14.73
	24th week	35.86	12.25	37.06	14.73
FABER D. (cm) (*n*=60)	**Preoperative**	22.69	7.24	22.44	6.55
	2nd week	23.91	7.23	23.47	6.79
	8th week	20.66	5.73	20.93	7.30
	24th week	18.55	5.58	19.52	7.07

N, newton; FABER D, distance between the knee and bed; *M*, mean; *SD*, standard deviation.

Appendix B. 12 - Functional level and pain scores obtained from the RM-ANOVA test.

Measurement		*M*	*SD*
MHHS (*n*=27)	**Preoperative**	65.41	13.83
	8th week	75.15	15.72
	24th week	84.78	13.01
VAS (*n*=49)	**Preoperative**	7.14	2.16
	2nd week	2.61	2.13
	8th week	2.39	2.06
	24th week	1.70	1.93

N, newton; MHHS, modified Harris Hip Score; VAS, visual analogue scale; FABER D, distance between the knee and bed; *M*, mean; *SD*, standard deviation.

Appendix C Chapter 4 Supporting Material

Appendix C. 1 - A pilot study using the platform version P2.3

The study was performed on six volunteers using muscle length testing and manual muscle testing. Shoulder and knee flexibility tests were conducted to assess if they affect the participant's comfort and range of motion. These were followed by evaluating the shoulder and upper limbs muscle strength. Lastly, subjective questions were asked about the preferred position and the reason (Appendix C. 2). Placing the board on a typical physiotherapy bed showed that it could compromise testing safety. The position of the board will not be safe nor comfortable especially for those with internal hip pathologies.

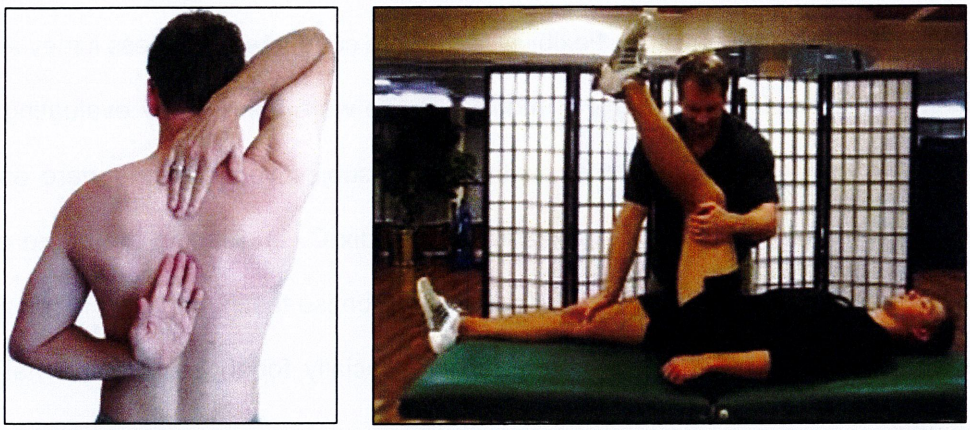
Appendix C. 2 - Tests and Procedures Performed in the Pilot Testing of Concept 2.3.

Potential Source of Error	Test / Procedure
Position of the board – on the bed or the floor	Lab trial by positioning it on both locations and check
Position of the subject – long sitting or supine	Pilot study: Positioning the volunteers in both positions and ask them about their preference and reason
Shoulder and thigh flexibility	Pilot study: Perform Apley's Scratch Test for the shoulder and Active Knee Extension Test for the hamstrings
Shoulder, arm and forearm strength (for support during long sitting)	Pilot study: Perform manual muscle testing for the related shoulder, arm and forearm muscles
Position of the arms during testing (wrapped around chest or supports trunk in long sitting)	Pilot study: Positioning the volunteers in both positions and ask them about their preference and reason

The testing results showed that the participants who scored poor to fair on the Apley's Scratch Test* (TopendSports, 2015), indicating reduced shoulder flexibility, did not complain of discomfort in the shoulder region when they were asked to support the upper body in a long sitting position (Appendix C. 3). The only participant who scored fair in

* Performed by asking the participant to "scratch" the upper back by externally rotating a shoulder and internally rotating the other. The test is performed again with reversing the action. The distance between the middle fingers is compared in the two positions. Scoring criteria: Good: fingers are touching, Fair: < 5 cm between fingers, Poor: > 5 cm distance between fingers.

the Active Knee Extension Test* (Rebodybuilding, 2014) complained of discomfort and inflexibility in the hamstring muscles in the long sitting position (Appendix C. 4). Overall, all volunteers preferred the supine position to the long sitting position because of the enhanced force production and recruitment of the hip abductor and adductor muscles.



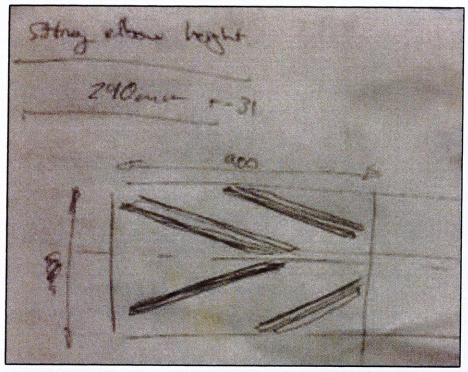
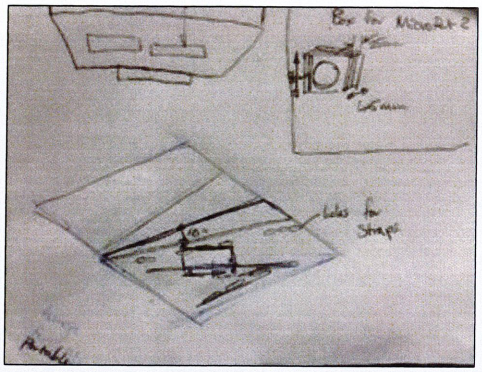
Appendix C. 3 - Apley's scratch test (TopendSports, 2015) (left); Active knee extension test (Rebodybuilding, 2014) (right).

Appendix C. 4 - Results of concept 2.3 pilot study.

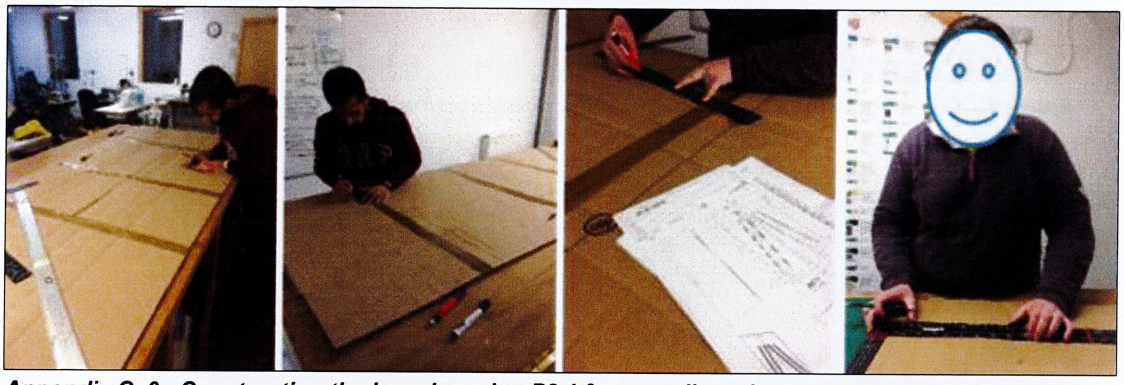
No.	Sex	Age	AKET	Apley's	SAB	SHAB	SER	SIR	EEX	WFL	Preferred position	Comments/Feedback
1	M	29	Good	Fair	5	5	5	5	5	5	Supine	Feels increased force production in supine
2	M	29	Good	Fair	5	5	5	5	5	5	Supine	Feels increased force production in supine
3	M	23	Fair	Good	5	5	5	5	5	5	Supine	Bothered with inflexibility and prefers to have the opposite limb supported
4	F	31	Good	Good	5	5	4	4	5	5	Supine	Got cramp in quads with long sitting → used it as secondary muscles to compensate
5	F	31	Good	Poor	5	5	4	5	5	5	Supine	Easier in supine compared to long sitting

No., number; M, Male, F, Female, AKET, Active Knee Extension Test; SAB, Shoulder Abduction; SHAB, Shoulder Horizontal Abduction; SER, Shoulder External Rotation; SIR, Shoulder Internal Rotation; EEX, Elbow Extensors; WFL, Wrist Flexors.

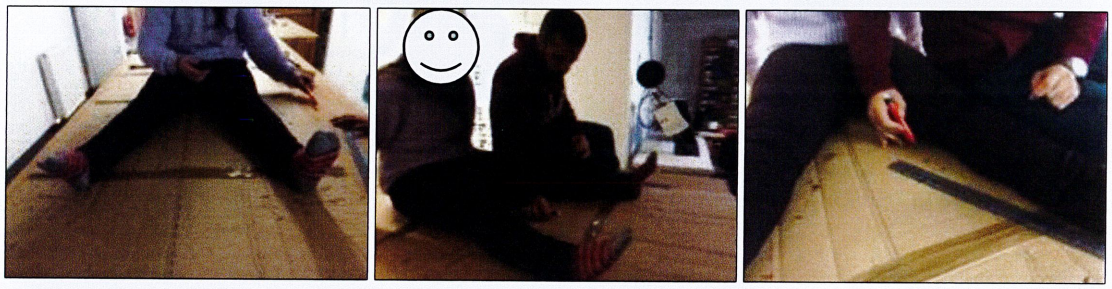
* Performed by placing the hip and knee in 90° of flexion. The subject is asked to extend the knee while maintaining hip position. Scoring criteria: Poor: Fixed at approximately 90°, Fair: 0-20° increase in angle, Good: > 20° increase in angle.

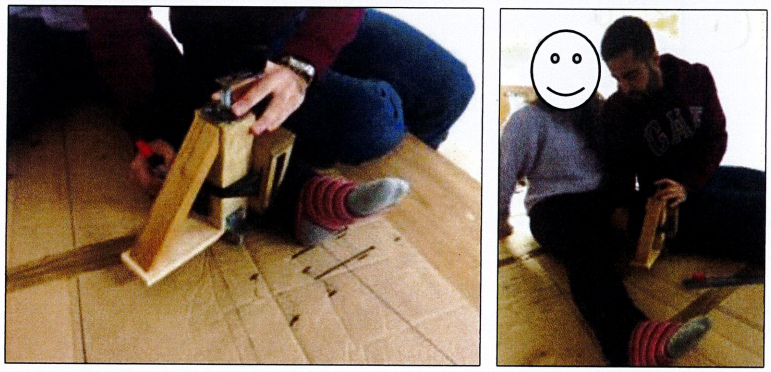
Appendix C. 5 - Sketching the board version P2.0.



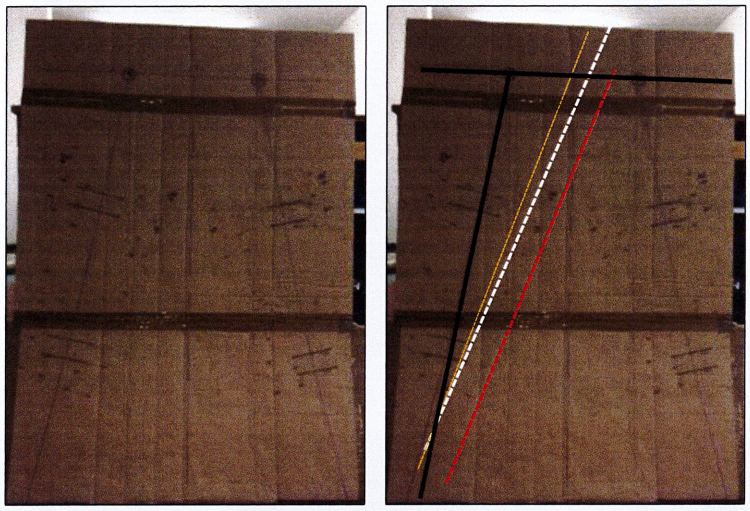
Appendix C. 6 - Constructing the board version P2.1 from cardboard.



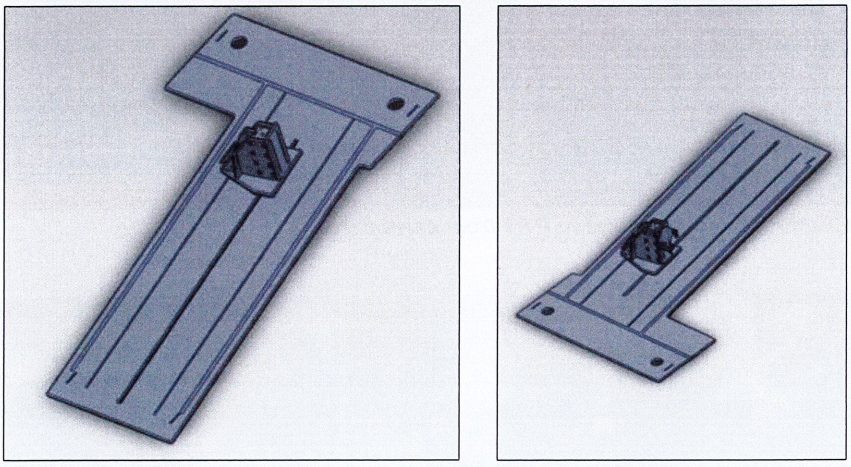
Appendix C. 7 - Quick testing of the cardboard model of board version P2.1.



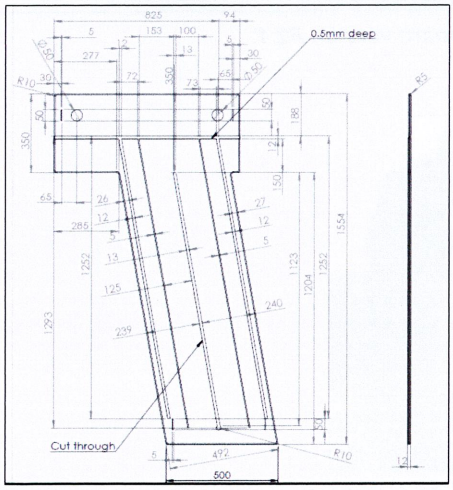
Appendix C. 8 - Testing the dynamometer support version S2.0.



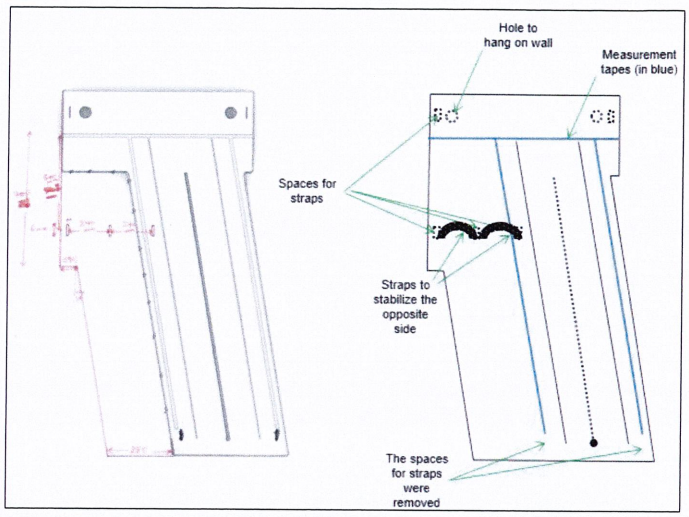
Appendix C. 9 - The cardboard prototype of the board version P2.1.



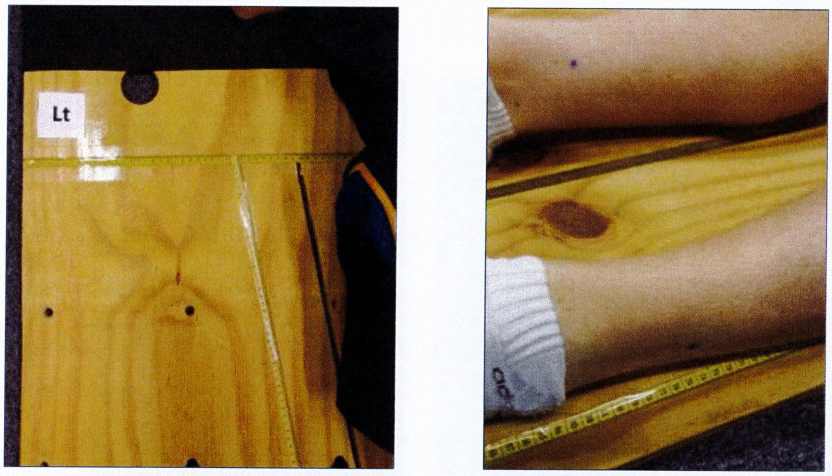
Appendix C. 10 - SolidWorks drawings of the board version P2.3 and dynamometer support version S2.1.



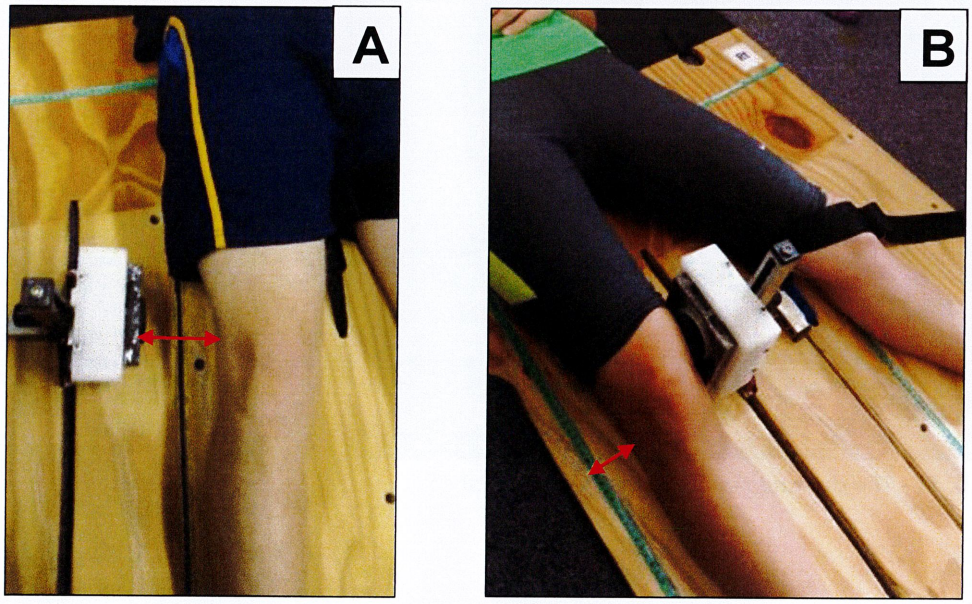
Appendix C. 11 - The dimensions of the board version P2.3.



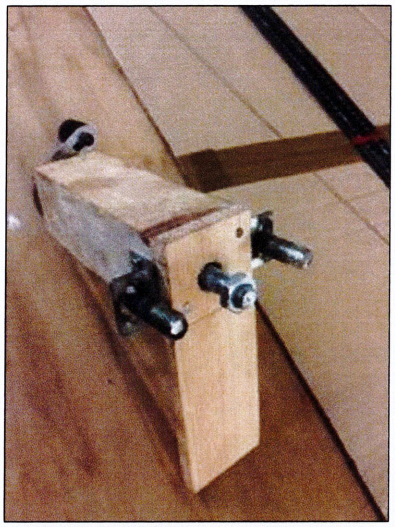
Appendix C. 12 - Sketches of the board version P2.4.



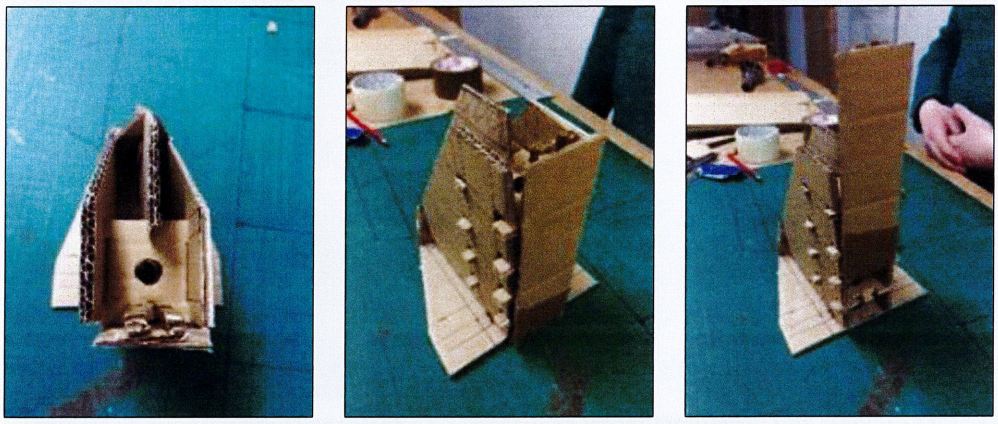
Appendix C. 13 - Labels and measurement tapes locations on the board version P2.4.



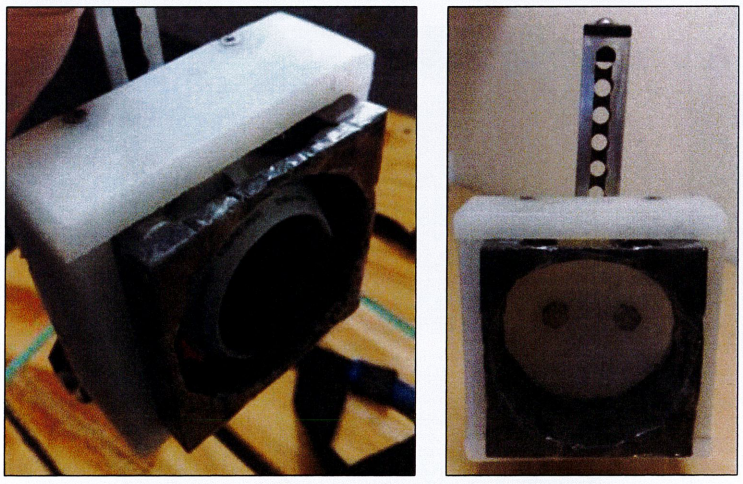
Appendix C. 14 - (A) A gap formed between the dynamometer and skin, (B) Loss of the 10° of hip abduction to properly align the dynamometer with the skin while testing using the



Appendix C. 15 – The dynamometer support version S2.0



Appendix C. 16 - Cardboard prototype of version S2.1 of the dynamometer support.



Appendix C. 17 - The dynamometer support version S3.0.

Appendix D Chapter 5 Supporting Material



Appendix D. 1 - Screenshot of the Leeds Beckett University Research Ethics Applications webpage.

Page 1

School of Sport – Physiology Group Screening Questionnaire

LEEDS BECKETT UNIVERSITY

Name:_____

Risk Factors		Risk Factor	No Risk Factor
Q 1. Age = _____ *years*	Males	≥ 45	< 45
	Females	≥ 55	< 55
Q 2. Have any parents, brothers or sisters had a heart attack, bypass surgery, angioplasty, or sudden death, prior to 55 years (male relatives) or 65 years (female relatives)*		Yes	No
Q 3. Are you currently a smoker – have you quit within the past 6 months – are you exposed to environmental tobacco smoke?		Yes	No
Q 4. In the past 3 months have you performed at least 30 minutes of moderate intensity physical activity or equivalent on at least 3 days of the week (total of 90min mod. act.)?		No**	Yes
Q 5a. Body mass index = _____ $kg.m^{-2}$ (weight divided by height squared)		≥ 30	< 30
Q 5b. Waist girth = _____ cm	Males	> 102	≤ 102
	Females	> 88	≤ 88
Q 6a. Do you take blood pressure medication		Yes	No
Q 6b. Resting blood pressure: SBP = _____ *mmHg, DBP = _____ mmHg*		≥ 140/90***	< 140/90
TOTAL NUMBER OF RISK FACTORS * If Yes to early sudden death in family history advise pre-participation screening for SCD ** If currently sedentary consider whether max test is essential ***If BP ≥140/90mmHg treat as High Risk and advise pre-participation screening for SCD			

Signs or Symptoms	S/S	No S/S
Q 7. Do you ever have pain or discomfort in your chest or surrounding areas (neck, jaw, arms or other areas)?	Yes	No
Q 8. Are you ever short of breath at rest or with mild exertion?	Yes	No
Q 9. Have you ever experienced dizziness or loss of consciousness during or shortly after exercise?	Yes	No
Q10. Have you ever been short of breath at rest in the recumbent position or had an attack of breathlessness in the middle of the night which was relieved by sitting up?	Yes	No
Q 11. Do your ankles ever become swollen (other than as a result of an injury)?	Yes	No
Q 12. Do you ever have palpitations (=the unpleasant awareness of the heart beating in your chest) or an unusual period of rapid heart rate?	Yes	No
Q 13. Do you ever suffer from cramp-like pains in your legs, brought on by exertion and relieved after 1-2 minutes of rest?	Yes	No
Q 14. Has a doctor ever said you have a heart murmur?	Yes	No
Q 15. Do you feel unusually fatigued or find it difficult to breathe with usual activities?	Yes	No
SIGNS/SYMPTOMS OF DISEASE	YES / NO	

Appendix D. 2 - The physiology group screening questionnaire that was used in phases one and two of the platforms study.

Page

History of Disease	History of Disease	No History of Disease
Q 16. Heart disease	Yes	No
Q 17. Peripheral vascular disease	Yes	No
Q 18. Cerebrovascular disease (e.g. stroke)	Yes	No
Q 19. Chronic obstructive pulmonary disease (emphysema/chronic bronchitis)	Yes	No
Q 20. Asthma	Yes	No
Q 21. Interstitial lung disease	Yes	No
Q 22. Cystic fibrosis	Yes	No
Q 23. Diabetes mellitus	Yes	No
Q 24. Thyroid disorder	Yes	No
Q 25. Renal disease	Yes	No
Q 26. Liver disease	Yes	No
HISTORY OF DISEASE	YES / NO	

Other conditions	Condition	Condition
Q 27. Do you have any bone or joint problems such as arthritis or a past injury that might get worse with exercise? (Exercise testing may need delaying or modifying)	Yes	No
Q 28. Do you have any other problem that might make it difficult for you to do strenuous exercise? Details:	Yes	No
Q29. Are you or have you recently been pregnant?	Yes	No
Q30. Are you on any prescription medications? List:	Yes	No

*If HDL>1.55 – subtract 1 from the total number of risk factors

Risk Analysis

Total no. of RF (inc blood analysis, if necessary) = _____

SS/HoD = YES / NO – Other conditions to consider: YES / NO

Low risk: No more than one risk factor – safe to do submax/max exercise testing or enter a vigorous programme

Moderate risk: More than one risk factors – safe to do submax exercise testing or enter a moderate programme (max test must be conducted by a person trained in clinical exercise testing)

High risk: One or more signs/symptoms of disease or HoD - cannot do any testing or exercise without physician clearance

Final risk category =

- I confirm that the above information which I have provided to Leeds Beckett University is true and accurate to the best of my knowledge and belief and I understand that I must notify promptly of any changes to the information.
- I understand that the information I have provided above may be used as part of an anonymised dataset by staff or students of the School of Sport for completion of coursework or for research or audit purposes (with the appropriate ethical approval in place).

Signature: **Date:**

Assessors Signature: **Date:**

(continue)

LEEDS BECKETT UNIVERSITY
School of Sport

Dynamometer Study Specific Screening Questionnaire

Personal Information

Name (First, last): _____

Gender: ◯ Male ◯ Female

Contact Telephone / Mobile number: _____

Emergency contact telephone number: _____

Email address: _____

Date of Birth: _____ / _____ / _____

Would you like to be contacted in the future via email to participate in another study?

◯ Yes ◯ No

Specific health and physical activity screening (please tick)

	Signs and Symptoms	Yes	No
1	Do you ever feel faint or have dizzy spells while performing resistive training?		
2	Are you recovering from a recent illness or operation? If yes, please mention		
3	Are you currently taking any prescribed medication? If yes, please mention		
4	Do you suffer from any muscle or joint disorder which may be aggravated by physical activity?		
5	Did you suffer from any lower back, pelvic or lower extremities pain in the past 3 months? If yes, please explain:		
6	Did you experience any recent injury in the pelvic, hip, groin or knee regions in the past 3 months? If yes, please explain:		
7	Do you complain from any other condition or problem that may affect your physical performance? If yes, please explain:		

Kindly proceed to the second part

8	How many times per week do you exercise or perform physical activity that involves your lower limbs? ◯ Zero ◯ 1-2 times ◯ 3 times or more
9	What is the total amount of time per week do you spend exercising or being physically active? ◯ Zero ◯ 1-2 hours ◯ 3-4 hours ◯ 4.5 hours or more

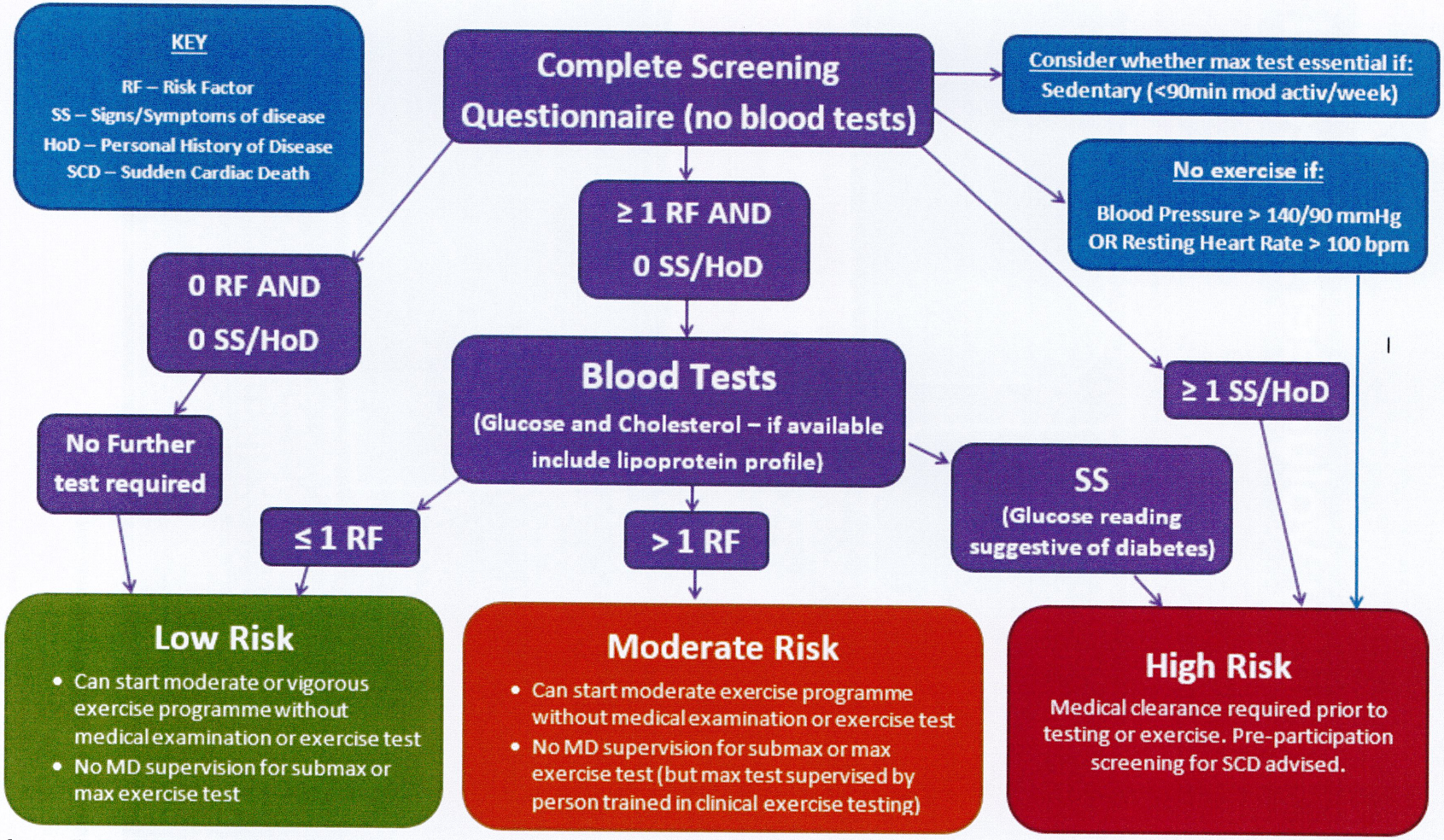
- I confirm that the above information which I have provided to Leeds Beckett University is true and accurate to the best of my knowledge and belief and I understand that I must notify promptly of any changes to the information.
- I understand that the information I have provided above may be used as part of an anonymised dataset by staff or students of the School of Sport for completion of coursework or for research or audit purposes (with the appropriate ethical approval in place).

Signature: **Date:**

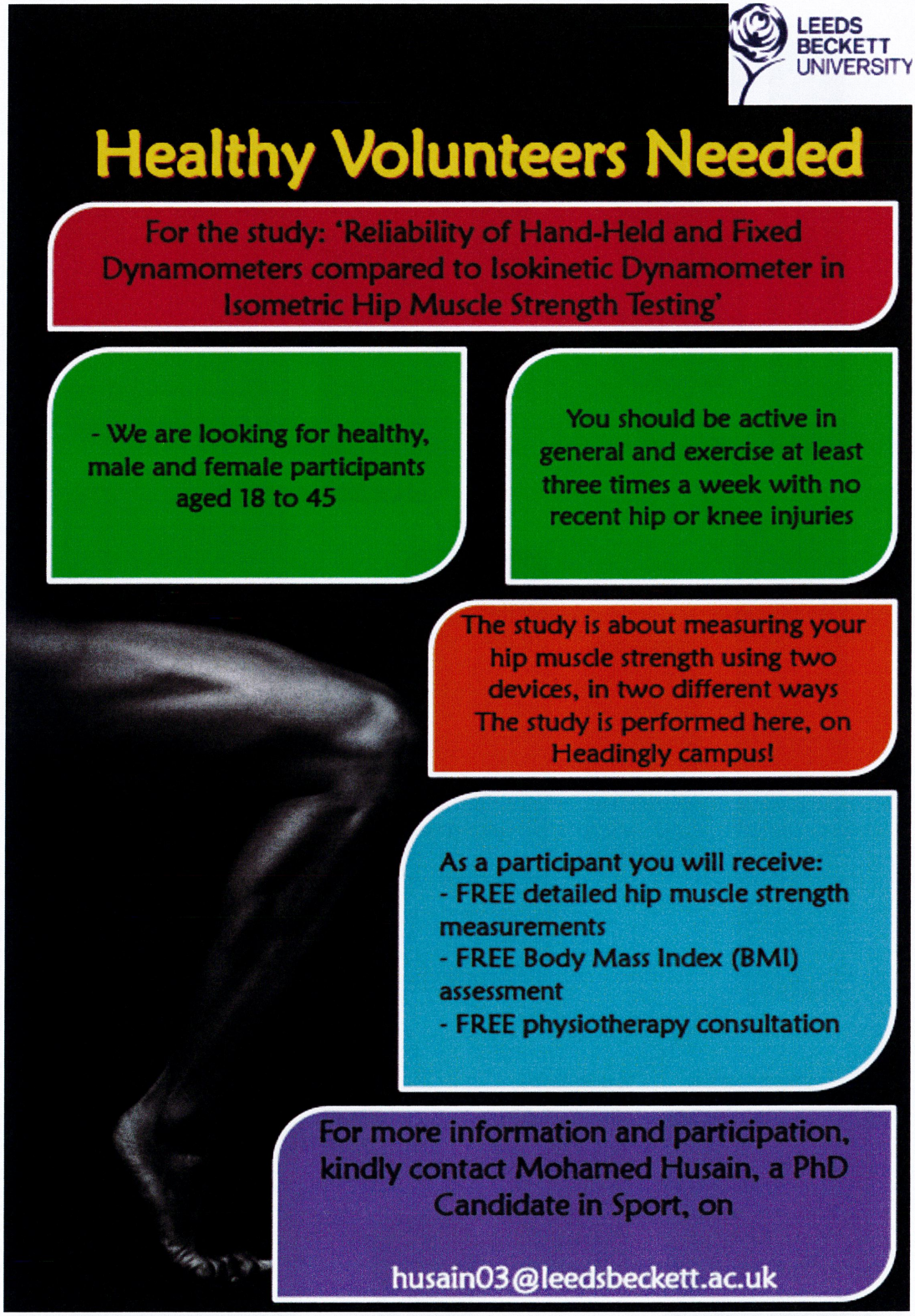
Assessors Signature: **Date:**

Proceed to Physiology Screening Questionnaire YES NO

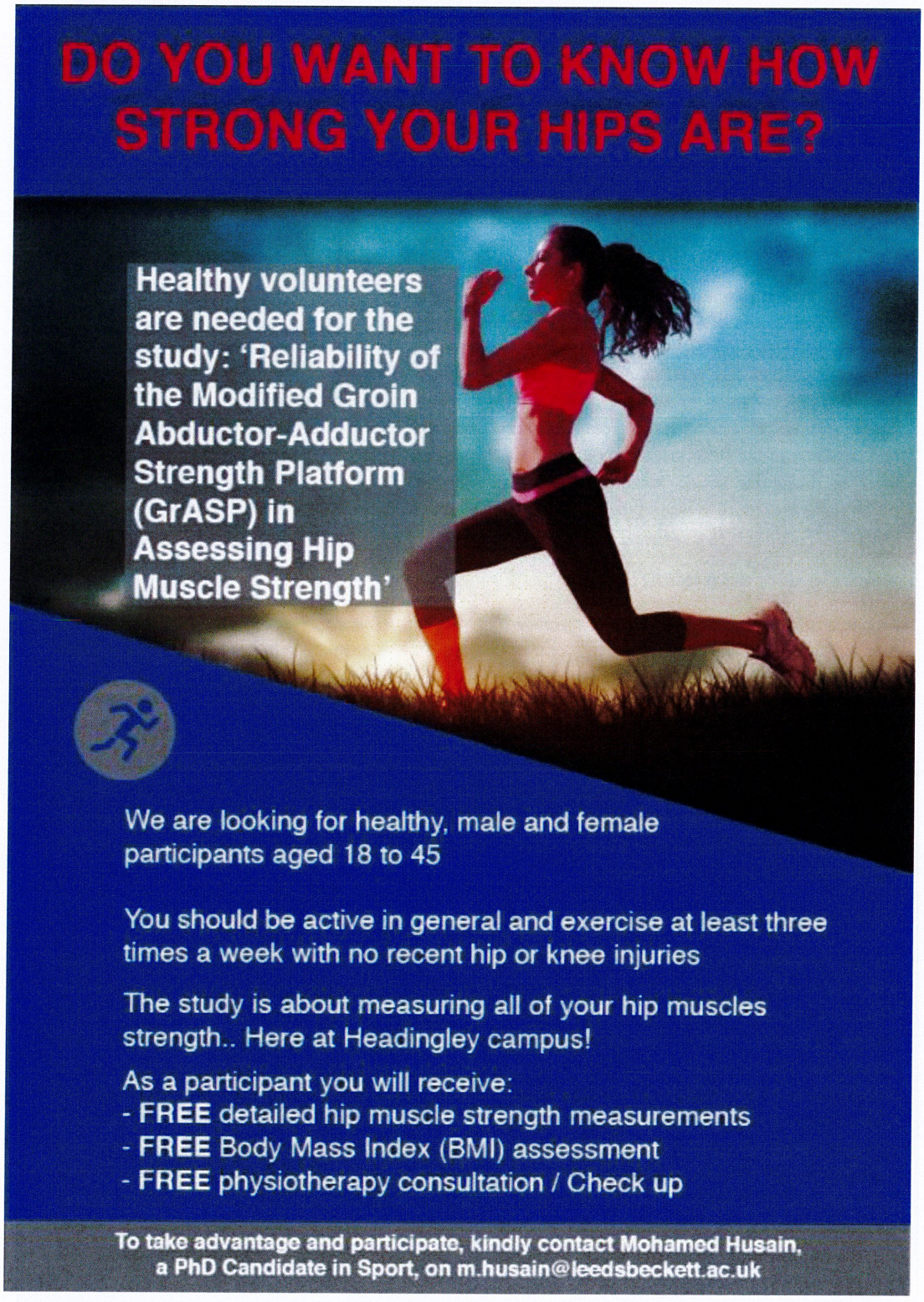
Appendix D. 3 - The dynamometer study specific screening questionnaire that was used in phased one and two of the platforms study.



Appendix D. 4 – The screening process.



Appendix D. 5 – Phase one recruitment poster.



Appendix D. 6 – Phase two recruitment poster.

LEEDS BECKETT UNIVERSITY
School of Sport

Information Sheet

Study: Reliability of Hand-Held and Fixed Dynamometers compared to Isokinetic Dynamometer in Isometric Hip Muscle Strength Testing: A Pilot Study

Overview:



This study aims at comparing two different methods of using a small device (hand-held dynamometer – microFET2) to quantify hip/thigh muscle strength. The device measures the amount of strength your muscles can produce without causing pain. You will be asked to perform certain hip movements against the device while lying down on your back. During the first session, you will be assessed using a larger state of the art machine called the Biodex. After that, two examiners will perform the procedure in three separate sessions by placing the microFET2 device close to your knees and ankles.

Why are we doing this research?

Proper strength assessment is vital to determine if players can go back to play or not. Data obtained from this study will help clinicians decide which of the two methods and positions is better to assess the strength of hip muscles. It may also lay down the foundations for bigger future research projects.

Am I eligible to participate?

You are if you meet the following criteria: You don't complain from systemic problems/diseases related to the heart or respiratory systems, blood hypertension or diabetes. Moreover, you shouldn't be currently complaining from pelvic, groin or lower extremity pain or had an injury in the past 3 months. Also, you should be performing any kind of physical activity involving your lower limbs for more than 3 times/week with a total of 4.5 hours or more/week.

What will happen if I agreed to participate?

You will be asked to attend three sessions of hip abductor and adductor strength assessment. Your height and weight will be measured to calculate Body Mass Index (BMI). Before each session, necessary skin landmarks on your thigh and leg will be made for the device/pad placement and lever measurements. The starting test side, muscle, dynamometer stabilisation method and placement will all be randomly determined. The examiner will ask you to perform a 5 minutes warm up on a stationary bicycle or run outdoors prior to any strength assessment. In the first session, you will be positioned appropriately on The Biodex System 4 machine which is available at the Carnegie Research Institute labs. You will be allowed 2-3 attempts before the actual procedure, which consists of 5 contractions for each muscle group on each side using two modes of muscle activity. You will be given one minute of rest in between repetitions. It may take approximately 20 minutes to assess each side. In the following two sessions, muscle strength will be measured using the microFET2 (small measurement device). You will be positioned appropriately on a board and the pad of the dynamometer will be positioned perpendicular to the skin in the thigh and leg regions.

Appendix D. 7 – Phase one study Information Sheet.

You will be allowed 2-3 trials followed by 1 minute rest period. For each strength measurement, you will be asked to perform 3 muscle contractions for 5 seconds, followed by 30 seconds rest period. This will be followed by testing the other side using the same method and sequence of placement. The procedure will be repeated again but with changing the method the microFET2 is stabilised (once by hand and once to a board). Once done, you will be allowed to perform stretching exercises and will be scheduled for a third session within seven days. The same procedure as the second session will be performed. In general, each session lasts about an hour.

What are the possible benefits from participating in the study?

By participating in this study, you will receive an analysed version of your results compared to normal values. In addition, you will be given a FREE physiotherapy consultation session and/or advice to any sports related physical problem within 1 month from a qualified and experienced physiotherapist.

What are the possible risks associated with participating in the study?

There are no adverse effects resulting from participating in the study. However, you might feel some soreness in the buttock and/or thigh regions due to the increased number of repetitions. The type of soreness you might experience is similar to mild post-exercise muscular fatigue or ache. To minimise the effect, you will be asked to warm-up on a stationary bicycle prior to testing and perform stretching exercises afterword. Light to moderate fatigue is also expected. However, part of the testing will be performed using the Biodex machine, which if not handled with care, might result in minor injuries. Cardiovascular and respiratory emergencies could occur if prior medical history was not revealed in details to the examiner. The machine will be operated by qualified researchers following the procedure mentioned in the manual to avoid any incidents.

How are we going to keep your information anonymous and confidential?

Any information obtained in this study, whether from personal information or data, will remain confidential and the researcher(s) will maintain the anonymity of participants by not mentioning names or contact details. In addition, your personal information will not be sold or given to third parties. We may take photographic and/or videotape recordings of the procedure when needed. However, the researcher will maintain your confidentiality by not showing your face. The images may be published in an appropriate journal or textbook or presented in a lecture / conference.

What if I don't want to continue participating in the study?

You have the right to withdraw from the study at any stage and withdraw your data up to completion of data analysis. You do not have to give any reasons or clarifications, just inform the researcher of your decision.

What happens if something was not right?

The devices used in this study are calibrated and checked to meet safety requirements. However, in the unlikely event of an adverse effect, you will need to contact Mohamed Husain (below) as soon as possible and the issue will be addressed as appropriate. If serious harm occurred while participating in the study, Leeds Beckett University has clinical trial Indemnity Insurance. The Indemnity Insurance will only cover only what the University is legally liable for as a direct result of participating in the study.

(continue)

LEEDS BECKETT UNIVERSITY
School of Sport

Contact details:

If after reading this Information Sheet you feel like you have unanswered questions, please do not hesitate to contact any of the researchers below.

Principal researcher:

Mohamed Husain
PhD Candidate in Sport
Carnegie Faculty, Leeds Beckett University
Room G03 Macaulay building,
Headingley Campus
Leeds LS6 3QT
Email: husain03@leedsbeckett.ac.uk / mob: 07475250022

Supervisory team:

1. Prof. Carlton Cooke
Associate Dean Research Carnegie Faculty
Carnegie Professor of Sport and Exercise Science
Carnegie Faculty, Leeds Beckett University
Room 101 Fairfax building,
Headingley Campus
Leeds LS6 3QT
Email: C.Cooke@leedsbeckett.ac.uk

2. Prof. Ernest Schilders
Hip and Groin Orthopaedic Surgeon
Yorkshire Clinic
Bradford Road
Bingley BD16 1TW
Email: e.schilders@btopenworld.com

If you wish to talk to someone who is not directly involved in the study but researches the same area, or would like to make any form of complaint, kindly contact:

Dr John O'Hara
Reader, Local Research Ethics Committee Coordinator
Carnegie Faculty of Sport and Education
Leeds Metropolitan University
Room 107 CRI hall
Headingley Campus
Leeds LS6 3QT
Email: J.OHara@leedsbeckett.ac.uk

Thank you for your interest in participating.
M Husain

(continue)

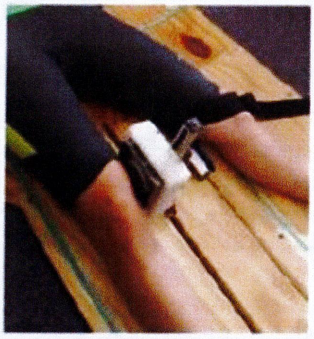
LEEDS BECKETT UNIVERSITY
School of Sport

Information Sheet

Study: Reliability of the Modified Groin Abductor-Adductor Strength Platform (GrASP) in Assessing Hip Muscle Strength

Overview:

This study aims at looking the reliability of a small device (microFET2) to quantify hip and thigh muscles strength. The device measures amount of strength your muscles can produce without causing pain. You will be asked to perform certain hip and knee movements against the device while lying down or sitting on an assessment platform. There will be two sessions within 7 days using the same technique followed by one session using another machine called the Biodex. The Biodex is considered the gold standard of muscle strength assessment.



Why are we doing this research?

It has been shown that using a board or a fixed method to stabilise the dynamometer is better than just holding it in hand and this study will help to support this conclusion. In addition, data obtained from this study will be used to establish reference strength values of various hip and knee muscles using the microFET2 and the Biodex.

Am I eligible to participate?

You are if you meet the following criteria: You don't complain from systemic problems/diseases related to the heart or respiratory systems, blood hypertension or diabetes. Moreover, you shouldn't be currently complaining from pelvic, groin or lower extremity pain or had an injury in the past 3 months. Also, you should be performing any kind of physical activity involving your lower limbs for more than 3 times/week with a total of 3 hours or more/week.

What will happen if I agreed to participate?

You will be asked to attend three sessions of hip and thigh muscle assessment. Before the first session, you will be given the consent form to read, sign and ask questions. After that you will be asked to fill in two questionnaires to determine your eligibility to participate. Stature and body mass will be measured to calculate Body Mass Index (BMI). Before each session, necessary skin marks on your leg will be made for device placement and lever arm measurements. The starting muscle group will be randomly determined. The examiner will ask you to warm up on a stationary bicycle or run outdoors prior to any strength assessment followed by stretching. You will be positioned appropriately on a board and the pad of the dynamometer will be positioned just above your ankle joint. For each strength measurement, you will be asked to perform 3 muscle contractions for 5 seconds, followed by 30-60 seconds rest period. The procedure will be repeated again but for the opposite muscle group. Once done, you will be allowed to perform stretching exercises and will be scheduled for the second session within 7 days. Within another 7 days, you will be scheduled for the third (final) session using the Biodex. In the Biodex testing session, you will be strapped against the machine in certain positions and will be given verbal and on-screen instructions. This test

Appendix D. 8 – Phase two study Information Sheet.

involves pushing against the machine for 5 seconds followed by 60 seconds rest period. In general, each session would last about an hour and a half.

What are the possible benefits from participating in the study?

By participating in this study, you will receive an analysed version of your results from both devices. In addition, you will be given a FREE physiotherapy consultation session and/or advice to any sports related physical problem within 1 month from a qualified and experienced physiotherapist. By participating in this study, you will be given the opportunity to apply your biomechanics and anatomy knowledge and ask any questions to enhance your experience.

What are the possible risks associated with participating in the study?

There are no adverse effects resulting from participating in the study. However, you might feel some soreness in the buttock and/or thigh regions due to repeated contractions. The type of soreness you might experience is similar to mild post-exercise muscular fatigue or ache. To minimise the effect, you will be asked to warm-up and perform stretching exercises prior to testing. Light to moderate fatigue is also expected. Cardiovascular and respiratory emergencies could occur if prior medical history was not revealed in details to the examiner. The testing device will be operated by a qualified researcher following the procedure mentioned in the manual to avoid any incidents and he will be with you during all testing phases.

How are we going to keep your information anonymous and confidential?

Any information obtained in this study, whether from personal information or data, will remain confidential and the researcher will maintain the anonymity of participants by not mentioning names or contact details. In addition, your personal information will not be sold or given to third parties. Photographic and/or videotape recordings of the procedure might be taken when needed. However, the researcher will maintain your confidentiality by not showing your face. The images may be published in an appropriate journal or textbook or presented in a lecture / conference.

What if I don't want to continue participating in the study?

You have the right to withdraw from the study at any stage and withdraw your data up to completion of data analysis. You do not have to give any reasons or clarifications, just inform the researcher of your decision.

What happens if something was not right?

The device used in this study is calibrated and checked to meet safety requirements. However, in the unlikely event of an adverse effect, you will need to contact Mohamed Husain (below) as soon as possible and the issue will be addressed as appropriate. If serious harm occurred while participating in the study, Leeds Beckett University has clinical trial Indemnity Insurance. The Indemnity Insurance will only cover only what the University is legally liable for as a direct result of participating in the study.

I would like to thank you for your interest in participating in this study.
Mohamed.

(Continue reading for contact details)

(continue)

LEEDS BECKETT UNIVERSITY
School of Sport

Contact details:
If after reading this Information Sheet you feel like you have unanswered questions, please do not hesitate to contact any of the researchers below

Principal researcher
Mohamed Husain, PhD Candidate in Sport
Carnegie Faculty, Leeds Beckett University
105 Churchwood Hall
Headingley Campus
Leeds LS6 3QT
E-mail: m.husain@leedsbeckett.ac.uk / mob: 07475250022

Supervisory team.
1. Dr. Claire Griffiths, Senior Lecturer in Physical Activity, Exercise & Health
Tel: 0113 812 6566
E-mail: C.Griffiths@leedsbeckett.ac.uk

2. Prof. Carlton Cooke, Head of the School of Sport
Leeds Trinity University
E-mail: C.Cooke@leedstrinity.ac.uk

3. Prof. Ernest Schilders, Hip and Groin Orthopaedic Surgeon
Yorkshire Clinic
Bradford Road
Bingley BD16 1TW
E-mail: e.schilders@btopenworld.com

If you wish to talk to someone who is not directly involved in the study but researches the same area, or would like to make any form of complaint, kindly contact:

Prof. John O'Hara, Professor in Sport and Exercise Physiology
Carnegie Faculty of Sport and Education
Leeds Beckett University
Room 107 CRI hall
Headingley Campus
Leeds LS6 3QT
Email: J.OHara@leedsbeckett.ac.uk

(continue)

LEEDS BECKETT UNIVERSITY
School of Sport

Informed Consent Form

Name_____

I hereby give my consent to participate in the hip muscle strength testing using the dynamometer which is part of the study:
Titled: Reliability of Hand-Held and Fixed Dynamometers compared to Isokinetic Dynamometer in Isometric Hip Muscle Strength Testing: A Pilot Study
Researcher's name: Mohamed Husain **Course Title:** Doctor of Philosophy in Sport
Student Number: 77149318 **Supervisors' names:** Prof. Carlton Cooke / Prof. Ernest Schilders

I, the undersigned, confirm that (please circle as appropriate):

• I have read and understood the information about the project, potential adverse effects as provided in the Information Sheet and I am satisfied that I understand the procedures involved.	Yes	No
• I agree to participate in all stages of the study as outlined in the Information Sheet.	Yes	No
• I hereby agree to have photographic and/or videotape recordings made of me when needed.	Yes	No
• I have been given the opportunity to ask questions about the project and my participation.	Yes	No
• I understand that, during the course of the test/activity, I have the right to ask further questions about it.	Yes	No
• The information which I have supplied to Leeds Beckett University prior to taking part in the test/activity is true and accurate to the best of my knowledge and belief and I understand that I must notify promptly of any changes to the information.	Yes	No
• The procedures regarding confidentiality have been clearly explained to me.	Yes	No
• I understand that my participation in the test/activity is voluntary and I am therefore at liberty to withdraw my involvement at any stage.	Yes	No
• I understand that, if there is any concern about the appropriateness of my continuing in the test/activity, I may be asked to withdraw my involvement at any stage.	Yes	No
• I understand that once the test/activity has been completed, the information gained as a result of it may be used as part of an anonymised dataset by the lead researcher for completion of a doctoral thesis, publication and future research projects. Further to this the dataset may be used by undergraduate and postgraduate students for the completion of Major Independent Studies (dissertations) (with the appropriate ethical approval in place).	Yes	No
• I understand that other researchers will have access to this data only if they agree to preserve the confidentiality of the data and if they agree to the terms I have specified in this form.	Yes	No
• I, along with the Researcher, agree to sign and date this informed consent form.	Yes	No

SIGNATURE OF THE PARTICIPANT DATE

_____ _____

SIGNATURE OF THE RESEARCHER DATE

_____ _____

M Husain

Appendix D. 9 – Phase one study informed consent.

LEEDS BECKETT UNIVERSITY

School of Sport

Informed Consent Form

Name_____

I hereby give my consent to participate in the hip muscle strength testing using the dynamometer which is part of the study

Titled: Reliability of the Modified Groin Abductor-Adductor Strength Platform (GrASP) in Assessing Hip Muscle Strength

Researcher's name: Mohamed Husain. **Student Number:** 77149318. **Course Title:** Doctor of Philosophy in Sport

Supervisors' names: Dr Claire Griffiths / Prof Carlton Cooke / Prof Ernest Schilders

I, the undersigned, confirm that (please circle as appropriate).

• I have read and understood the information about the project, potential adverse effects as provided in the Information Sheet and I am satisfied that I understand the procedures involved	Yes	No
• I agree to participate in all stages of the study as outlined in the Information Sheet.	Yes	No
• I hereby agree to have photographic and/or videotape recordings made of me when needed	Yes	No
• I have been given the opportunity to ask questions about the project and my participation	Yes	No
• I understand that, during the course of the test/activity, I have the right to ask further questions about it.	Yes	No
• The information which I have supplied to Leeds Beckett University prior to taking part in the test/activity is true and accurate to the best of my knowledge and belief and I understand that I must notify promptly of any changes to the information	Yes	No
• The procedures regarding confidentiality have been clearly explained to me.	Yes	No
• I understand that my participation in the test/activity is voluntary and I am therefore at liberty to withdraw my involvement at any stage.	Yes	No
• I understand that, if there is any concern about the appropriateness of my continuing in the test/activity, I may be asked to withdraw my involvement at any stage.	Yes	No
• I understand that once the test/activity has been completed the information gained as a result of it may be used as part of an anonymised dataset by the lead researcher for completion of a doctoral thesis, publication and future research projects. Further to this the dataset may be used by undergraduate and postgraduate students for the completion of Major Independent Studies (dissertations) (with the appropriate ethical approval in place)	Yes	No
• I understand that other researchers will have access to this data only if they agree to preserve the confidentiality of the data and if they agree to the terms I have specified in this form.	Yes	No
• I along with the Researcher agree to sign and date this informed consent form	Yes	No

SIGNATURE OF THE PARTICIPANT **DATE**

_____ _____

SIGNATURE OF THE RESEARCHER **DATE**

_____ _____

M Husain

1

Appendix D. 10 – Phase two study informed consent.

LEEDS BECKETT UNIVERSITY
School of Sport

Dynamometer Study Risk Assessment Form

Procedure	Hand-held dynamometer and Biodex
Assessment number	1
Date Assessed	05 May 2015
Assessed by	Mohamed Husain
Signed by	**Carlton Cooke DOS**

Severity of hazard consequence (H):

3 = Major (death or severe injury)

2 = Serious (injuries requiring more than three days off work)

1 = Slight (minor injuries requiring no or brief absence from work)

Likelihood of occurrence (L):

3 = High (event is likely to occur frequently)

2 = Medium (event is likely to occur occasionally)

1 = Low (event is unlikely to occur)

Risk rating will be calculated as follows:

Risk rating = hazard consequence x likelihood

Appendix D. 11 – Phase one study Risk Assessment.

LEEDS BECKETT UNIVERSITY

School of Sport

Nature and Method of Work	Hazards/Risks	Safeguards	Risk Rating
Record series of maximal isometric and concentric Adduction/abduction muscle contractions on a hand-held dynamometer and the Biodex	Muscle soreness and fatigue	- Warm-up period of approximately 5 minutes prior to every assessment session. It will consist of low-intensity aerobic exercise, dynamic stretching and low intensity trials for the intended activities, this will reduce the risk of soft tissue injuries (e.g. muscle strains) resulting from inadequate preparation - Participants will be instructed to avoid extraneous physical activity the day prior to any assessment session	H1 x L3 = R3; Moderate Risk after implementing safeguards: H1 x L2 = R2; Acceptable
	Musculoskeletal injuries (e.g. muscle strain, ligament sprains, hyperextension)	- All testing procedures will be standardised and detailed information will be given beforehand - Familiarization period - Use of mechanical emergency stops - Assessment will be performed within the physiological limits of each subject - The screening will include a medical questionnaire, excluding participants with (a history of) musculoskeletal pathology - A first aider will be present in the CRI/Sports centre building. - The researcher involved has been fully trained in all the procedures being performed. - Strict use of protocol - Participants will not be left alone during the tests and will be instructed not to press any button or manipulate the settings of the devices	H2 x L2 = R4; Moderate Risk after implementing safeguards: H2 x L1 = R2; Acceptable
	Cardiovascular compilications or other medical emergencies	- The screening will include a medical and physiology questionnaires, excluding participants with (a history of) serious medical issues - All testing procedures will be standardised and detailed information will be given beforehand - Participants will be informed to stop at any time they deem necessary - First aid support will be present in the building at all times, as well as immediate access to a mobile phone	H3 x L1 = R3; Moderate Risk after implementing safeguards: H2 x L1 = R2; Acceptable
	Electrical hazard (e.g. exposed wires, poor fitting and attachment)	- Cable are checked prior to testing and replaced when necessary - Routine machine check up by the lab technicians	H2 x L2 = R4; Moderate Risk after implementing safeguards: H2 x L1 = R2; Acceptable

Student's Signature: Mohamed Husain Date: 06 / 05 / 2015

Supervisor's Signature: ...Carlton Cooke..... Date: ...17/6/15.....

(continue)

LEEDS BECKETT UNIVERSITY

School of Sport

Board and Dynamometer Risk Assessment Form

Procedure	Board and Dynamometer testing
Assessment number	1
Date Assessed	25 May 2016
Assessed by	Mohamed Husain
Signed by	 **Dr Claire Griffiths (DOS)**

Severity of hazard consequence (H):

3 = Major (death or severe injury)

2 = Serious (injuries requiring more than three days off work)

1 = Slight (minor injuries requiring no or brief absence from work)

Likelihood of occurrence (L):

3 = High (event is likely to occur frequently)

2 = Medium (event is likely to occur occasionally)

1 = Low (event is unlikely to occur)

Risk rating will be calculated as follows

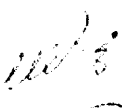
Risk rating = hazard consequence x likelihood

Appendix D. 12 – Phase two study Risk Assessment.

LEEDS BECKETT UNIVERSITY

School of Sport

Nature and Method of Work	Hazards/Risks	Safeguards and Control Measures	Risk Rating
Record series of maximal isometric hip and thigh muscular contractions using the Modified GrASP	Muscle soreness and fatigue / Musculoskeletal injuries (e.g. muscle strain ligament sprains, hyperextension)	- All participants will be medically screened using two questionnaires excluding those with a history of recent (3 months) and related musculoskeletal pain or injuries - Warm-up period of approximately 10 minutes prior to dynamometry testing using a bicycle (low intensity) followed by dynamic mobility exercises (e.g. squats and lunges) to prepare the muscles for activity - Static stretching (30 sec x 3 reps) for each muscle to improve flexibility - All testing procedures will be standardised and detailed information will be given beforehand including how to avoid getting injured. This will be accompanied with a familiarisation period to prepare the participant physically and mentally for the actual test. The assessment will be performed within the physiological limits of the muscles and joints and will not approach end of range - Participants will be instructed to avoid extraneous physical activity the day prior to any assessment session - Participants will be informed of the possibility of delayed onset muscle soreness before signing the consent form - Participants will not be left alone during testing and will be instructed change the set-up of the board or the straps - All testing procedures will be performed with strict application of the protocol and safeguards to avoid unwanted complications. The participants will be aligned properly and stabilised in place (when needed) using straps. All participants will be asked about their comfort prior to testing and will be asked to inform the researcher if they experience any unusual sensation - The tester has been trained on using the dynamometer and setting up the board - The board was manufactured based on medical standards and can support more than 300 lb of body mass	H1 x L2 = R2: Low Risk after implementing safeguards: H1 x L1 = R1: Acceptable
	Cardiovascular complications or other medical emergencies	In addition to the above mentioned safeguards: - Any participant with a history of fainting, heart or respiratory medically diagnosed conditions or signs of underlying pathologies will be immediately excluded and referred to a GP if needed - Participants will be informed to stop at any time they deem necessary. - Participants will be instructed to inform the researcher whenever they feel dizzy or about to faint or if they are short of breath. Although, this is not expected because the procedure does not test maximum cardiorespiratory physiological limits - First aid support will be present in the building at all times, as well as immediate access to a mobile phone.	H3 x L1 = R3: Moderate Risk after implementing safeguards: H2 x L1 = R2: Low
	Other hazards (e.g. exposed wires, poor fitting and attachment)	- The board will be fixed and secured to provide a stable platform for testing. - Any unwanted item around the testing area will be removed. - The board and the dynamometer do not contain any sharp edges that may cause injury.	H1 x L2 = R2: Low Risk after implementing safeguards: H2 x L1 = R2: Acceptable

Student's Signature:  Date: 25/05/2016

Supervisor's Signature:  Date: 26/05/2016

(continue)

Appendix E Chapter 6 Supporting Material

Appendix E. 1 - Normality Tests of the GrAASP data in phase one.

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
HAB Isometric at 0 (DS)	0.158	23	> 0.05	0.958	23	>0.05
HAD Isometric at 0 (DS)	0.122	23	> 0.05	0.949	23	>0.05
HAB Isometric at 15 (DS)	0.123	23	> 0.05	0.973	23	>0.05
HAD Isometric at 15 (DS)	0.143	23	> 0.05	0.942	23	>0.05
HAB Isometric at 30 (DS)	0.124	23	> 0.05	0.968	23	>0.05
HAD Isometric at 30 (DS)	0.103	23	> 0.05	0.972	23	>0.05
HAB Isokinetic at 60.s^{-1} (DS)	0.108	23	> 0.05	0.969	23	>0.05
HAD Isokinetic at 60.s^{-1} (DS)	0.156	23	> 0.05	0.833	23	0.001*
HAB Isokinetic at 120.s^{-1} (DS)	0.078	23	> 0.05	0.985	23	>0.05
HAD Isokinetic at 120.s^{-1} (DS)	0.266	23	< 0.001*	0.797	23	< .001*
HAB Isometric at 0 (NS)	0.132	23	> 0.05	0.941	23	>0.05
HAD Isometric at 0 (NS)	0.123	23	> 0.05	0.911	23	0.043*
HAB Isometric at 15 (NS)	0.087	23	> 0.05	0.984	23	>0.05
HAD Isometric at 15 (NS)	0.168	23	> 0.05	0.925	23	>0.05
HAB Isometric at 30 (NS)	0.102	23	> 0.05	0.986	23	>0.05
HAD Isometric at 30 (NS)	0.122	23	> 0.05	0.950	23	>0.05
HAB Isokinetic at 60.s^{-1} (NS)	0.111	23	> 0.05	0.958	23	>0.05
HAD Isokinetic at 60.s^{-1} (NS)	0.158	23	> 0.05	0.880	23	0.001*
HAB Isokinetic at 120.s^{-1} (NS)	0.121	23	> 0.05	0.969	23	>0.05
HAD Isokinetic at 120.s^{-1} (NS)	0.139	23	> 0.05	0.933	23	>0.05
D1 HAB Manual Proximal (DS)	0.106	21	> 0.05	0.968	21	>0.05
D1 HAB Manual Distal (DS)	0.098	21	> 0.05	0.963	21	>0.05
D1 HAD Manual Proximal (DS)	0.091	21	> 0.05	0.972	21	>0.05
D1 HAD Manual Distal (DS)	0.102	21	> 0.05	0.951	21	>0.05
D1 HAB Fixed Proximal (DS)	0.123	21	> 0.05	0.950	21	>0.05
D1 HAB Fixed Distal (DS)	0.141	21	> 0.05	0.944	21	>0.05
D1 HAD Fixed Proximal (DS)	0.125	21	> 0.05	0.958	21	>0.05
D1 HAD Fixed Distal (DS)	0.134	21	> 0.05	0.951	21	>0.05
D1 HAB Manual Proximal (NS)	0.176	21	> 0.05	0.907	21	0.048*
D1 HAB Manual Distal (NS)	0.150	21	> 0.05	0.957	21	>0.05
D1 HAD Manual Proximal (NS)	0.154	21	> 0.05	0.922	21	>0.05
D1 HAD Manual Distal (NS)	0.133	21	> 0.05	0.937	21	>0.05
D1 HAB Fixed Proximal (NS)	0.096	21	> 0.05	0.948	21	>0.05
D1 HAB Fixed Distal (NS)	0.121	21	> 0.05	0.961	21	>0.05
D1 HAD Fixed Proximal (NS)	0.137	21	> 0.05	0.899	21	0.033*
D1 HAD Fixed Distal (NS)	0.171	21	> 0.05	0.947	21	>0.05
D2 HAB Manual Proximal (DS)	0.138	20	> 0.05	0.881	20	0.019*
D2 HAB Manual Distal (DS)	0.105	20	> 0.05	0.971	20	>0.05
D2 HAD Manual Proximal (DS)	0.124	20	> 0.05	0.975	20	>0.05
D2 HAD Manual Distal (DS)	0.121	20	> 0.05	0.949	20	>0.05
D2 HAB Fixed Proximal (DS)	0.142	20	> 0.05	0.926	20	>0.05
D2 HAB Fixed Distal (DS)	0.173	20	> 0.05	0.932	20	>0.05
D2 HAD Fixed Proximal (DS)	0.159	20	> 0.05	0.889	20	0.026*
D2 HAD Fixed Distal (DS)	0.150	20	> 0.05	0.939	20	>0.05
D2 HAB Manual Proximal (NS)	0.132	20	> 0.05	0.938	20	>0.05
D2 HAB Manual Distal (NS)	0.074	20	> 0.05	0.977	20	>0.05
D2 HAD Manual Proximal (NS)	0.109	20	> 0.05	0.919	20	>0.05
D2 HAD Manual Distal (NS)	0.153	20	> 0.05	0.943	20	>0.05
D2 HAB Fixed Proximal (NS)	0.133	20	> 0.05	0.932	20	>0.05
D2 HAB Fixed Distal (NS)	0.146	20	> 0.05	0.944	20	>0.05
D2 HAD Fixed Proximal (NS)	0.128	20	> 0.05	0.934	20	>0.05
D2 HAD Fixed Distal (NS)	0.140	20	> 0.05	0.937	20	>0.05

*. Significant at 0.05, D1, MicroFET2 first session; D2, MicroFET2 second session; DS, dominant side; NS, non-dominant side; HAB, hip abduction; HAD, hip adduction.

Appendix E. 2 - Normality Tests of the STAPH data in phase two.

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
S1_DS_HFL	0.226	18	0.02	0.896	18	>0.05
S1_DS_HEX	0.189	18	>0.05	0.877	18	0.02
S1_DS_HAB	0.183	18	>0.05	0.911	18	>0.05
S1_DS_HAD	0.197	18	>0.05	0.777	18	0.00
S1_DS_HIR	0.182	18	>0.05	0.938	18	>0.05
S1_DS_HER	0.193	18	>0.05	0.847	18	0.01
S1_DS_KFL	0.176	18	>0.05	0.884	18	0.03
S2_DS_HFL	0.189	18	>0.05	0.889	18	0.04
S2_DS_HEX	0.180	18	>0.05	0.838	18	0.01
S2_DS_HAB	0.152	18	>0.05	0.952	18	>0.05
S2_DS_HAD	0.239	18	0.01	0.772	18	0.00
S2_DS_HIR	0.202	18	0.05	0.921	18	>0.05
S2_DS_HER	0.192	18	>0.05	0.807	18	0.00
S2_DS_KFL	0.153	18	>0.05	0.884	18	0.03
S1_ND_HFL	0.212	18	0.03	0.907	18	>0.05
S1_ND_HEX	0.180	18	>0.05	0.940	18	>0.05
S1_ND_HAB	0.115	18	>0.05	0.954	18	>0.05
S1_ND_HAD	0.184	18	>0.05	0.884	18	0.03
S1_ND_HIR	0.156	18	>0.05	0.940	18	>0.05
S1_ND_HER	0.175	18	>0.05	0.914	18	>0.05
S1_ND_KFL	0.120	18	>0.05	0.926	18	>0.05
S2_ND_HFL	0.188	18	>0.05	0.847	18	0.01
S2_ND_HEX	0.183	18	>0.05	0.917	18	>0.05
S2_ND_HAB	0.152	18	>0.05	0.929	18	>0.05
S2_ND_HAD	0.131	18	>0.05	0.943	18	>0.05
S2_ND_HIR	0.173	18	>0.05	0.916	18	>0.05
S2_ND_HER	0.130	18	>0.05	0.913	18	>0.05
S2_ND_KFL	0.149	18	>0.05	0.945	18	>0.05

a. Lilliefors Significance Correction; S1, first session; S2, second session; DS, dominant side; NS, non-dominant side; HFL, hip flexion; HEX, hip extension; HAB, hip abduction; HAD, hip adduction; HIR, hip internal rotation; HER, hip external rotation; KFL, knee flexion

Appendix E. 3 - Muscle torque values of the dominant and non-dominant sides obtained by the Biodex in phase one.

Measurement (*N·m*)		Dominant (*n*=20)		Non-dominant (*n*=20)		Difference				
		M	*SD*	*M*	*SD*	*M*	*SD*	%	*t*	*p*
Isometric at 0°	**HAB**	114.41	41.67	116.55	36.57	-2.14	18.73	1.9	-0.510	0.616
	HAD	61.16	32.18	64.00	35.65	-2.84	14.33	4.6	-0.886	0.387
Isometric at 15°	**HAB**	113.70	41.56	113.26	35.41	0.45	20.67	-0.4	0.096	0.924
	HAD	70.00	31.12	68.56	30.03	1.39	28.09	-2.1	0.221	0.827
Isometric at 30°	**HAB**	114.95	43.13	109.94	35.24	5.01	17.50	-4.4	1.689	0.216
	HAD	106.89	38.54	106.39	39.19	0.50	27.01	-0.5	0.830	0.935
Isokinetic at 60°/s	**HAB**	130.59	38.49	124.15	38.12	6.44	16.71	-4.9	1.720	0.101
	HAD	81.41	48.09	98.20	45.79	-16.79	36.81	20.6	-2.040	0.056
Isokinetic at 120°/s	**HAB**	135.30	49.32	135.38	46.61	-0.08	24.06	0.1	-0.015	0.988
	HAD	62.52	43.95	72.88	36.34	-10.35	27.94	16.6	-1.656	0.114

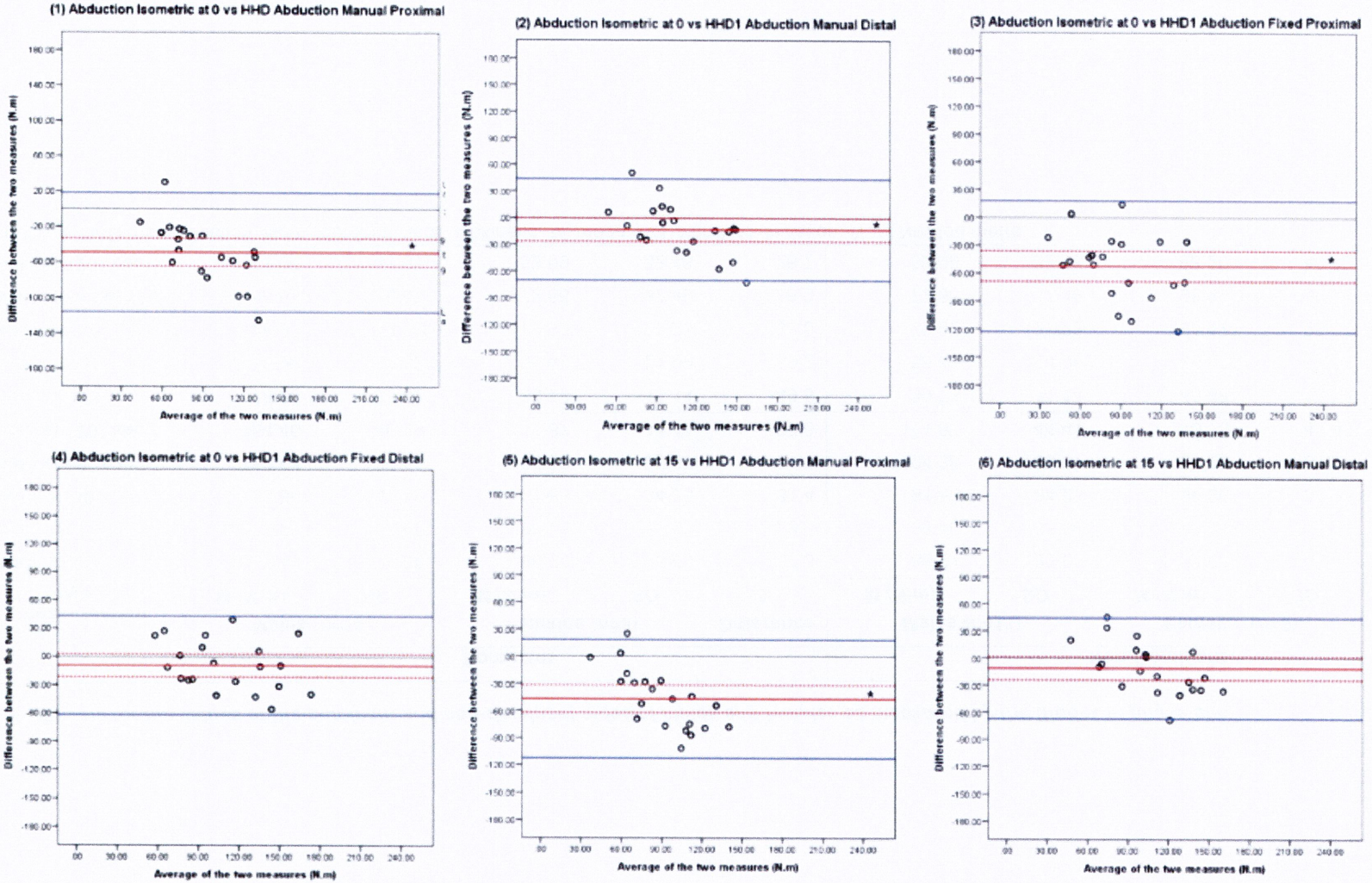
HAB, hip abduction; HAD, hip adduction; *M*, mean; *SD*, standard deviation; df, degrees of freedom; ηp^2, partial eta square; MD, mean of difference.

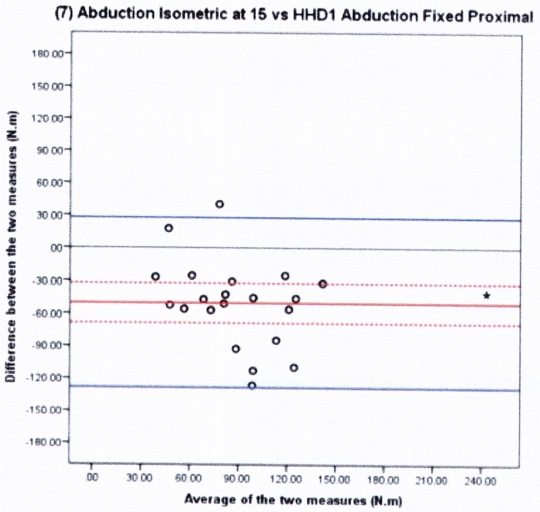
Appendix E. 4 - Muscle torque values of the dominant and non-dominant sides of males and females as measured by the Biodex in phase one.

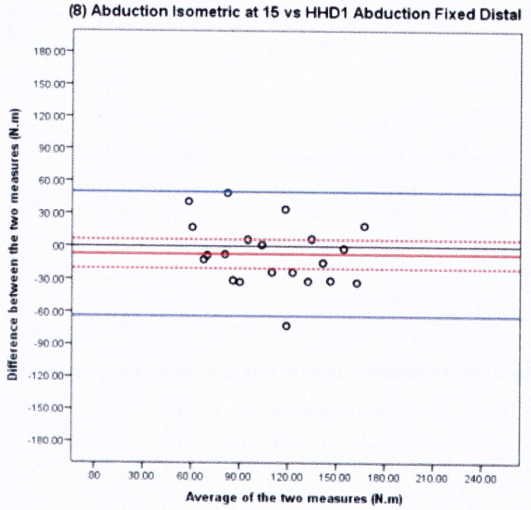
		Dominant					Non-Dominant				
		Males (n=11)		Females (n=9)		Difference	Males (n=11)		Females (n=9)		Difference
Muscle	Measurement ($N \cdot m$)	*M (N.m)*	*SD*	*M (N.m)*	*SD*	*%*	*M (N.m)*	*SD*	*M (N.m)*	*SD*	*%*
HAB	**Isom. at 0°**	137.49	38.68	86.21	25.05	-37.3	86.21	25.05	94.04	20.94	9.1
	Isom. at 15°	130.26	39.99	93.46	35.56	-28.3	93.46	35.56	88.13	27.46	-5.7
	Isom. at 30°	133.86	41.50	91.83	34.23	-31.4	91.83	34.23	90.98	31.98	-0.9
	Isok. at 60°.sec^{-1}	152.04	30.62	104.38	30.67	-31.3	104.38	30.67	106.79	31.21	2.3
	Isok. at 120°.sec^{-1}	154.46	46.29	111.87	44.44	-27.6	111.87	44.44	116.78	41.45	4.4
HAD	**Isom. at 0°**	78.69	33.42	39.73	11.14	-49.5	39.73	11.14	40.88	18.16	2.9
	Isom. at 15°	80.96	32.32	56.49	24.99	-30.2	56.49	24.99	49.83	24.98	-11.8
	Isom. at 30°	128.92	32.29	79.97	27.23	-38.0	79.97	27.23	81.17	29.24	1.5
	Isok. at 60°.sec^{-1}	88.56	56.07	72.66	37.49	-18.0	72.66	37.49	95.49	35.71	31.4
	Isok. at 120°.sec^{-1}	71.06	51.61	52.09	32.18	-26.7	52.09	32.18	68.20	40.72	30.9

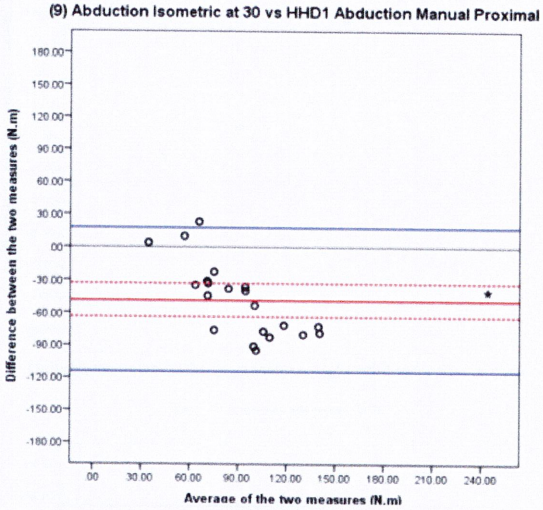
HAB, hip abduction; HAD, hip adduction; Isom, isometric; Isok, isokinetic; *M,* mean; *SD,* standard deviation; $N \cdot m$, Newton metre.

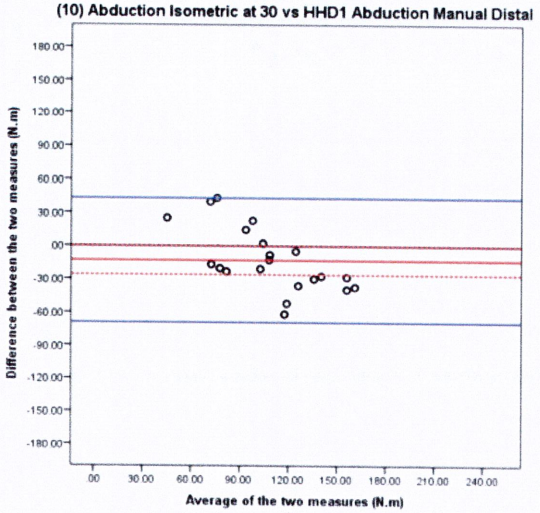
Appendix E. 5 - Bland-Altman agreement plots between the Biodex and MicroFET2 of phase one.

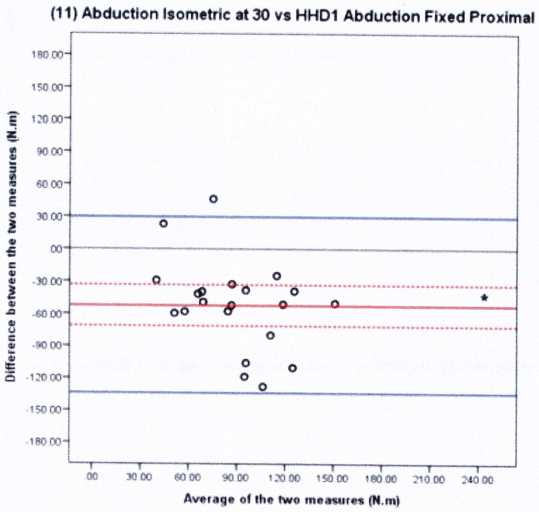


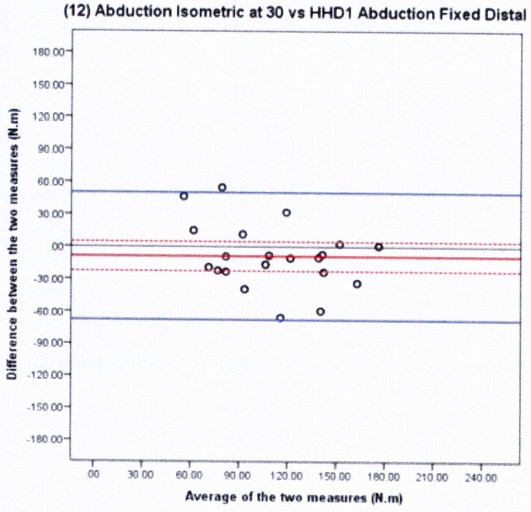

(7) Abduction Isometric at 15 vs HHD1 Abduction Fixed Proximal
Difference between the two measures (N.m)
Average of the two measures (N.m)


(8) Abduction Isometric at 15 vs HHD1 Abduction Fixed Distal
Difference between the two measures (N.m)
Average of the two measures (N.m)

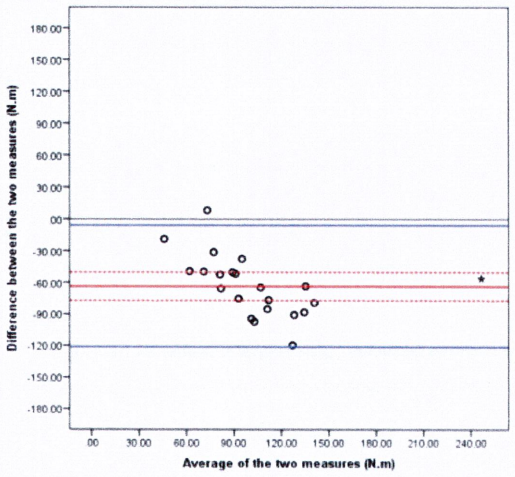

(9) Abduction Isometric at 30 vs HHD1 Abduction Manual Proximal
Difference between the two measures (N.m)
Average of the two measures (N.m)


(10) Abduction Isometric at 30 vs HHD1 Abduction Manual Distal
Difference between the two measures (N.m)
Average of the two measures (N.m)

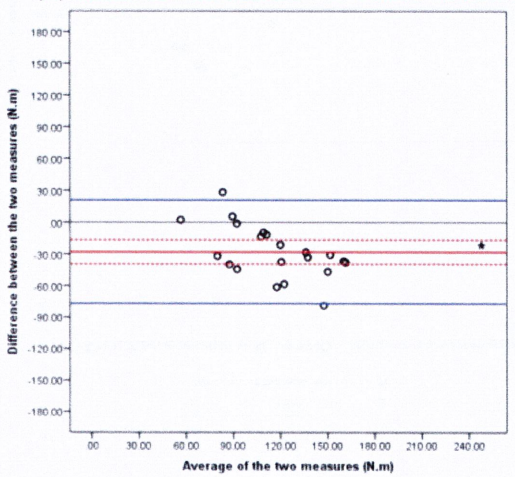

(11) Abduction Isometric at 30 vs HHD1 Abduction Fixed Proximal
Difference between the two measures (N.m)
Average of the two measures (N.m)


(12) Abduction Isometric at 30 vs HHD1 Abduction Fixed Distal
Difference between the two measures (N.m)
Average of the two measures (N.m)

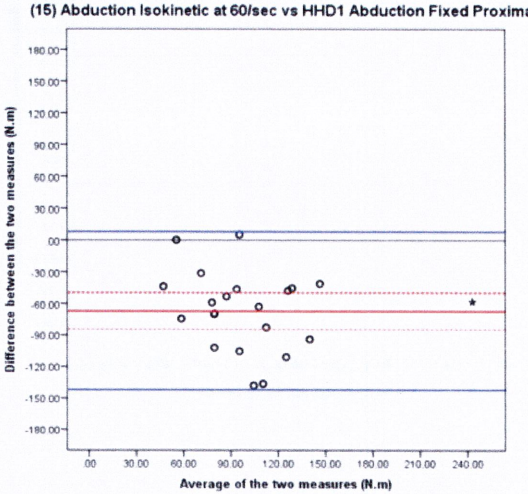
(13) Abduction Isokinetic at 60/sec vs HHD1 Abduction Manual Proximal



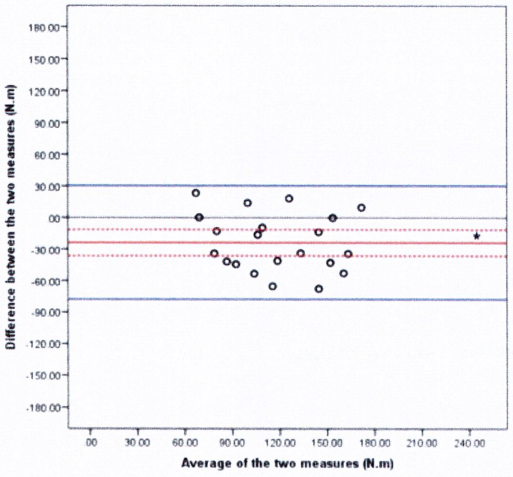
(14) Abduction Isokinetic at 60/sec vs HHD1 Abduction Manual Distal



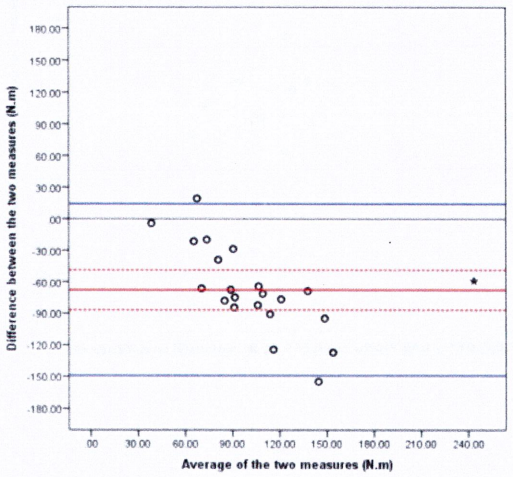
(15) Abduction Isokinetic at 60/sec vs HHD1 Abduction Fixed Proximal



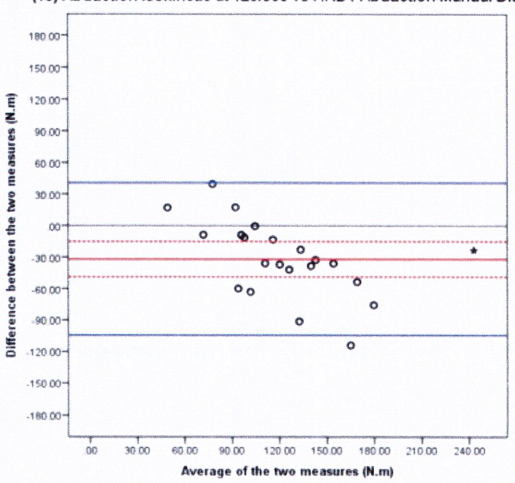
(16) Abduction Isokinetic at 60/sec vs HHD1 Abduction Fixed Distal



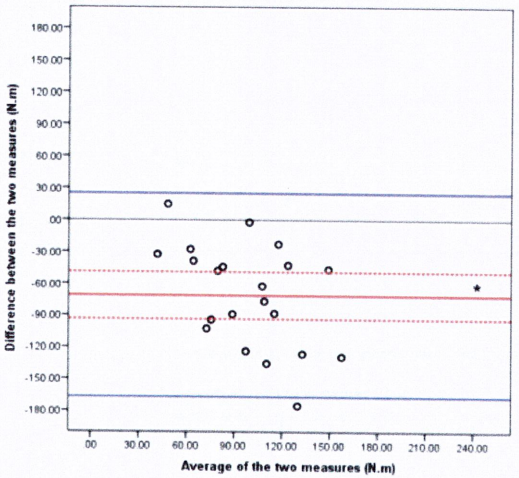
(17) Abduction Isokinetic at 120/sec vs HHD1 Abduction Manual Proximal



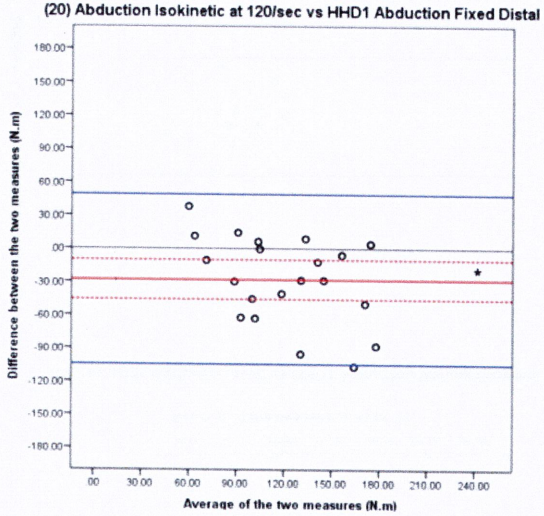
(18) Abduction Isokinetic at 120/sec vs HHD1 Abduction Manual Distal



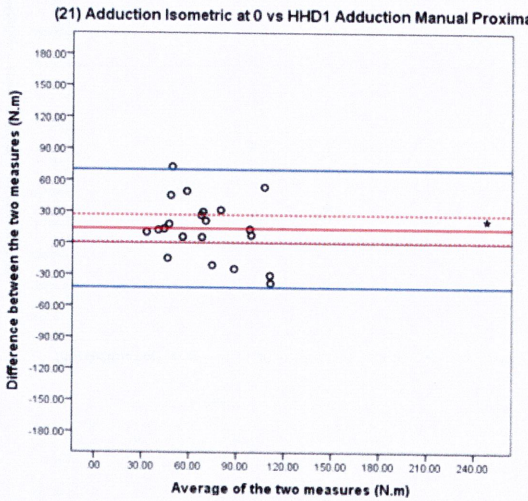
(19) Abduction Isokinetic at 120/sec vs HHD1 Abduction Fixed Proximal



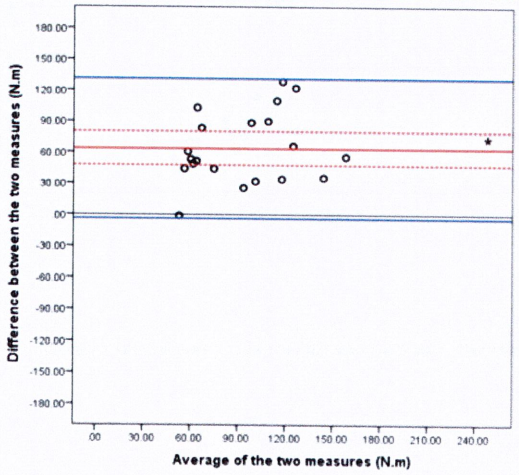
(20) Abduction Isokinetic at 120/sec vs HHD1 Abduction Fixed Distal



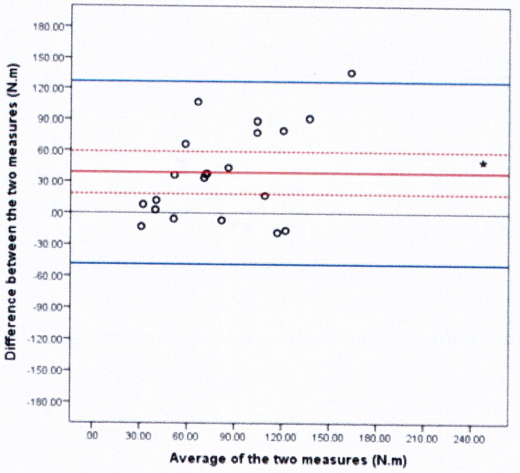
(21) Adduction Isometric at 0 vs HHD1 Adduction Manual Proximal



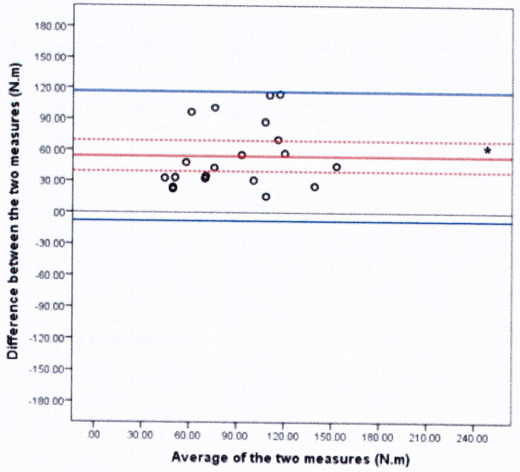
(22) Isometric Adduction at 0 vs HHD1 Adduction Manual Distal

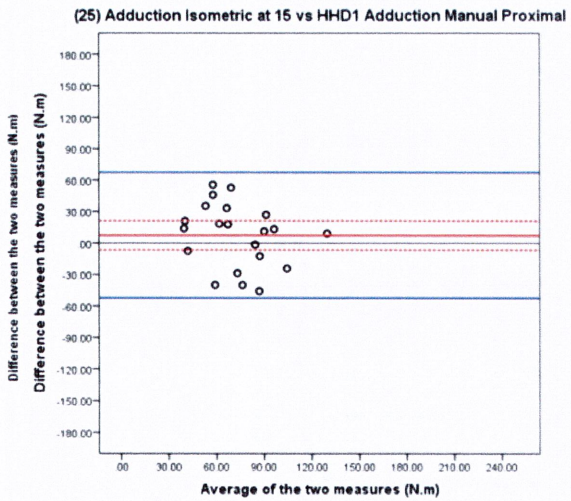


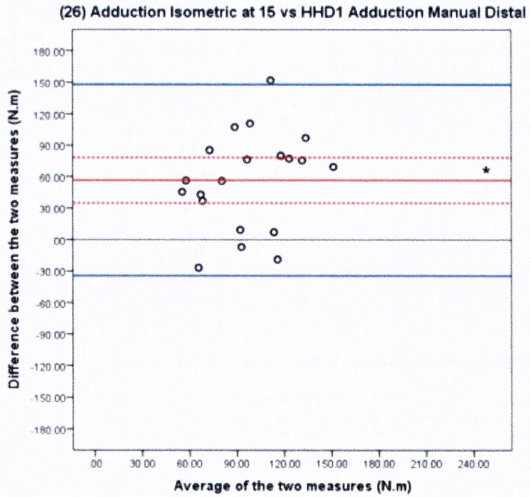
(23) Adduction Isometric at 0 vs HHD1 Adduction Fixed Proximal

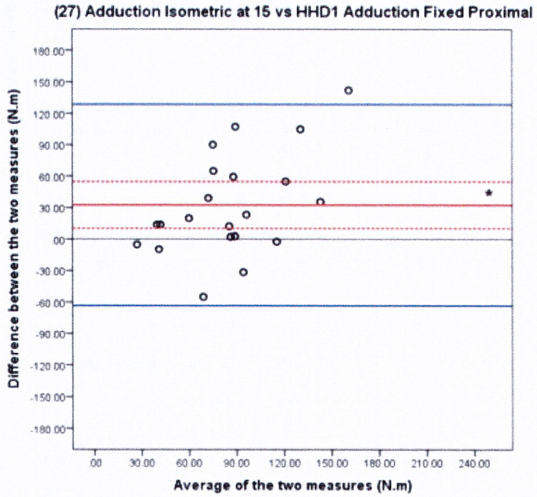


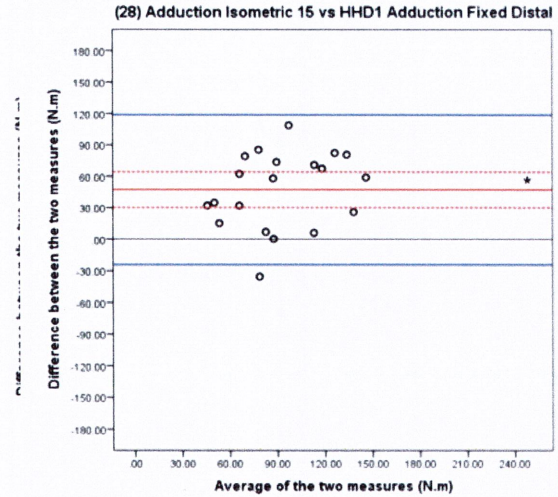
(24) Adduction Isometric at 0 vs HHD1 Adductor Fixed Distal

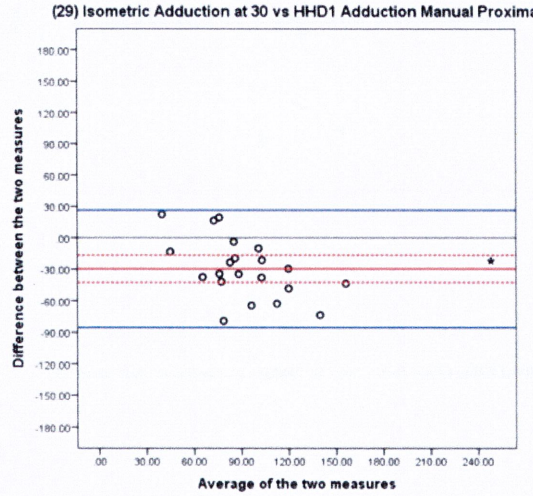


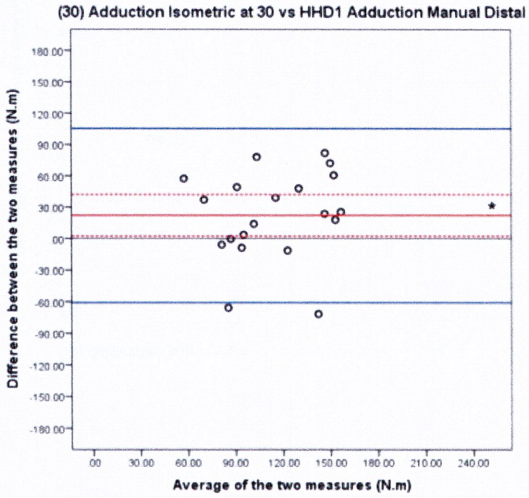

(25) Adduction Isometric at 15 vs HHD1 Adduction Manual Proximal
Difference between the two measures (N.m)
Average of the two measures (N.m)

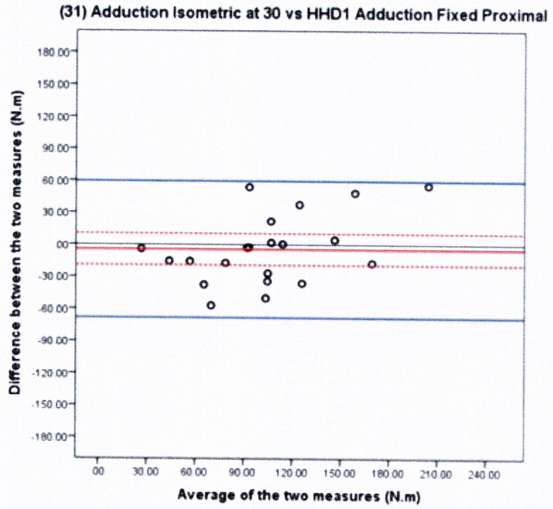

(26) Adduction Isometric at 15 vs HHD1 Adduction Manual Distal
Difference between the two measures (N.m)
Average of the two measures (N.m)

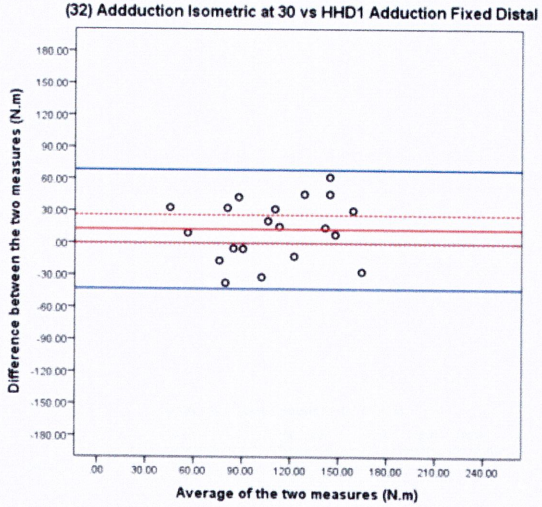

(27) Adduction Isometric at 15 vs HHD1 Adduction Fixed Proximal
Difference between the two measures (N.m)
Average of the two measures (N.m)

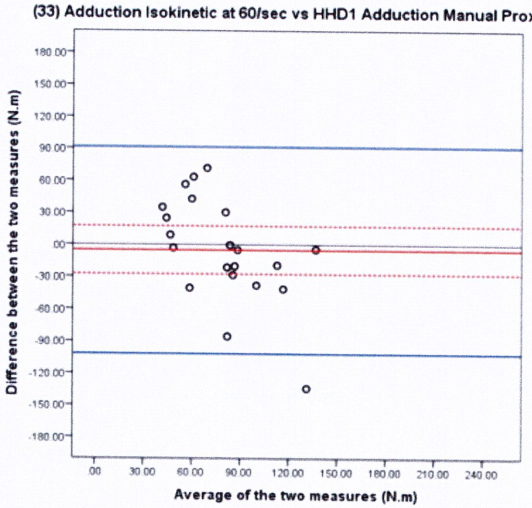

(28) Adduction Isometric 15 vs HHD1 Adduction Fixed Distal
Difference between the two measures (N.m)
Average of the two measures (N.m)

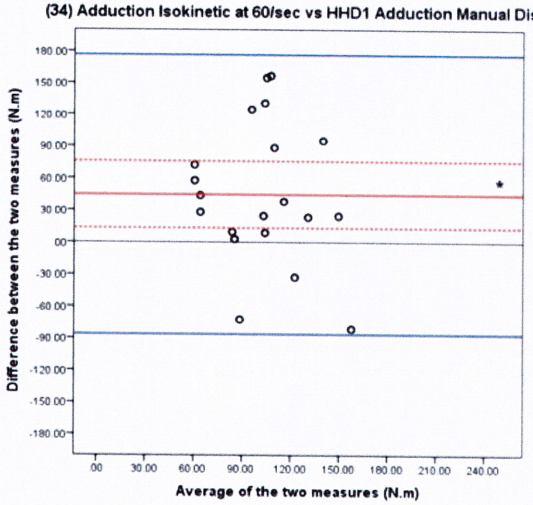

(29) Isometric Adduction at 30 vs HHD1 Adduction Manual Proximal
Difference between the two measures
Average of the two measures

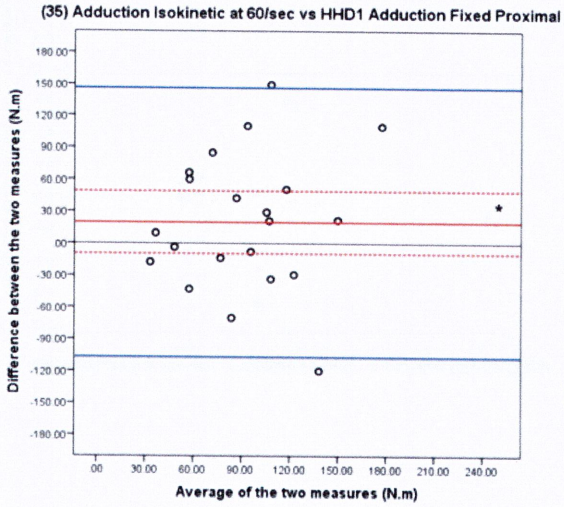

(30) Adduction Isometric at 30 vs HHD1 Adduction Manual Distal
Difference between the two measures (N.m)
Average of the two measures (N.m)

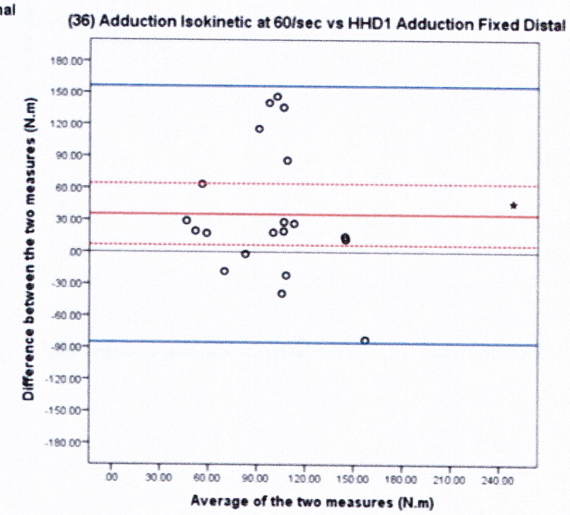

(31) Adduction Isometric at 30 vs HHD1 Adduction Fixed Proximal
Difference between the two measures (N.m)
Average of the two measures (N.m)


(32) Addduction Isometric at 30 vs HHD1 Adduction Fixed Distal
Difference between the two measures (N.m)
Average of the two measures (N.m)

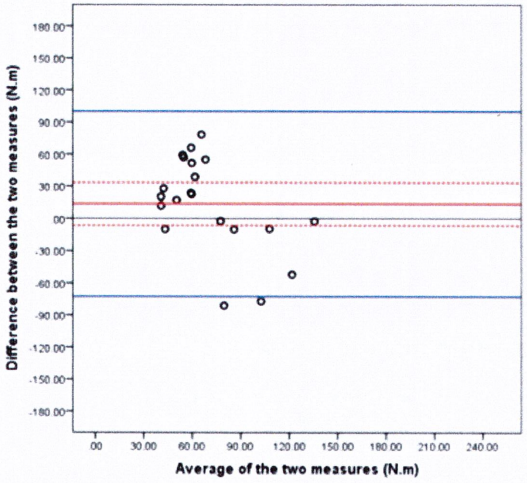

(33) Adduction Isokinetic at 60/sec vs HHD1 Adduction Manual Proximal
Difference between the two measures (N.m)
Average of the two measures (N.m)


(34) Adduction Isokinetic at 60/sec vs HHD1 Adduction Manual Distal
Difference between the two measures (N.m)
Average of the two measures (N.m)

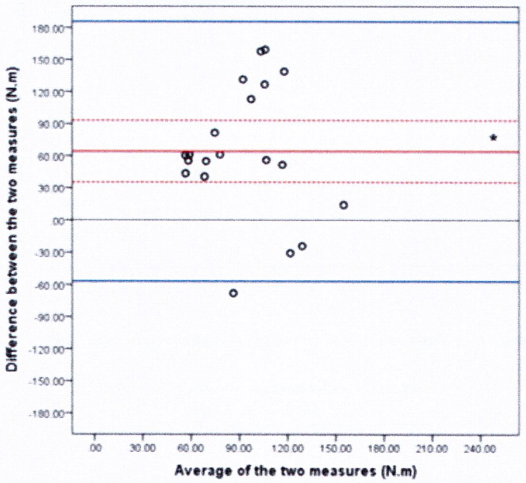

(35) Adduction Isokinetic at 60/sec vs HHD1 Adduction Fixed Proximal
Difference between the two measures (N.m)
Average of the two measures (N.m)


(36) Adduction Isokinetic at 60/sec vs HHD1 Adduction Fixed Distal
Difference between the two measures (N.m)
Average of the two measures (N.m)

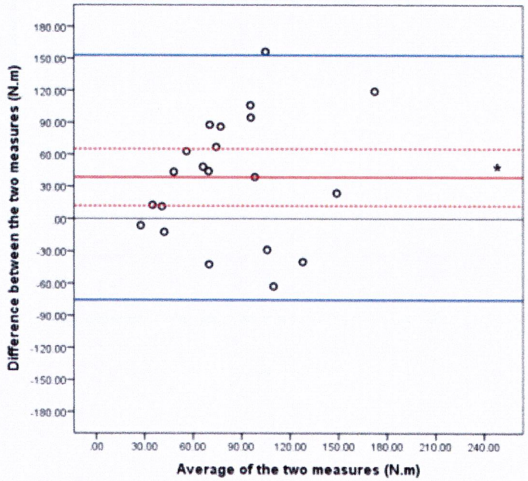
(37) Adduction Isokinetic at 120/sec vs HHD1 Adduction Manual Proximal



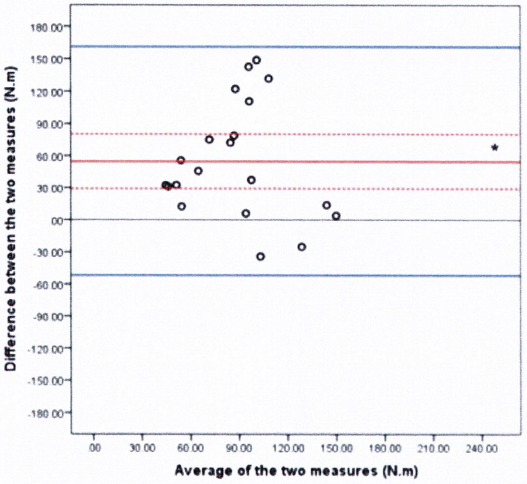
(38) Adduction Isokinetic at 120/sec vs HHD1 Adduction Manual Distal



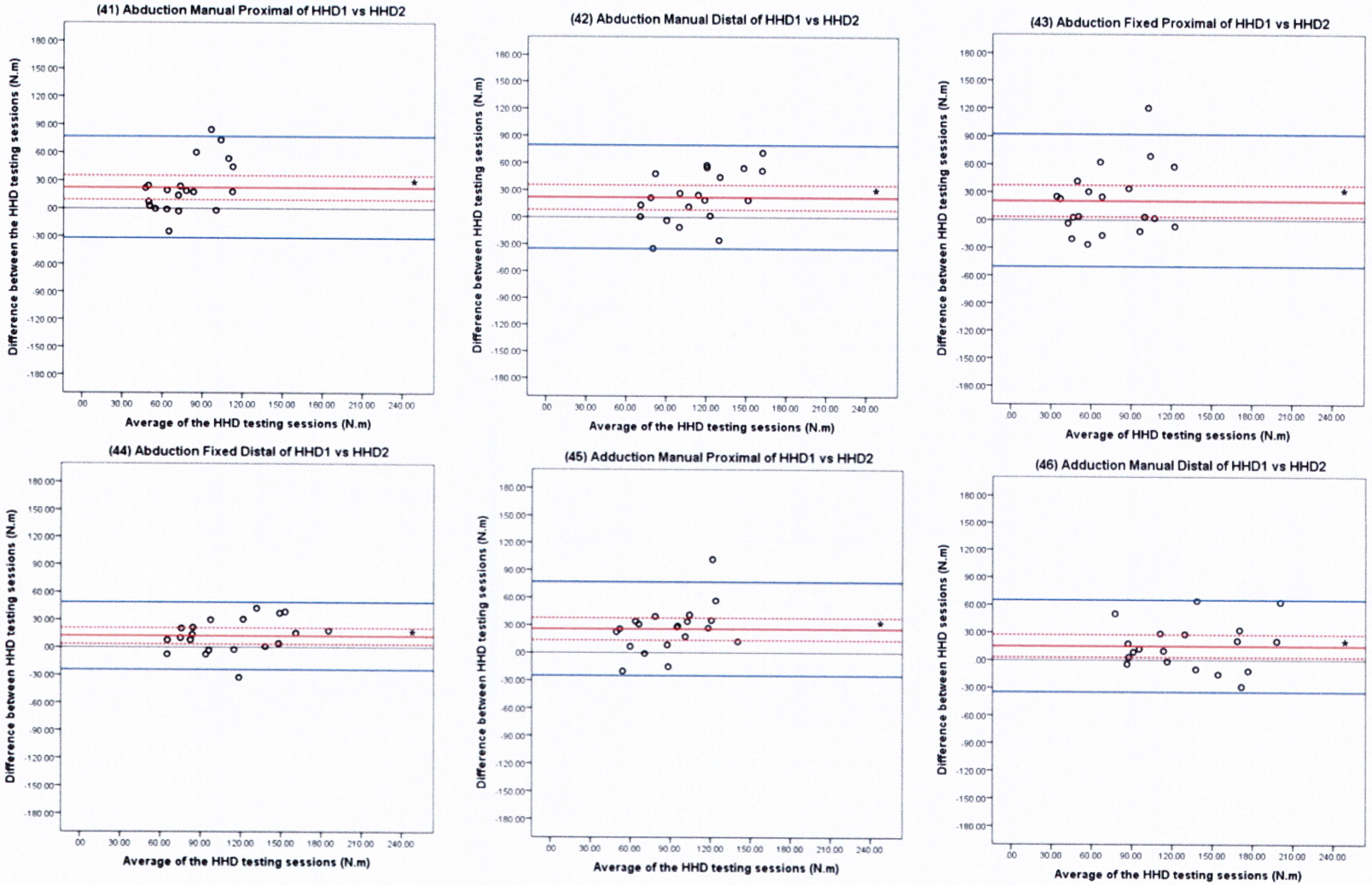
(39) Adduction Isokinetic at 120/sec vs HHD1 Adduction Fixed Proximal

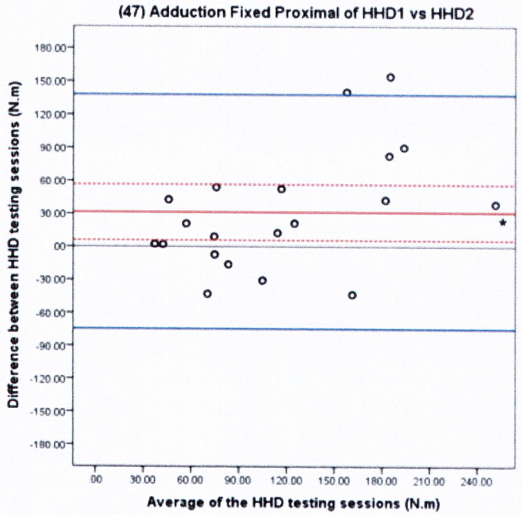


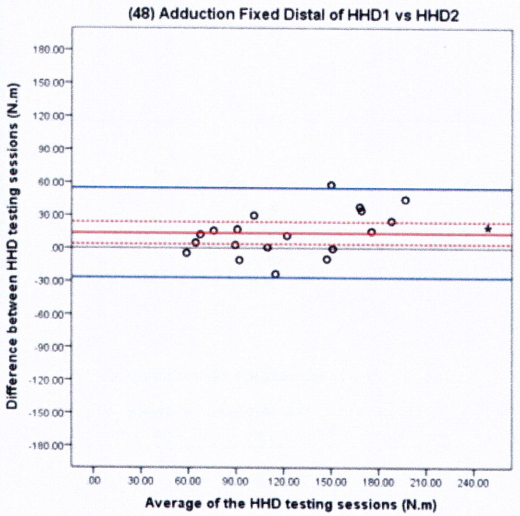
(40) Adduction Isokinetic at 120/sec vs HHD1 Adduction Fixed Distal



Appendix E. 6 - Bland-Altman agreement plots between the two MicroFET2 sessions of phase one.




(47) Adduction Fixed Proximal of HHD1 vs HHD2
Difference between HHD testing sessions (N.m)
180.00
150.00
120.00
90.00
60.00
30.00
.00
-30.00
-60.00
-90.00
-120.00
-150.00
-180.00
.00
30.00
60.00
90.00
120.00
150.00
180.00
210.00
240.00
Average of the HHD testing sessions (N.m)


(48) Adduction Fixed Distal of HHD1 vs HHD2
Difference between HHD testing sessions (N.m)
180.00
150.00
120.00
90.00
60.00
30.00
.00
-30.00
-60.00
-90.00
-120.00
-150.00
-180.00
.00
30.00
60.00
90.00
120.00
150.00
180.00
210.00
240.00
Average of the HHD testing sessions (N.m)

Appendix E. 7 - Muscle torque values of the dominant and non-dominant sides obtained by the MicroFET2 in phase one.

			Session 1							Session 2						
			Dominant		Non-Dominant		Difference			Dominant		Non-Dominant		Difference		
Stab.	Mov.	Loc.	*M*	*SD*	*M*	*SD*	*M*	*SD*	%	*M*	*SD*	*M*	*SD*	*M*	*SD*	%
MAN	HAB	PROX	66.26	18.48	60.56	15.84	5.70	14.70	-8.6	88.39	32.06	86.74	29.80	1.66	14.60	-1.9
		DIS	101.86	24.52	101.93	23.83	-0.08	15.69	0.1	124.16	38.62	126.02	36.69	-1.85	9.40	1.5
	HAD	PROX	76.77	23.89	80.30	31.66	-3.53	18.09	4.6	102.56	34.79	112.27	36.70	-9.71	29.24	9.5
		DIS	119.11	47.99	121.55	40.44	-2.44	46.73	2.1	141.69	42.41	139.32	43.94	2.37	11.96	-1.7
FIX	HAB	PROX	62.74	28.95	72.92	34.50	-10.18	29.49	16.2	83.66	38.46	82.92	40.08	0.74	19.79	-0.9
		DIS	105.98	32.98	106.64	30.99	-0.67	16.08	0.6	118.15	38.93	114.07	31.48	4.08	15.69	-3.5
	HAD	PROX	101.93	51.79	99.82	52.97	2.11	26.84	-2.1	132.62	77.93	131.21	71.52	1.41	32.00	-1.1
		DIS	110.14	45.78	118.65	38.56	-8.51	50.00	7.7	131.70	49.94	130.50	49.45	1.20	27.28	-0.9

MAN, manual, FIX, fixed; Stab., stabilisation; Mov., movement; Loc., location; PROX, proximal; DIS, distal; HAB, abduction; HAD, hip adduction; *M,* mean; *SD,* standard deviation.

Appendix E. 8 - Muscle torque values of the dominant and non-dominant sides obtained by the MicroFET2 in phase two.

Muscle group	Session 1 (*n*=18)								Session 2 (*n*=18)							
	Dominant (1)		Non-Dominant (2)		Difference (2-1)				Dominant (1)		Non-Dominant (2)		Difference (2-1)			
	M (N.m)	*SD*	*M (N.m)*	*SD*	*M (N.m)*	*SD*	*p*	*%*	*M (N.m)*	*SD*	*M (N.m)*	*SD*	*M (N.m)*	*SD*	*p*	*%*
HFL	56.10	16.21	52.12	14.89	4.34	10.50	0.132	-7.1	57.20	14.43	54.76	14.34	2.63	7.91	0.217	-4.3
HEX	55.23	22.39	52.57	17.32	3.69	10.99	0.289	-4.8	56.59	17.34	54.52	17.60	3.48	9.39	0.245	-3.7
HAB	75.17	20.18	76.09	21.04	-0.73	10.20	0.713	1.2	75.67	18.74	74.76	18.72	1.48	8.10	0.630	-1.2
HAD	59.89	19.45	58.36	14.98	1.15	11.34	0.579	-2.6	64.89	18.69	62.03	15.36	2.45	9.92	0.243	-4.4
HIR	68.13	17.61	68.33	19.47	-0.96	11.94	0.944	0.3	68.28	15.93	67.76	19.31	-0.66	14.27	0.873	-0.8
HER	51.87	14.28	55.27	16.90	-2.73	10.08	0.165	6.5	54.63	13.94	55.72	15.99	0.48	9.33	0.492	2.0
KFL	125.19	27.71	130.08	33.48	-7.52	18.79	0.193	3.9	126.27	25.75	129.47	31.24	-6.12	18.04	0.316	2.5

HFL, hip flexion; HEX, hip extension; HAB, hip abduction; HAD, hip adduction; HIR, hip internal rotation; HER, hip external rotation; KFL, knee flexor; *M,* mean; *SD,* standard deviation.

Appendix E. 9 - Muscle torque values of the dominant and non-dominant sides of males and females as measured by the MicroFET2 in phase two.

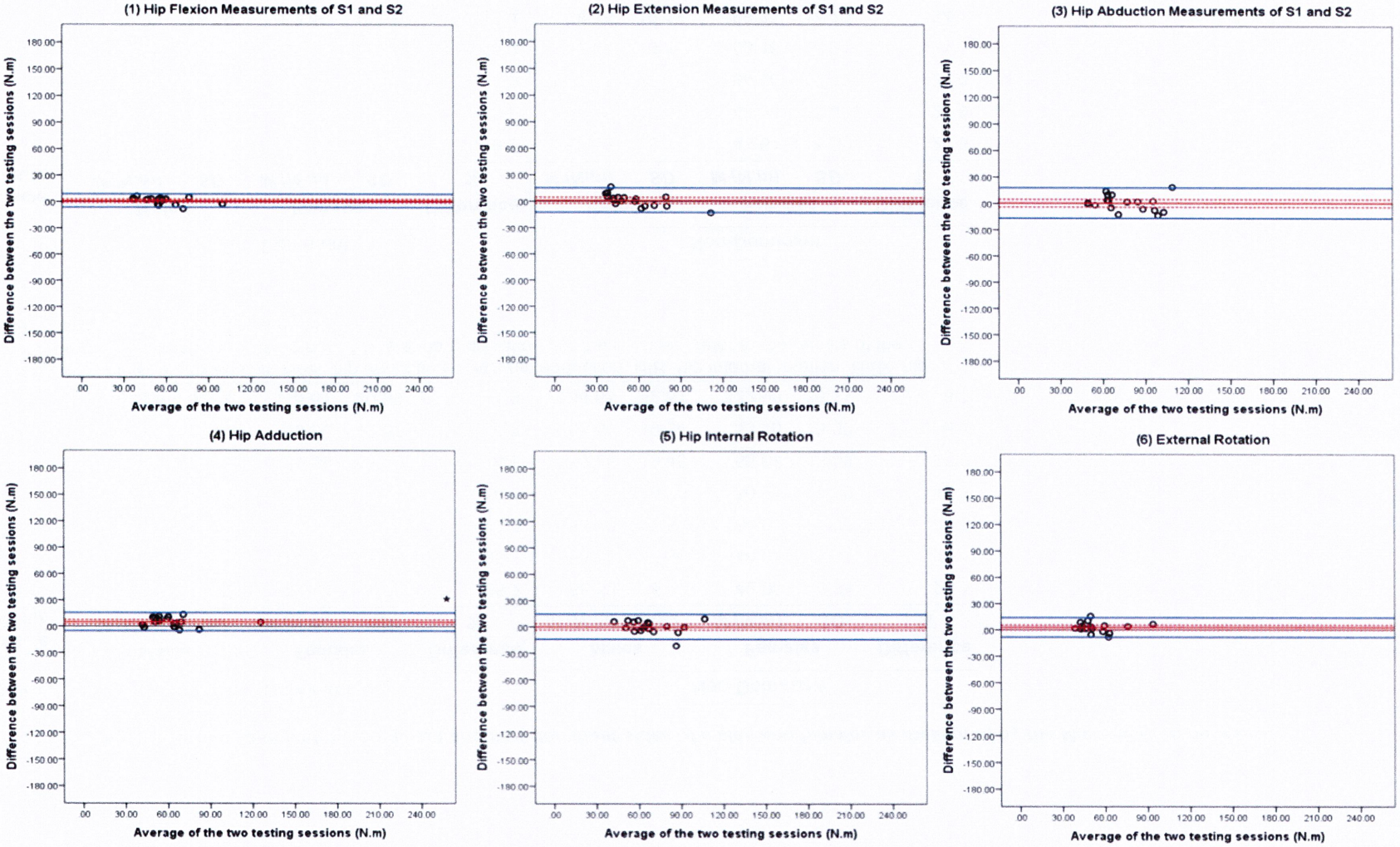
Session 1	Dominant					Non-Dominant				
Muscle group	Males		Females		Difference	Males		Females		Difference
	M (N.m)	*SD*	*M (N.m)*	*SD*	*%*	*M (N.m)*	*SD*	*M (N.m)*	*SD*	*%*
HFL	62.18	17.22	46.54	8.70	-25.2	58.52	15.27	42.06	6.81	-28.1
HEX	61.20	24.38	45.86	16.18	-25.1	58.00	17.96	44.04	13.17	-24.1
HAB	81.91	17.83	64.59	20.25	-21.2	83.40	17.83	64.61	21.75	-22.5
HAD	62.31	23.59	56.10	10.79	-10.0	63.09	16.76	50.91	7.89	-19.3
HIR	72.23	16.75	61.69	18.20	-14.6	75.75	19.45	56.67	13.58	-25.2
HER	50.65	15.69	53.80	12.67	6.2	56.61	19.34	53.16	13.30	-6.1
KFL	131.28	30.76	115.61	20.59	-11.9	134.65	32.61	122.90	36.11	-8.7

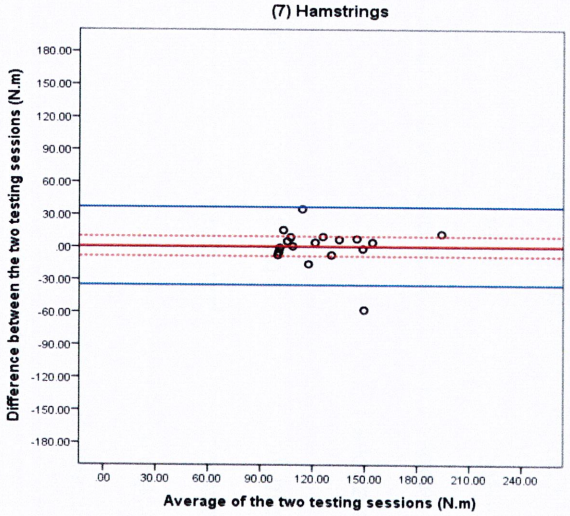
HFL, hip flexion; HEX, hip extension; HAB, hip abduction; HAD, hip adduction; HIR, hip internal rotation; HER, hip external rotation; KFL, knee flexor; *M,* mean; *SD,* standard deviation; for the p values, refer to the results of the repeated measures ANOVA in section 5.3.3.

Session 2	Dominant					Non-Dominant				
Muscle group	Males		Females		Difference	Males		Females		Difference
	M (N.m)	*SD*	*M (N.m)*	*SD*	*%*	*M (N.m)*	*SD*	*M (N.m)*	*SD*	*%*
HFL	61.98	15.84	49.69	8.02	-19.8	60.55	15.18	45.64	6.30	-24.6
HEX	60.29	18.22	50.77	15.31	-15.8	61.18	17.57	44.04	12.40	-28.0
HAB	82.40	17.77	65.10	15.97	-21.0	81.05	17.48	64.86	17.23	-20.0
HAD	68.95	21.52	58.50	11.84	-15.2	67.67	16.45	53.16	8.21	-21.5
HIR	72.03	18.19	62.40	10.05	-13.4	74.75	19.13	56.77	14.70	-24.1
HER	54.54	15.56	54.79	12.11	0.5	56.12	17.87	55.09	13.84	-1.8
KFL	131.69	27.55	117.74	21.79	-10.6	133.47	31.04	123.17	32.91	-7.7

HFL, hip flexion; HEX, hip extension; HAB, hip abduction; HAD, hip adduction; HIR, hip internal rotation; HER, hip external rotation; KFL, knee flexor; *M,* mean; *SD,* standard deviation; for the p values, refer to the results of the repeated measures ANOVA in section 5.3.3.

Appendix E. 10 - Bland-Altman agreement plots between the two MicroFET2 sessions of phase two.

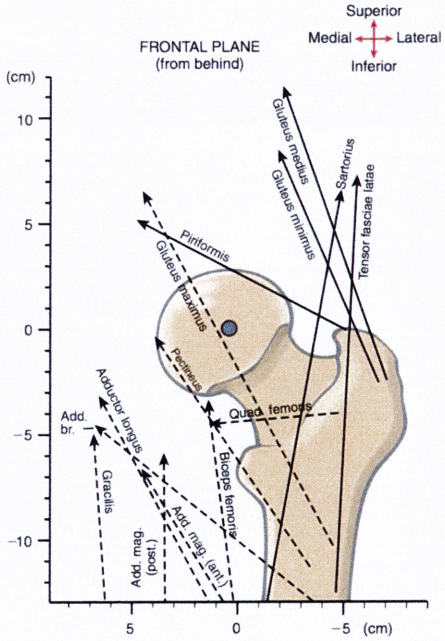



(7) Hamstrings
Difference between the two testing sessions (N.m)
180.00
150.00
120.00
90.00
60.00
30.00
.00
-30.00
-60.00
-90.00
-120.00
-150.00
-180.00
.00
30.00
60.00
90.00
120.00
150.00
180.00
210.00
240.00
Average of the two testing sessions (N.m)

Appendix F Chapter 7 Supporting Material

The Proposed Biomechanical Model of the Hip

The findings of Chapter 6 suggest that the distal fixed location is superior to the proximal location in testing. Furthermore, the muscle strength ratios in the MicroFET2 testing of phase one were higher than 1 (HAD to HAB) compared to Biodex testing. Several factors influence the magnitude and direction of the torque produced around the hip joint. The hip adductor, abductor and part of the extensor muscles control hip motion in the frontal plane through complex muscular contractions (Appendix F. 1).



Appendix F. 1 - Posterior view of the femoral head showing the force lines of muscles crossing the right hip in the frontal plane only (Neumann, 2009b).

Movement in the human body is based on three classes of the lever system that are used to move a body segment (resistance) (Appendix F. 2). The joint acts as a fulcrum, the bone acts as a rigid bar (lever), and muscle contraction produces force (Hall, 2003). Each muscle has a unique anatomical properties and direction of pull. Some muscles insert proximal in the femur while others go beyond the knee joint and insert into the tibia (Appendix A. 1). Such orientation affects force production when an external resistance is applied.

In a first-class lever system, the resistance and force are on opposite sides of the fulcrum. In the hip, none of these muscle groups form a first-class lever system (Appendix F. 2). In a second-class system, the applied force and resistance are on the same side of the axis, with the resistance moment arm ($MA_{resistance}$) being shorter. It results in mechanical advantage since $MA_{force} > MA_{resistance}$. In the third-class system, the force and resistance are also on the same side of the force moment arm (MA_{force}) being

shorter leading to a mechanical disadvantage because of $MA_{force} < MA_{resistance}$ (Hall, 2003). The proposed biomechanical model of the hip joint shows the direction of muscular pull in the frontal plane, resulting in hip abduction and adduction (Appendix F. 2).

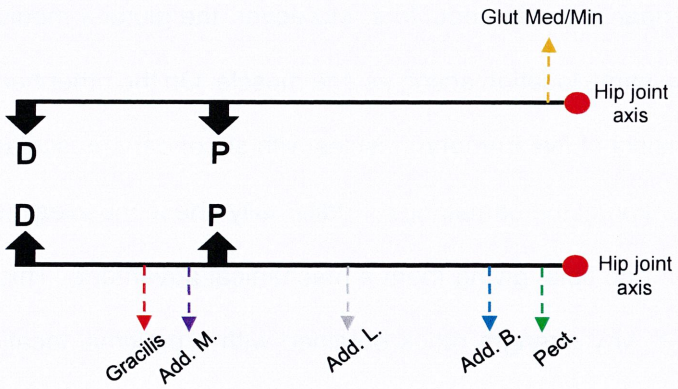
When the MicroFET2 was applied on a proximal location just above the knee, the force of pull created by the gluteus medius and minimus resulted in a third-class system (Appendix F. 2). All adductor muscles formed a second-class system except the gracilis and adductor magnus, which formed a third-class system (Appendix F. 2). As a result, all abductor and three-fifths (60%) of the adductor muscles were in a mechanical disadvantage because of the shorter MA_{force}. Shifting the location of stabilisation distally resulted in a mechanical disadvantage for both muscle groups forming a third-class lever system (Appendix F. 2).

Placing the dynamometer on a proximal location resulted in a 40% difference between the abductors and adductors based on the number of muscles. The difference between a true clinical measurement and the biomechanical model is due to the muscle cross section, secondary muscular activity, type and orientation of muscle fibres (Methenitis, et al., 2016). The abductor muscles do not possess a wide cross section or secondary muscle action compared to the adductors. Moreover, the gluteus medius and minimus insert in almost the same location acting as one muscle. On the other hand, the adductor muscle group consists of five primary muscles with secondary muscular activity arising from the medial portion of the hamstrings. Additionally, these muscles insert into varying levels of the femur and tibia, giving them a mechanical advantage. This biomechanical model may explain why strength ratios obtained with a proximal location were higher than the distal location in phase one.

The same biomechanical principles apply in phase two of testing. However, the results show that the HAD-HAB ratios were <1 in both MicroFET2 sessions. It may be attributed to the enhanced stability of the STAPH, which reflected on the accuracy of testing, reduced measurement error and a smaller difference between the two sessions.

Moreover, the participants in phase two were not allowed to stabilise themselves using their upper or lower extremities. It may have resulted in the reduced adductor torque values.

The proximal location of stabilisation was also used in Biodex testing. Even though, there is a marked difference between strength ratios obtained using the Biodex and those obtained using the MicroFET2. Having the participants in a standing position during Biodex testing added a few more elements that need to be addressed. First of all, the body weight adds more resistance to the pelvic stabilisers and the contracting muscles. For instance, the contralateral muscles are activated to keep the pelvis properly aligned and stable while the other side is being tested. As mentioned earlier, the gluteus medius and minimus share the role of abducting and internally rotating the hip based on the activated muscle fibres. Isolating abduction from internal rotation during testing is not an easy task. Second, limb weight and joint reaction force should also be considered. Third, single limb support during testing is challenging for some. These factors should not be neglected when the torque and strength ratios of Biodex testing are reported and applied.



Appendix F. 2 – The second the third lever systems of the hip muscles (Glut, gluteus; med, medius; min, minimus; add, adductor; M., magnus; L, longus; B., brevis; Pect., pectineus).

Appendix G Supplementary Material

Appendix G. 1 - Missing Data

Missing data is a problem that many researchers come across in surveys and longitudinal studies. Several procedures were developed and described to complete data sets ranging from simple guidelines to complicated methods using statistical software packages (Dong & Peng, 2013; Newman, 2014).

If the issue of missing data was ignored, it could lead to inaccurate statistical results, bias and reduced generalisability. Some researchers tend to include only the participants with complete data and exclude those with missing information such as age, weight or occupation. This procedure results in overlooking some important information that may affect the results, increases standard error and reduces power. Bias, in which the results or analysis are systematically different from the population of interest, can be in the form of over- or under-estimation (Newman, 2014).

Error, which is the deviation of the observed value from of the true value, is another consequence of missing data. It results from associating inaccurate standard errors with a parameter in sample sizes that are either too high or too low (Newman, 2014). Both bias and error can lead to substantial deviation from true values and misinterpretation of the results. Although controlling and minimising missing values are crucial before statistical analysis, the decision of handling missing values or not is based on the pattern of distribution or percentage.

- *Patterns of Missing Data*

The degree of bias in analysing the data with missing values is attributed to the pattern of missing values (Sterne, et al., 2009):

- Missing completely at random: There is no relation between the missing and observed values, which means that the probability of missing values does not depend on the observed values.
- Missing at random: There is some relation between the missing and observed values that can be explained. It also means that the probability of missing values partially depends on the observed values.
- Missing not at random: In this case, the pattern of the missing values is associated with the way the observed values are distributed and presented. It means that the missing values depend on the observed values.

- *Handling Missing Data*

Many statistical techniques have been introduced to complete data sets, handle missing data, and reduce the effect of error and bias. They can be divided into two types: techniques that exclude the data of individuals with incomplete information and those that consider filling or completing the missing information. Listwise deletion or complete case analysis involves excluding the sets of data of individuals with one or more missing values from the analysis. Therefore, in longitudinal studies, any participant with incomplete data will be discarded. This technique reduces power and increases bias by over- or underestimating the effect sizes and should be avoided (Newman, 2014). It works best in datasets with few missing values in which the missing completely at random rule is more likely applicable. Nevertheless, it is less favoured with higher percentages of missing values especially when the predictor variables are missing a large number of measurements (Pigott, 2001).

Pairwise deletion or available case analysis is another simple tool that is readily available in most statistical software. It is based on analysing the whole data set without excluding any participant. This method is useful when the researcher is analysing univariate (one variable) datasets to obtain descriptive information about the sample and

the mean of values. It tends to produce problems when bi- or multivariates are included in the analysis. It is because pairwise deletion is based on weakly correlated variables, which means that the absence of the value of one variable will not affect the results of another variable in the same participant. When variables are highly correlated, the results will be secondary to complete case analysis. Therefore, the findings may be unreliable (Pigott, 2001).

Intention to Treat (ITT) analysis is typically used in randomised control trials in which the participants are allocated to an intervention or control groups. This method is applied to maximise the number of participants in the analysis phase. The principle is based on including the participants in their original randomised group regardless if they adhered to it or discontinued to participate in the study (Gupta, 2011, Streiner and Geddes, 2001). Including those who did not receive the treatment at all stages with those who did may affect the efficacy of a drug or an intervention (Gupta, 2011). The secondary data analysis of the hip arthroscopy patients is considered a longitudinal study, not a randomised controlled study. Furthermore, there are a lot of missing values. Hence, this method of handling missing data cannot be applied (Bell et al., 2014).

Last observation carried forward (LOCF) is an alternative to ITT in longitudinal studies with repeated measures. When the participants drop out from a study, some researchers do not exclude their results but attempt to complete the remaining missing values using the last observed (measured) value before dropping out. The drawback of this method is that it assumes no effect of the treatment and time after the last measurement (Streiner and Geddes, 2001).

Single Imputation (SI, impute = fill in) is a technique applied to replace the missing values with logical values based on a set of rules (Andridge and Little, 2010). Examples of this technique include the mean of observed values and regression imputation. By applying this technique, the available data from all participants - regardless of containing missing values or not - are maintained. The most commonly used method here is substituting the missing values with the mean value of that particular variable. Therefore, the

analysis can be carried on without excluding any participant or reducing the sample size. The potential downfall is reducing the variance of variables and smaller SEs, which both come from the assumption that missing values are equal to the mean without considering any variation between what the real value could have been and the imputed value (Pigott, 2001, Sterne et al., 2009, Newman, 2014).

Multiple Imputation (MI) was introduced as a method of avoiding the pitfalls of using listwise or SI in data analysis. Instead of performing a single imputation, the statistical software is programmed to create several and unbiased copies of the observed values combined with the expected missing values (Sterne et al., 2009). Missing values of the original set are then replaced by the imputed values by taking advantage of the different guesses in the sets and the differences in the SEs related to each variable. Thus, it reduces bias and increases the accuracy of hypothesis testing accuracy (Newman, 2014). The recommended number of imputations is debatable and varies from one source to another (Dong and Peng, 2013, Pigott, 2001, Newman, 2014, Graham et al., 2007). The average of the imputed values is taken from all data sets, and the SEs are combined based on the formula suggested by Donald Rubin (Appendix G. 2) (Rubin, 2008). The detailed steps of the imputation model can be found elsewhere (Dong and Peng, 2013).

$$Rubin's\ Formula: SE = \sqrt{\frac{1}{M}\sum_{m=1}^{M} SE_m^2 + \left(1 + \frac{1}{M}\right)\left(\frac{1}{M-1}\right)\sum_{m=1}^{M}(b_m - b^-)^2}$$

Where $\frac{1}{M}\sum_{m=1}^{M} SE_m^2$ is the average squared SE, $\left(\frac{1}{M-1}\right)\sum_{m=1}^{M}(b_m - b^-)^2$ is the variance of the parameter estimates across imputations, and $\left(1 + \frac{1}{M}\right)$ is a correction factor (Newman 2014)

Appendix G. 2 - Rubin's Formula

Hot deck imputation (HDI) methods are a form of multiple imputations. In these methods, the missing values of the participants (known as recipients) are replaced with values from other participants with similar characteristics (known as donors). The donors are selected based on variables that are similar between the donors and recipients. In random hot deck methods, possible donors are gathered or pooled together, and one donor is selected randomly. Even though it may seem easy to apply, it can get tricky when the missing data do not follow a monotone or fixed pattern. Furthermore, copying the information from other participants and the weak theoretical basis are sufficient reasons to avoid using this method in the secondary data analysis (Andridge and Little, 2010).

Maximum Likelihood (ML) methods are similar to MIs but can be quite complex. When the researcher has a complete set of data, the means and linear regression coefficients are obtained. These parameters are used to maximise the likelihood of the observed data (Pigott, 2001). In other words, ML procedures are based on the likelihood function which defines the connection between the probability of the observed data and different values of the parameter estimates (Newman, 2014). However, when there are missing values, it can be difficult to obtain the maximising parameters (e.g. mean). The ML approach is based on Expectation-Maximization (EM), which is described elsewhere (Newman, 2014, Pigott, 2001).

Appendix G. 3 - Statistical methods to analyse repeated measures

- *Analysis of Variance Methods*

The analysis of variance (ANOVA) methods with its subtypes is one of the most commonly used statistical tests. The comparison is made to determine the differences between means of continuous variables. Repeated measures ANOVA (RM-ANOVA) is related to the analysis of means of a single or multiple variables that were obtained at multiple time points. Using this method, each mean can be compared and contrasted against other means in several ways (Field, 2013).

The ANOVA procedures are based on two assumptions that need to be met. The first assumption is related to sphericity or homogeneity of variance. In sphericity, which is measured using the Mauchly test, the covariances of all possible pair combinations of groups or levels of independence are equal. If this assumption is not met, any conclusion based on a statistical procedure that assumes sphericity will be false. ANOVA tests are based on the *F*-test which assumes sphericity and requires at least three measurements for the analysis. When there are only two measurements, sphericity cannot be calculated. In case it is not met, a more conservative testing using the Huynh and Feldt or Greenhouse-Geisser tests are applied to adjust the degrees of freedom. Ignoring the sphericity assumption and adjusted degrees of freedom will result in a Type (I) error and rejecting the null hypothesis (Misangyi, et al., 2006; Field, 2013). The second assumption is related to the normal distribution of variables and violation of this assumption may also result in inaccurate conclusions (Quene & van den Bergh, 2004; Field, 2013). Even though some statisticians suggest that a standard ANOVA is robust against moderate violation of non-normality and sphericity if the sample sizes were equal, it can't be generalised to include RM-ANOVAs (Oberfeld & Franke, 2013). Also, missing data are not allowed in RM-ANOVA procedures (Quene & van den Bergh, 2004).

Multivariate repeated measures ANOVA (MRM-ANOVA) is an alternative to the typical RM-ANOVA. It is used when multiple dependant variables are assessed simultaneously. MRM-ANOVA is based on testing the relationship between dependent variables using matrix algebra and cross-products. This approach does not assume sphericity between the variances and covariances in each group but assumes that the variances-covariances of groups are equal (Field, 2013).

- *Multilevel Modelling*

Multilevel modelling (MLM) has emerged as a flexible and an all-in-one statistical approach in the past few years. It is based on hierarchical linear models in which the data are nested in clusters and levels. For example, patients with hip pain can be considered the lowest level (Level 1). Those who were operated on by the same surgeon can be clustered together forming the second level (Level 2). Surgeons located in the same area form the highest level (Level 3). This is a typical three-level model with a single membership. The complexity of the model can be increased by having multiple memberships or cross-classifications such as a patient seeing a surgeon who works in two different areas (Steele, 2008).

Multilevel modelling is more advantageous over other statistical methods in comparing within- and between groups. Also, the assumptions made in other methods do not have to be met when MLM is applied. The unique characteristics of MLM can be summarised as (Misangyi, et al., 2006; Quene & van den Bergh, 2004; Field, 2013; Steele, 2014):

- It does not assume a normal distribution of the data or sphericity of variances and covariances. Also, it is not affected by the sample size when groups of an unequal number of participants are compared together.
- Missing values do not affect multilevel modelling of repeated measures data. In repeated measures MLM, the amount of data obtained from participants can vary, and they do not have to be equal.

- The residuals are allowed to be dependant and related to each other. Independence of the residuals is assumed in other statistical methods.
- It can be performed with the incomplete datasets. Therefore, multiple imputations or other methods of handling missing values are not necessarily applied beforehand.
- Estimation of within- and between-group variances can be achieved using MLMs. It can also give an insight if the scores of participants were affected by their membership in certain groups.

- *Conclusion*

The brief review above covered the advantages and disadvantages of the most suitable methods to analyse data with repeated measures. Repeated measures ANOVA is the easiest and most commonly used method of the three, but it is based on certain assumptions. Several solutions were suggested to handle non-normality such as log transformation, which is based on taking the logarithm or the square of the observed values. Using log transformation is still debatable among researchers because changing the values of data may change the way they are interpreted. Another solution would be using parametric tests with trimmed means, which hasn't shown sufficient control of Type (I) error. The third solution is using a statistical software to analyse data directly. However, how reliable the outcome would be with a small sample size is not clear yet (Field, 2013; Oberfeld & Franke, 2013). Multivariate repeated measures ANOVA (MRM-ANOVA) can be used instead of repeated measures ANOVA. The *F*-test in this subtype of ANOVAs is not affected by sphericity and will typically produce stronger results than univariate ANOVAs.

Multilevel modelling is the best option given the fact that normality of the hip arthroscopy dataset was not met. MLM is a comparable replacement to multivariate repeated

measures and is usually preferred over various ANOVA procedures when the conditions mentioned in Appendix G. 4 are present (Misangyi, et al., 2006).

Appendix G. 4 - Methods Used For Repeated-Measures Analysis in Multilevel Research (Misangyi* et al.*, 2006).

	Relevant Questions/Assumptions and Recommendations			
		Method to Use With		
Multilevel Design	Sphericity Assumption Met?	Balanced Data	Missing Data	Alternative Assumptions of Variance/Covariance
Designs with one within-groups factor	Yes	ANOVA, MANOVA, RMR, MLM	MLM	MLM
	No	ANOVA, MANOVA, MLM	MLM	MLM
	No	ANOVA, MANOVA, MLM		
		Balanced Data?	Method to Use When Estimating σ_B^2 and σ_W^2	Method to Use When Testing the Hypothesis $\sigma_B^2 = 0$
Designs with group membership as the only between-groups factor		Yes	ANOVA, MLM	ANOVA
		No	MLM	ANOVA
		Balanced Data?	Method to Use	
Designs with one continuous group-level independent variable		Yes	RMR, MLM	
		No	MLM	
		Assumption Regarding the Slope of the Relationship	Balanced Data?	Method to Use
Designs with one continuous individual-level independent variable		Same across groups	Yes	RMR, MLM
			No	MLM
		Varies across groups	Yes	MLM
			No	MLM

Note: ANOVA = repeated-measures analysis of variance; MANOVA = multivariate repeated-measures analysis of variance; RMR = repeated-measures regression; MLM = multilevel modeling.

References

1. Abdel-aziem, A. A., Mosaad, D. M. and Ragaa Abdelraouf, O. R. (2013) Influence of body position on concentric peak torque of hip abductor and adductor muscles. **J Nov Physiother**, 3(6).

2. Abrams, D., Davidson, M., Harrick, J., Harcourt, P., Zylinski, M. and Clancy, J. (2006) Monitoring the change: Current trends in outcome measure usage in physiotherapy. **Manual Therapy**, 11(1) 2006/02/01/, pp.46-53.

3. Agre, J. C., Magness, J. L., Hull, S. Z., Wright, K. C., Baxter, T. L., Patterson, R. and Stradel, L. (1987) Strength testing with a portable dynamometer: Reliability for upper and lower extremities. **Arch Phys Med Rehabil**, 68(7) Jul, pp.454-458.

4. Ahtiainen, J. P., Pakarinen, A., Alen, M., Kraemer, W. J. and Häkkinen, K. (2003) Muscle hypertrophy, hormonal adaptations and strength development during strength training in strength-trained and untrained men. **Eur J Appl Physiol**, 89(6) 2003/08/01, pp.555-563.

5. Alnahdi, A. H., Zeni, J. A. and Snyder-Mackler, L. (2014) Hip abductor strength reliability and association with physical function after unilateral total knee arthroplasty: A cross-sectional study. **Physical Therapy**, 94(8), pp.1154-1162.

6. Amenabar, T. and O'Donnell, J. (2013) Return to sport in australian football league footballers after hip arthroscopy and midterm outcome. **Arthroscopy**, 29(7) Jul, pp.1188-1194.

7. Amundsen, L. R., Bohannon, R. W., Delitto, A., Lynch, L., Mathiowetz, V. and Wilk, K. (1990) **Muscle strength testing: Instrumented and non-instrumented systems**.

8. Anderson, D. E. and Madigan, M. L. (2014) Healthy older adults have insufficient hip range of motion and plantar flexor strength to walk like healthy young adults. **J Biomech**, 47(5) Mar 21, pp.1104-1109.

9. Anderson, K., Strickland, S. M. and Warren, R. (2001) Hip and groin injuries in athletes. **The American Journal of Sports Medicine**, 29(4) July 1, 2001, pp.521-533.

10. Andrews, A. W., Thomas, M. W. and Bohannon, R. W. (1996) Normative values for isometric muscle force measurements obtained with hand-held dynamometers. **Physical Therapy**, 76(3) March 1, 1996, pp.248-259.

11. Aprato, A., Jayasekera, N. and Villar, R. N. (2012) Does the modified harris hip score reflect patient satisfaction after hip arthroscopy? **American Journal of Sports Medicine**, 40(11) 11//, p.2557.

12. Arnold, C. M., Warkentin, K. D., Chilibeck, P. D. and Magnus, C. R. (2010) The reliability and validity of handheld dynamometry for the measurement of lower-extremity muscle strength in older adults. **J Strength Cond Res**, 24(3) Mar, pp.815-824.

13. Atkinson, G. and Nevill, A. M. (1998) Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. **Sports Med**, 26(4) Oct, pp.217-238.

14. Aufsesser, P., Horvat, M. and Croce, R. (1996) A critical examination of selected hand-held dynamometers to assess isometric muscle strength. **Adapted Physical Activity Quarterly**, 13(2) 1996/04/01, pp.153-165.

15. Baldon, R. D., Nakagawa, T. H., Muniz, T. B., Amorim, C. F., Maciel, C. D. and Serrao, F. V. (2009) Eccentric hip muscle function in females with and without patellofemoral pain syndrome. **Journal of Athletic Training**, 44(5) Sep-Oct, pp.490-496.

16. Belhaj, K., Meftah, S., Mahir, L., Lmidmani, F. and Elfatimi, A. (2016) Isokinetic imbalance of adductor-abductor hip muscles in professional soccer players with chronic adductor-related groin pain. **Eur J Sport Sci**, 16(8) Nov, pp.1226-1231.

17. Beshay, N., Lam, P. H. and Murrell, G. A. C. (2011) Assessing the reliability of shoulder strength measurement: Hand-held versus fixed dynamometry. **Shoulder & Elbow**, 3(4), pp.244-251.

18. Biggs, K. (2010) **Locked hip** [Internet]. Available from: <http://en.allexperts.com/q/Chiropractors-965/2010/11/Locked-hip.htm> [Accessed 29th September, 2014].

19. Bijur, P. E., Silver, W. and Gallagher, E. J. (2001) Reliability of the visual analog scale for measurement of acute pain. **Academic Emergency Medicine**, 8(12), pp.1153-1157.

20. Biodex (2015). **Biodex multi-joint system: Clinical resource manual**. **Isokinetic Source Book**, Biodex Medical Systems.

21. Bittencourt, N. F. N., Santos, T. R. T., Gonçalves, G. G. P., Coelho, A. P., Gomes, B. G. B. d. M., Mendonça, L. D. M. and Fonseca, S. T. (2016) Reference values of hip abductor torque among youth athletes: Influence of age, sex and sports. **Physical Therapy in Sport**, 219//, pp.1-6.

22. Bizzini, M., Notzli, H. P. and Maffiuletti, N. A. (2007) Femoroacetabular impingement in professional ice hockey players: A case series of 5 athletes after open surgical decompression of the hip. **The American Journal of Sports Medicine**, 35(11) November 1, 2007, pp.1955-1959.

23. Blacklock, R. E., Rhodes, R. E. and Brown, S. G. (2007) Relationship between regular walking, physical activity, and health-related quality of life. **J Phys Act Health**, 4(2) Apr, pp.138-152.

24. Bland, J. M. and Altman, D. G. (2003) Applying the right statistics: Analyses of measurement studies. **Ultrasound Obstet Gynecol**, 22(1) Jul, pp.85-93.

25. BlueCross. (2016) **Outerbridge classification** [Internet]. Available from: <https://www.bcbst.com/mpmanual/Outerbridge_Classification.htm> [Accessed 16 August, 2016].

26. Bogunovic, L., Gottlieb, M., Pashos, G., Baca, G. and Clohisy, J. C. (2013) Why do hip arthroscopy procedures fail? **Clinical Orthopaedics and Related Research**, 471(8) 05/01, pp.2523-2529.

27. Bohannon, R. W. (1986) Manual muscle test scores and dynamometer test scores of knee extension strength. **Arch Phys Med Rehabil**, 67(6) Jun, pp.390-392.

28. Bohannon, R. W. (1997a) Discriminant construct validity of hand-held dynamometry and manual muscle testing in a home care setting. **Journal of Physical Therapy Science**, 9(2), pp.57-61.

29. Bohannon, R. W. (1997b) Internal consistency of manual muscle testing scores. **Perceptual and Motor Skills**, 85(2) 1997/10/01, pp.736-738.

30. Bohannon, R. W. (1997c) Reference values for extremity muscle strength obtained by hand-held dynamometry from adults aged 20 to 79 years. **Arch Phys Med Rehabil**, 78(1), pp.26-32.

31. Bohannon, R. W. (2002) Quantitative testing of muscle strength: Issues and practical options for the geriatric population. **Topics in Geriatric Rehabilitation**, 18(2), pp.1-17.

32. Bohannon, R. W. (2005) Manual muscle testing: Does it meet the standards of an adequate screening test? **Clin Rehabil**, 19(6) Sep, pp.662-667.

33. Bohannon, R. W. (2006) Hand-held dynamometry: Adoption 1900-2005. **Perceptual & Motor Skills**, 103(1), pp.3-4.

34. Bohannon, R. W. (2012) Hand-held dynamometry: A practicable alternative for obtaining objective measures of muscle strength. **Isokinetics and Exercise Science**, 20(4), pp.301-315.

35. Bohannon, R. W. and Andrews, A. W. (1987) Interrater reliability of hand-held dynamometry. **Phys Ther**, 67(6) Jun, pp.931-933.

36. Bohannon, R. W., Bubela, D. J., Wang, Y.-C., Magasi, S. R. and Gershon, R. C. (2011) Adequacy of belt-stabilized testing of knee extension strength. **Journal Of Strength And Conditioning Research / National Strength & Conditioning Association**, 25(7), pp.1963-1967.

37. Brinckmann, P., Hoefert, H. and Jongen, H. T. (1981) Sex differences in the skeletal geometry of the human pelvis and hip joint. **Journal of Biomechanics**, 14(6) 1981/01/01, pp.427-430.

38. Brown, M., Sinacore, D. R. and Host, H. H. (1995) The relationship of strength to function in the older adult. **J Gerontol A Biol Sci Med Sci**, 50 Spec NoNov, pp.55-59.

39. Brunner, A., Horisberger, M. and Herzog, R. F. (2009) Sports and recreation activity of patients with femoroacetabular impingement before and after arthroscopic osteoplasty. **The American Journal of Sports Medicine**, 37(5) May 1, 2009, pp.917-922.

40. Burkett, L. N. (1970) Causative factors in hamstring strains. **Med Sci Sports**, 2(1) Spring, pp.39-42.

41. Burman, M. S. (2001) Arthroscopy or the direct visualization of joints: An experimental cadaver study. 1931. **Clin Orthop Relat Res**,(390) Sep, pp.5-9.

42. Burnett, C. N., Betts, E. F. and King, W. M. (1990) Reliability of isokinetic measurements of hip muscle torque in young boys. **Physical Therapy**, 70(4) April 1, 1990, pp.244-249.

43. Byrd, J. W. (2010) Femoroacetabular impingement in athletes, part i: Cause and assessment. **Sports Health**, 2(4) Jul, pp.321-333.

44. Byrd, J. W. and Jones, K. S. (2001) Hip arthroscopy in athletes. **Clin Sports Med**, 20(4) Oct, pp.749-761.

45. Byrd, J. W. and Jones, K. S. (2015) Hip arthroscopy in high-level baseball players. **Arthroscopy**, 31(8) Aug, pp.1507-1510.

46. Byrd, J. W. T. (2005) Hip arthroscopy in athletes. **Operative Techniques in Sports Medicine**, 13(1) 1//, pp.24-36.

47. Byrd, J. W. T. and Jones, K. (2010) Prospective analysis of hip arthroscopy with 10-year followup. **Clinical Orthopaedics and Related Research®**, 468(3) 2010/03/01, pp.741-746.

48. Byrd, J. W. T. and Jones, K. S. (2000) Prospective analysis of hip arthroscopy with 2-year follow-up. **Arthroscopy: The Journal of Arthroscopic & Related Surgery**, 16(6) 9//, pp.578-587.

49. Byrd, J. W. T. and Jones, K. S. (2009) Hip arthroscopy in athletes: 10-year follow-up. **American Journal of Sports Medicine**, 37(11) 11//, p.2140.

50. Calmels, P. M., Nellen, M., van der Borne, I., Jourdin, P. and Minaire, P. (1997) Concentric and eccentric isokinetic assessment of flexorextensor torque ratios at the hip, knee, and ankle in a sample population of healthy subjects. **Arch Phys Med Rehabil**, 78(11), pp.1224-1230.

51. Casartelli, N. C., Maffiuletti, N. A., Item-Glatthorn, J. F., Impellizzeri, F. M. and Leunig, M. (2014) Hip muscle strength recovery after hip arthroscopy in a series of patients with symptomatic femoroacetabular impingement. **Hip International**, 24(4), pp.387-393.

52. Casartelli, N. C., Maffiuletti, N. A., Item-Glatthorn, J. F., Staehli, S., Bizzini, M., Impellizzeri, F. M. and Leunig, M. (2011) Hip muscle weakness in patients with symptomatic femoroacetabular impingement. **Osteoarthritis and Cartilage**, 19(7) 7//, pp.816-821.

53. Chahal, J., Thiel, G. S. V., Mather, R. C., Lee, S., Salata, M. J. and Nho, S. J. (2014) The minimal clinical important difference (mcid) and patient acceptable symptomatic state (pass) for the modified harris hip score and hip outcome score among patients undergoing surgical treatment for femoroacetabular impingement. **Orthopaedic Journal of Sports Medicine**, 2(2 suppl) July 1, 2014.

54. Charlton, P. C., Bryant, A. L., Kemp, J. L., Clark, R. A., Crossley, K. M. and Collins, N. J. (2016) Single-leg squat performance is impaired 1 to 2 years after hip arthroscopy. **PM&R**, 8(4) 4//, pp.321-330.

55. Christensen, C. P., Althausen, P. L., Mittleman, M. A., Lee, J. A. and McCarthy, J. C. (2003) The nonarthritic hip score: Reliable and validated. **Clin Orthop Relat Res**,(406) Jan, pp.75-83.

56. Clohisy, J. C., Kamath, G. V., Byrd, G. D., Steger-May, K. and Wright, R. W. (2008) Patient compliance with clinical follow-up after total joint arthroplasty. **J Bone Joint Surg Am**, 90(9) Sep, pp.1848-1854.

57. Coldwells, A., Atkinson, G. and Reilly, T. (1994) Sources of variation in back and leg dynamometry. **Ergonomics**, 37(1) Jan, pp.79-86.

58. Cometti, G., Maffiuletti, N. A., Pousson, M., Chatard, J. C. and Maffulli, N. (2001) Isokinetic strength and anaerobic power of elite, subelite and amateur french soccer players. **Int J Sports Med**, 22(01) //

59. 31.12.2002, pp.45-51.

60. Cuthbert, S. C. and Goodheart, G. J. (2007) On the reliability and validity of manual muscle testing: A literature review. **Chiropractic & Osteopathy**, 1503/06

61. 02/14/received

62. 03/06/accepted, pp.4-4.

63. Daivajna, S., Bajwa, A. and Villar, R. (2015) Outcome of arthroscopy in patients with advanced osteoarthritis of the hip. **PLoS ONE**, 10(1) 01/30

64. 06/06/received

65. 11/02/accepted, p.e0113970.

66. Daneshjoo, A., Rahnama, N., Mokhtar, A. H. and Yusof, A. (2013) Bilateral and unilateral asymmetries of isokinetic strength and flexibility in male young professional soccer players. **J Hum Kinet**, 36Mar, pp.45-53.

67. Dangin, A., Tardy, N., Wettstein, M., May, O. and Bonin, N. (2016) Microinstability of the hip: A review. **Orthop Traumatol Surg Res**, 102(8s) Dec, pp.S301-s309.

68. Degen, R. M., Nawabi, D. H., Fields, K. G., Wentzel, C. S., Kelly, B. T. and Coleman, S. H. (2016) Simultaneous versus staged bilateral hip arthroscopy in the treatment of femoroacetabular impingement. **Arthroscopy**, 32(7) Jul, pp.1300-1307.

69. Dellal, A., Hill-Haas, S., Lago-Penas, C. and Chamari, K. (2011) Small-sided games in soccer: Amateur vs. Professional players' physiological responses, physical, and technical activities. **The Journal of Strength & Conditioning Research**, 25(9), pp.2371-2381.

70. Delp, S. L., Hess, W. E., Hungerford, D. S. and Jones, L. C. (1999) Variation of rotation moment arms with hip flexion. **J Biomech**, 32(5) May, pp.493-501.

71. Denton, T. P., Jones, M. A. and Saunders, S. W. (2014) Test-retest reliability of isometric hip muscle strength measured using handheld dynamometry: A pilot study. **Australian Medical Student Journal**, 5(1) 01//, p.44.

72. Dienst, M. (2005) Hip arthroscopy: Technique and anatomy. **Operative Techniques in Sports Medicine**, 13(1) Jan, pp.13-23.

73. Dippmann, C., Thorborg, K., Kraemer, O., Winge, S., Palm, H. and Hölmich, P. (2014) Hip arthroscopy with labral repair for femoroacetabular impingement: Short-term outcomes. **Knee Surgery, Sports Traumatology, Arthroscopy**, 22(4) 2014/04/01, pp.744-749.

74. Domb, B. G., Dunne, K. F., Martin, T., Gui, C., Finch, N. and Stake, C. E. (2015a) Return to sports in a general hip arthroscopy cohort: Minimum two-year follow-up. **Orthopaedic Journal of Sports Medicine**, 3(7 suppl2) 07/17, p.2325967115S2325900083.

75. Domb, B. G., El Bitar, Y. F., Stake, C. E., Trenga, A. P., Jackson, T. J. and Lindner, D. (2014) Arthroscopic labral reconstruction is superior to segmental resection for irreparable labral tears in the hip: A matched-pair controlled study

with minimum 2-year follow-up. **American Journal of Sports Medicine**, 42(1) 01//, p.122.

76. Domb, B. G., Linder, D., Finley, Z., Botser, I. B., Chen, A., Williamson, J. and Gupta, A. (2015b) Outcomes of hip arthroscopy in patients aged 50 years or older compared with a matched-pair control of patients aged 30 years or younger. **Arthroscopy**, 31(2) Feb, pp.231-238.

77. Dong, Y. and Peng, C. Y. (2013) Principled missing data methods for researchers. **Springerplus**, 2(1) Dec, p.222.

78. Dorfmann, H. and Boyer, T. (1999) Arthroscopy of the hip: 12 years of experience. **Arthroscopy**, 15(1) Jan-Feb, pp.67-72.

79. Drake, R. L., Vogl, W. and Mitchell, A. (2005) **Gray's anatomy for students**. Elsevier Churchill Livingstone.

80. Drouin, J. M., Valovich-mcLeod, T. C., Shultz, S. J., Gansneder, B. M. and Perrin, D. H. (2004) Reliability and validity of the biodex system 3 pro isokinetic dynamometer velocity, torque and position measurements. **Eur J Appl Physiol**, 91(1) Jan, pp.22-29.

81. Dudda, M., Kim, Y.-J., Zhang, Y., Nevitt, M. C., Xu, L., Niu, J., Goggins, J., Doherty, M. and Felson, D. T. (2011) Morphological differences between chinese and caucasian female hips: Could they account for the ethnic difference in hip osteoarthritis? **Arthritis and rheumatism**, 63(10), pp.2992-2999.

82. Dvorak, M., Duncan, C. P. and Day, B. (1990) Arthroscopic anatomy of the hip. **Arthroscopy: The Journal of Arthroscopic & Related Surgery**, 6(4) 12//, pp.264-273.

83. Edelstein, J., Ranawat, A., Enseki, K., Yun, R. and Draovitch, P. (2012) Post-operative guidelines following hip arthroscopy. **Current Reviews in Musculoskeletal Medicine**, 5(1) 2012/03/01, pp.15-23.

84. Ejnisman, L., Philippon, M. J., Lertwanich, P., Pennock, A. T., Herzog, M. M., Briggs, K. K. and Ho, C. P. (2013) Relationship between femoral anteversion and findings in hips with femoroacetabular impingement. **Orthopedics**, 36(3) Mar, pp.e293-300.

85. Emara, K., Samir, W., Motasem el, H. and Ghafar, K. A. (2011) Conservative treatment for mild femoroacetabular impingement. **J Orthop Surg (Hong Kong)**, 19(1) Apr, pp.41-45.

86. Emery, C. A., Maitland, M. E. and Meeuwisse, W. H. (1999) Test-retest reliability of isokinetic hip adductor and flexor muscle strength. **Clinical Journal of Sport Medicine**, 9(2), pp.79-85.

87. Enseki, K. R., Martin, R. and Kelly, B. T. (2010) Rehabilitation after arthroscopic decompression for femoroacetabular impingement. **Clin Sports Med**, 29(2) Apr, pp.247-255, viii.

88. Espinosa, N., Beck, M., Rothenfluh, D. A., Ganz, R. and Leunig, M. (2007) Treatment of femoro-acetabular impingement: Preliminary results of labral refixation. Surgical technique. **J Bone Joint Surg Am**, 89 Suppl 2 Pt.1Mar, pp.36-53.

89. Espinosa, N., Rothenfluh, D. A., Beck, M., Ganz, R. and Leunig, M. (2006) Treatment of femoro-acetabular impingement: Preliminary results of labral refixation. **J Bone Joint Surg Am**, 88(5) May, pp.925-935.

90. Falla, D., Jull, G. and Hodges, P. W. (2004) Feedforward activity of the cervical flexor muscles during voluntary arm movements is delayed in chronic neck pain. **Exp Brain Res**, 157(1) Jul, pp.43-48.

91. Feiring, D., Ellenbecker, T. and Derscheid, G. (1990) Test-retest reliability of the biodex isokinetic dynamometer. **Journal of Orthopaedic & Sports Physical Therapy**, 11(7) 1990/01/01, pp.298-300.

92. Fenter, P. C., Bellew, J. W., Pitts, T. and Kay, R. (2003a) A comparison of 3 hand-held dynamometers used to measure hip abduction strength. **Journal of Strength and Conditioning Research**, 17(3) Aug, pp.531-535.

93. Fenter, P. C., Bellew, J. W., Pitts, T. A. and Kay, R. E. (2003b) Reliability of stabilised commercial dynamometers for measuring hip abduction strength: A pilot study. **British Journal of Sports Medicine**, 37(4) Aug 1, pp.331-334.

94. Fernandez, E. and Gastaldi, P. (2012) Hip pain from the orthopeadic point of view. **European Journal of Radiology**, 81(12) 12//, pp.3737-3739.

95. Field, A. (2013) **Discovering statistics using ibm spss statistics**. Fourth Edition ed., SAGE Publications Ltd.

96. Florence, J. M., Pandya, S., King, W. M., Robison, J. D., Baty, J., Miller, J. P., Schierbecker, J. and Signore, L. C. (1992) Intrarater reliability of manual muscle test (medical research council scale) grades in duchenne's muscular dystrophy. **Physical Therapy**, 72(2) February 1, 1992, pp.115-122.

97. Fredericson, M., Cookingham, C. L., Chaudhari, A. M., Dowdell, B. C., Oestreicher, N. and Sahrmann, S. A. (2000) Hip abductor weakness in distance runners with iliotibial band syndrome. **Clin J Sport Med**, 10(3) Jul, pp.169-175.

98. Freke, M. D., Kemp, J., Svege, I., Risberg, M. A., Semciw, A. and Crossley, K. M. (2016) Physical impairments in symptomatic femoroacetabular impingement: A systematic review of the evidence. **British Journal of Sports Medicine**,June 14, 2016.

99. Frese, E., Brown, M. and Norton, B. J. (1987) Clinical reliability of manual muscle testing: Middle trapezius and gluteus medius muscles. **Physical Therapy**, 67(7) July 1, 1987, pp.1072-1076.

100. Friesenbichler, B., Casartelli, N. C., Wellauer, V., Item-Glatthorn, J. F., Ferguson, S. J., Leunig, M. and Maffiuletti, N. A. (2017) Explosive and maximal strength before and 6 months after total hip arthroplasty. **Journal Of Orthopaedic Research: Official Publication Of The Orthopaedic Research Society**.

101. Fukuda, T. Y., Rossetto, F. M., Magalhaes, E., Bryk, F. F., Lucareli, P. R. and de Almeida Aparecida Carvalho, N. (2010) Short-term effects of hip abductors and lateral rotators strengthening in females with patellofemoral pain syndrome: A randomized controlled clinical trial. **J Orthop Sports Phys Ther**, 40(11) Nov, pp.736-742.

102. Fulcher, M. L., Hanna, C. M. and Elley, C. R. (2010) Reliability of handheld dynamometry in assessment of hip strength in adult male football players. **Journal of Science and Medicine in Sport**, 13(1) Jan, pp.80-84.

103. Gautrey, C. N., Watson, T. and Mitchell, A. (2013) The effect of velocity on load range during isokinetic hip abduction and adduction exercise. **Int J Sports Med**, 34(7), pp.623-630.

104. Gerodimos, V., Karatrantou, K., Paschalis, V., Zafeiridis, A., Katsareli, E., Bilios, P. and Kellis, S. (2015) Reliability of concentric and eccentric strength of hip abductor and adductor muscles in young soccer players. **Biology of Sport**, 32(4), pp.351-356.

105. Giavarina, D. (2015) Understanding bland altman analysis. **Biochemia Medica**, 25(2) 06/05

106. 02/23/received

107. 04/30/accepted, pp.141-151.

108. Goodpaster, B. H., Park, S. W., Harris, T. B., Kritchevsky, S. B., Nevitt, M., Schwartz, A. V., Simonsick, E. M., Tylavsky, F. A., Visser, M. and Newman, A. B. (2006) The loss of skeletal muscle strength, mass, and quality in older adults: The health, aging and body composition study. **The Journals of Gerontology: Series A**, 61(10), pp.1059-1064.

109. Gupta, A., Redmond, J. M., Hammarstedt, J. E., Lindner, D., Stake, C. E. and Domb, B. G. (2015) Does obesity affect outcomes after hip arthroscopy? A cohort analysis. **J Bone Joint Surg Am**, 97(1) Jan 7, pp.16-23.

110. Hall, S. J. (2003) Chapter 13: Equilibrium and human movement. In: **Basic biomechanics**. McGraw Hill,

111. Hanneman, S. K. (2008) Design, analysis and interpretation of method-comparison studies. **AACN advanced critical care**, 19(2) Apr–Jun, pp.223-234.

112. Harikesavan, K., Chakravarty, R. D., Maiya, A. G., Hegde, S. P. and Y. Shivanna, S. (2017) Hip abductor strengthening improves physical function following total knee replacement: One-year follow-up of a randomized pilot study. **The Open Rheumatology Journal**, 1103/31

113. 09/09/received

114. 12/07/revised

115. 02/02/accepted, pp.30-42.

116. Hart, J. M., Pietrosimone, B., Hertel, J. and Ingersoll, C. D. (2010) Quadriceps activation following knee injuries: A systematic review. **Journal of Athletic Training**, 45(1) Jan-Feb, pp.87-97.

117. Hawker, G. A., Mian, S., Kendzerska, T. and French, M. (2011a) Measures of adult pain: Visual analog scale for pain (vas pain), numeric rating scale for pain (nrs pain), mcgill pain questionnaire (mpq), short-form mcgill pain questionnaire (sf-mpq), chronic pain grade scale (cpgs), short form-36 bodily pain scale (sf-36 bps), and measure of intermittent and constant osteoarthritis pain (icoap). **Arthritis Care Res (Hoboken)**, 63 Suppl 11Nov, pp.S240-252.

118. Hawker, G. A., Mian, S., Kendzerska, T. and French, M. (2011b) Measures of adult pain: Visual analog scale for pain (vas pain), numeric rating scale for pain (nrs pain), mcgill pain questionnaire (mpq), short-form mcgill pain questionnaire (sf-mpq), chronic pain grade scale (cpgs), short form-36 bodily pain scale (sf-36 bps), and measure of intermittent and constant osteoarthritis pain (icoap). **Arthritis Care & Research**, 63(S11), pp.S240-S252.

119. Heck, R. H., Thomas, S. L. and Tabata, L. N. (2014) **Multilevel and longitudinal modeling with ibm spss**. Quantitative methodology series.Second edition. ed.

120. Henderson, E. R., Marulanda, G. A., Cheong, D., Temple, H. T. and Letson, G. D. (2011) Hip abductor moment arm--a mathematical analysis for proximal femoral replacement. **J Orthop Surg Res**, 6Jan 25, p.6.

121. Hickey, M. S., Costill, D. L., McConell, G. K., Widrick, J. J. and Tanaka, H. (1992) Day to day variation in time trial cycling performance. **Int J Sports Med**, 13(6) Aug, pp.467-470.

122. Hinman, R. S., Dobson, F., Takla, A., O'Donnell, J. and Bennell, K. L. (2014) Which is the most useful patient-reported outcome in femoroacetabular impingement? Test–retest reliability of six questionnaires. **British Journal of Sports Medicine**, 48(6) March 1, 2014, pp.458-463.

123. Hislop, H., Avers, D. and Brown, M. (2013) **Daniels and worthingham's muscle testing: Techniques of manual examination and performance testing, edition 9**. 9 ed., Elsevier Health Sciences.

124. Hislop, H. and Montgomery, J. (2007) **Daniels and worthingham's muscle testing: Techniques of manual examination**. 8th edition ed., Saunders.

125. Hodges, P. W. and Richardson, C. A. (1996) Inefficient muscular stabilization of the lumbar spine associated with low back pain. A motor control evaluation of transversus abdominis. **Spine (Phila Pa 1976)**, 21(22) Nov 15, pp.2640-2650.

126. Holm, B., Thorborg, K., Husted, H., Kehlet, H. and Bandholm, T. (2013) Surgery-induced changes and early recovery of hip-muscle strength, leg-press power, and functional performance after fast-track total hip arthroplasty: A prospective cohort study. **PLoS ONE**, 8(4) 04/16

127. 10/12/received

128. 03/17/accepted, p.e62109.

129. Holmich, P., Holmich, L. R. and Bjerg, A. M. (2004) Clinical examination of athletes with groin pain: An intraobserver and interobserver reliability study. **Br J Sports Med**, 38(4) Aug, pp.446-451.

130. Hopkins, W. G. (2000) Measures of reliability in sports medicine and science. **Sports Med**, 30(1) Jul, pp.1-15.

131. Houck, P. R., Mazumdar, S., Koru-Sengul, T., Tang, G., Mulsant, B. H., Pollock, B. G. and Reynolds Iii, C. F. (2004) Estimating treatment effects from longitudinal clinical trial data with missing values: Comparative analyses using different methods. **Psychiatry Research**, 129(2) 12/15/, pp.209-215.

132. Hu, H., Meijer, O. G., van Dieën, J. H., Hodges, P. W., Bruijn, S. M., Strijers, R. L., Nanayakkara, P. W. B., van Royen, B. J., Wu, W. H. and Xia, C. (2011) Is the psoas a hip flexor in the active straight leg raise? **European Spine Journal**, 20(5) 07/13

133. 12/23/received

134. 05/17/revised

135. 06/27/accepted, pp.759-765.

136. Iddings, D. M., Smith, L. K. and Spencer, W. A. (1961) Muscle testing. 2. Reliability in clinical use. **Phys Ther Rev**, 41Apr, pp.249-256.

137. ISHA. (2017) **Hip arthroscopy outcome instruments** [Internet]. Available from: <http://www.isha.net/site-pdf/hip-arthroscopy-outcome-instruments.pdf> [Accessed 30 January, 2017].

138. Jackson, S. M., Cheng, M. S., Smith, A. R., Jr. and Kolber, M. J. (2016) Intrarater reliability of hand held dynamometry in measuring lower extremity isometric strength using a portable stabilization device. **Man Ther**,Jul 21.

139. Jacobs, C., Uhl, T. L., Seeley, M., Sterling, W. and Goodrich, L. (2005) Strength and fatigability of the dominant and nondominant hip abductors. **Journal of Athletic Training (National Athletic Trainers' Association)**, 40(3), pp.203-206.

140. Janda, V. (2014) **The janda approach** [Internet]. Available from: <http://www.muscleimbalancesyndromes.com> [Accessed 19 June, 2016].

141. Jensen, J., Hölmich, P., Bandholm, T., Zebis, M. K., Andersen, L. L. and Thorborg, K. (2014) Eccentric strengthening effect of hip-adductor training with elastic bands in soccer players: A randomised controlled trial. **British Journal of Sports Medicine**, 48(4) February 1, 2014, pp.332-338.

142. Jepsen, J., Laursen, L., Larsen, A. and Hagert, C. G. (2004) Manual strength testing in 14 upper limb muscles: A study of inter-rater reliability. **Acta Orthop Scand**, 75(4) Aug, pp.442-448.

143. Jung, M.-C. and Hallbeck, M. S. (2004) Quantification of the effects of instruction type, verbal encouragement, and visual feedback on static and peak handgrip strength. **International Journal of Industrial Ergonomics**, 34(5) 2004/11/01/, pp.367-374.

144. Kamath, A. F., Componovo, R., Baldwin, K., Israelite, C. L. and Nelson, C. L. (2009) Hip arthroscopy for labral tears: Review of clinical outcomes with 4.8-year mean follow-up. **The American Journal of Sports Medicine**, 37(9) September 1, 2009, pp.1721-1727.

145. Katoh, M. (2015) Test-retest reliability of isometric shoulder muscle strength measurement with a handheld dynamometer and belt. **J Phys Ther Sci**, 27(6) Jun, pp.1719-1722.

146. Katoh, M., Hiiragi, Y. and Uchida, M. (2011) Validity of isometric muscle strength measurements of the lower limbs using a handheld dynamometer and belt: A comparison with an isokinetic dynamometer. **Journal of Physical Therapy Science**, 23(4), pp.553-557.

147. Katoh, M. and Yamasaki, H. (2009) Comparison of reliability of isometric leg muscle strength measurements made using a hand-held dynamometer with and without a restraining belt. **Journal of Physical Therapy Science**, 21(1), pp.37-42.

148. Kawaguchi, J. K. and Babcock, G. (2010) Validity and reliability of handheld dynametric strength assessment of hip extensor and abductor muscles. **Athletic Training & Sports Health Care: The Journal for the Practicing Clinician**, 2(1), pp.11-17.

149. Keating, J. L. and Matyas, T. A. (1996) The influence of subject and test design on dynamometric measurements of extremity muscles. **Physical Therapy**, 76(8) August 1, 1996, pp.866-889.

150. Kelln, B. M., McKeon, P. O., Gontkof, L. M. and Hertel, J. (2008) Hand-held dynamometry: Reliability of lower extremity muscle testing in healthy, physically active, young adults. **Journal of Sport Rehabilitation**, 17(2) May, pp.160-170.

151. Kemp, J. L., Collins, N. J., Makdissi, M., Schache, A. G., Machotka, Z. and Crossley, K. (2012) Hip arthroscopy for intra-articular pathology: A systematic review of outcomes with and without femoral osteoplasty. **British Journal of Sports Medicine**, 46(9) July 1, 2012, pp.632-643.

152. Kemp, J. L., Collins, N. J., Roos, E. M. and Crossley, K. M. (2013) Psychometric properties of patient-reported outcome measures for hip arthroscopic surgery. **American Journal of Sports Medicine**, 41(9) 09//, p.2065.

153. Kemp, J. L., Makdissi, M., Schache, A. G., Finch, C. F., Pritchard, M. G. and Crossley, K. M. (2015) Is quality of life following hip arthroscopy in patients with chondrolabral pathology associated with impairments in hip strength or range of motion? **Knee Surg Sports Traumatol Arthrosc**,Jul 3.

154. Kemp, J. L., Makdissi, M., Schache, A. G., Pritchard, M. G., Pollard, T. C. and Crossley, K. M. (2014a) Hip chondropathy at arthroscopy: Prevalence and relationship to labral pathology, femoroacetabular impingement and patient-reported outcomes. **Br J Sports Med**, 48(14) Jul, pp.1102-1107.

155. Kemp, J. L., Schache, A. G., Makdissia, M., Pritchard, M. G., Sims, K. and Crossley, K. M. (2014b) Is hip range of motion and strength impaired in people with hip chondrolabral pathology? **J Musculoskelet Neuronal Interact**, 14(3) Sep, pp.334-342.

156. Kendall, F. P. (1993) **Muscles, testing and function**. Lippincott Williams & Wilkins.

157. Kennedy, M., Lamontagne, M. and Beaulé, P. (2009) The effect of cam femoroacetabular impingement on hip maximal dynamic range of motion. **Journal of Orthopedics**, 1(1), pp.41-50.

158. Khayambashi, K., Mohammadkhani, Z., Ghaznavi, K., Lyle, M. A. and Powers, C. M. (2012) The effects of isolated hip abductor and external rotator muscle strengthening on pain, health status, and hip strength in females with patellofemoral pain: A randomized controlled trial. **J Orthop Sports Phys Ther**, 42(1) Jan, pp.22-29.

159. Kierkegaard, S., Langeskov-Christensen, M., Lund, B., Naal, F. D., Mechlenburg, I., Dalgas, U. and Casartelli, N. C. (2017a) Pain, activities of daily living and sport function at different time points after hip arthroscopy in

patients with femoroacetabular impingement: A systematic review with meta-analysis. **British Journal of Sports Medicine**, 51(7), p.572.

160. Kierkegaard, S., Mechlenburg, I., Lund, B., Søballe, K. and Dalgas, U. (2017b) Impaired hip muscle strength in patients with femoroacetabular impingement syndrome. **Journal of Science and Medicine in Sport**,2017/05/25/.

161. Kim, H.-Y. (2013) Statistical notes for clinical researchers: Assessing normal distribution (2) using skewness and kurtosis. **Restorative Dentistry & Endodontics**, 38(1) 02/26, pp.52-54.

162. Kim, W. K., Kim, D.-K., Seo, K. M. and Kang, S. H. (2014) Reliability and validity of isometric knee extensor strength test with hand-held dynamometer depending on its fixation: A pilot study. **Annals of Rehabilitation Medicine**, 38(1) 02/25

163. 05/16/received

164. 10/01/accepted, pp.84-93.

165. King, E., Ward, J., Small, L., Falvey, E. and Franklyn-Miller, A. (2015) Athletic groin pain: A systematic review and meta-analysis of surgical versus physical therapy rehabilitation outcomes. **British Journal of Sports Medicine**,June 30, 2015.

166. Kokmeyer, D. J., van der Wurff, P., Aufdemkampe, G. and Fickenscher, T. C. M. (2002) The reliability of multitest regimens with sacroiliac pain provocation tests. **Journal of Manipulative and Physiological Therapeutics**, 25(1) 1//, pp.42-48.

167. Kolber, M. J. and Cleland, J. A. (2005) Strength testing using hand-held dynamometry. **Physical Therapy Reviews**, 10(2), pp.99-112.

168. Kollock Jr, R. O., Van Lunen, B., Linza, J. L. and Onate, J. A. (2013) Comparison of isometric portable fixed dynamometry to isokinetic dynamometry for assessment of hip strength. **International Journal of Athletic Therapy & Training**, 18(6), pp.1-6.

169. Kowalczuk, M., Bhandari, M., Farrokhyar, F., Wong, I., Chahal, M., Neely, S., Gandhi, R. and Ayeni, O. (2013) Complications following hip arthroscopy: A systematic review and meta-analysis. **Knee Surgery, Sports Traumatology, Arthroscopy**, 21(7), pp.1669-1675.

170. Kramer, J. F., Vaz, M. D. and Vandervoort, A. A. (1991) Reliability of isometric hip abductor torques during examiner- and belt-resisted tests. **Journal of Gerontology**, 46(2) March 1, 1991, pp.M47-M51.

171. Krause, D. A., Neuger, M. D., Lambert, K. A., Johnson, A. E., DeVinny, H. A. and Hollman, J. H. (2014) Effects of examiner strength on reliability of

hip-strength testing using a handheld dynamometer. **J Sport Rehabil**, 23(1) Feb, pp.56-64.

172. Krause, D. A., Schlagel, S. J., Stember, B. M., Zoetewey, J. E. and Hollman, J. H. (2007) Influence of lever arm and stabilization on measures of hip abduction and adduction torque obtained by hand-held dynamometry. **Arch Phys Med Rehabil**, 88(1), pp.37-42.

173. Kuhns, B. D., Weber, A. E., Levy, D. M. and Wuerz, T. H. (2015) The natural history of femoroacetabular impingement. **Front Surg**, 2, p.58.

174. Kwon, Y. J. and Lee, H. O. (2013) How different knee flexion angles influence the hip extensor in the prone position. **J Phys Ther Sci**, 25(10) Oct, pp.1295-1297.

175. Lakens, D. (2013) Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and anovas. **Front Psychol**, 4, p.863.

176. Lalkhen, A. G. and McCluskey, A. (2008) Clinical tests: Sensitivity and specificity. **Continuing Education in Anaesthesia, Critical Care & Pain**, 8(6) December 1, 2008, pp.221-223.

177. Lamontagne, M., Brisson, N., Kennedy, M. J. and Beaulé, P. E. (2011) Preoperative and postoperative lower-extremity joint and pelvic kinematics during maximal squatting of patients with cam femoro-acetabular impingement. **J Bone Joint Surg Am**, 93(Supplement 2), pp.40-45.

178. Larson, C. M. and Giveans, M. R. (2008) Arthroscopic management of femoroacetabular impingement: Early outcomes measures. **Arthroscopy: The Journal of Arthroscopic & Related Surgery**, 24(5) 5//, pp.540-546.

179. Leard, J. S., Crane, B. A. and Ball, K. A. (2009) Intrarater and interrater reliability of 22 clinical measures associated with lower quarter malalignment. **J Manipulative Physiol Ther**, 32(4) May, pp.270-276.

180. Lee, S. W. and Kim, S. Y. (2015) Effects of hip exercises for chronic low-back pain patients with lumbar instability. **J Phys Ther Sci**, 27(2) Feb, pp.345-348.

181. Levangie, P. K., Norkin, C. C. and Levangie, P. K. (2011) **Joint structure and function : A comprehensive analysis**. 5th ed. Philadelphia, F.A. Davis Co.

182. Levine, T. R. and Hullett, C. R. (2002) Eta squared, partial eta squared, and misreporting of effect size in communication research. **Human Communication Research**, 28(4) Oct 1, pp.612-625.

183. Levy, D. M., Kuhns, B. D., Chahal, J., Philippon, M. J., Kelly, B. T. and Nho, S. J. (2016) Hip arthroscopy outcomes with respect to patient acceptable symptomatic state (pass) and minimal clinically important difference (mcid). **Arthroscopy: The Journal of Arthroscopic & Related Surgery**.

184. Light, N. and Thorborg, K. (2016) The precision and torque production of common hip adductor squeeze tests used in elite football. **J Sci Med Sport**, 19(11) Nov, pp.888-892.

185. Lilienfeld, A. M., Jacobs, M. and Willis, M. (1954) A study of the reproducibility of muscle testing and certain other aspects of muscle scoring. **Phys Ther Rev**, 34(6) Jun, pp.279-289.

186. Lincoln, N. B., Nicholl, C. R., Flannaghan, T., Leonard, M. and Van der Gucht, E. (2003) The validity of questionnaire measures for assessing depression after stroke. **Clin Rehabil**, 17(8) Dec, pp.840-846.

187. Lindner, D., Stake, C. E., Jackson, T. J., Bitar, Y. E., Chen, A. and Domb, B. G. (2013) Two year follow-up of hip arthroscopies: A match-controlled study comparing patients over 50 years to under 30 years. **Orthopaedic Journal of Sports Medicine**, 1(4 Suppl) 09/20, p.2325967113S2325900048.

188. Lodhia, P., Martin, T., Gui, C., Stake, C. E., Vemula, S. P., Suarez-Ahedo, C., Chandrasekaran, S. and Domb, B. G. (2015) Outcomes of 1038 hip arthroscopies: A two-year follow-up study. **Orthopaedic Journal of Sports Medicine**, 3(7 suppl2) 07/17, p.2325967115S2325900135.

189. Loeser, R. F. (2010) Age-related changes in the musculoskeletal system and the development of osteoarthritis. **Clinics in geriatric medicine**, 26(3), pp.371-386.

190. Londers, J. and Van Melkebeek, J. (2007) Hip arthroscopy: Outcome and patient satisfaction after 5 to 10 years. **Acta Orthop Belg**, 73(4) Aug, pp.478-483.

191. Lourencin, F., Macedo, O., Scarpellini, E. and Greve, J. (2012) Evaluation of hip adductor and abductor muscles using an isokinetic dynamometer. **Acta Fisiátr**, 19(1), pp.16-20.

192. Lovett, R. W. and Martin, E. G. (1916) Certain aspects of infantile paralysis: With a description of a method of muscle testing. **Journal of the American Medical Association**, LXVI(10), pp.729-733.

193. Lu, Y.-M., Lin, J.-H., Hsiao, S.-F., Liu, M.-F., Chen, S.-M. and Lue, Y.-J. (2011) The relative and absolute reliability of leg muscle strength testing by a handheld dynamometer. **Journal Of Strength And Conditioning Research / National Strength & Conditioning Association**, 25(4), pp.1065-1071.

194. Lund, H., Sondergaard, K., Zachariassen, T., Christensen, R., Bulow, P., Henriksen, M., Bartels, E. M., Danneskiold-Samsoe, B. and Bliddal, H. (2005) Learning effect of isokinetic measurements in healthy subjects, and reliability and comparability of biodex and lido dynamometers. **Clin Physiol Funct Imaging**, 25(2) Mar, pp.75-82.

195. Lunnen, J. D., Yack, J. and LeVeau, B. F. (1981) Relationship between muscle length, muscle activity, and torque of the hamstring muscles. **Phys Ther**, 61(2) Feb, pp.190-195.

196. Lynch, T. S., Terry, M. A., Bedi, A. and Kelly, B. T. (2013) Hip arthroscopic surgery: Patient evaluation, current indications, and outcomes. **American Journal of Sports Medicine**, 41(5) 05//, p.1174.

197. Madsen, M. S., Ritter, M. A., Morris, H. H., Meding, J. B., Berend, M. E., Faris, P. M. and Vardaxis, V. G. (2004) The effect of total hip arthroplasty surgical approach on gait. **Journal of Orthopaedic Research**, 22(1), pp.44-50.

198. Magee, D. J. (2008) **Orthopedic physical assessment (5th ed.)**. Fifth Edition ed., Saunders Elsevier.

199. Magnusson, S. P., Gleim, G. W. and Nicholas, J. A. (1990) Subject variability of shoulder abduction strength testing. **Am J Sports Med**, 18(4) Jul-Aug, pp.349-353.

200. Malliaras, P., Hogan, A., Nawrocki, A., Crossley, K. and Schache, A. (2009) Hip flexibility and strength measures: Reliability and association with athletic groin pain. **British Journal of Sports Medicine**, 43(10), pp.739-744.

201. Marani, E. and Koch, W. F. R. M. (2014). **The pelvis structure, gender and society**, SpringerLINK ebooks - Biomedical and Life Sciences (2014): p. 1 online resource.

202. Mardones, R. M., Nemtala, F. and Tomic, A. (2010) Arthroscopic treatment of femoroacetabular impingement in patients over 60 years old: Preliminary report of a pilot study. **Cartilage**, 1(3), pp.188-193.

203. Martin, E. G. and Lovett, R. W. (1915) A method of testing muscular strength in infantile paralysis. **Journal of the American Medical Association**, LXV(18), pp.1512-1513.

204. Martin, H. J., Yule, V., Syddall, H. E., Dennison, E. M., Cooper, C. and Aihie Sayer, A. (2006a) Is hand-held dynamometry useful for the measurement of quadriceps strength in older people? A comparison with the gold standard bodex dynamometry. **Gerontology**, 52(3), pp.154-159.

205. Martin, R. L. (2005) Hip arthroscopy and outcome assessment. **Operative Techniques in Orthopaedics**, 15(3) 7//, pp.290-296.

206. Martin, R. L., Kelly, B. T. and Philippon, M. J. (2006b) Evidence of validity for the hip outcome score. **Arthroscopy: The Journal of Arthroscopic & Related Surgery**, 22(12) 12//, pp.1304-1311.

207. Martini, F. H. and Nath, J. L. (2008) **Fundamentals of anatomy & physiology**. 8th Ed ed., Pearson Benjamin Cummings.

208. Masonis, J. L. and Bourne, R. B. (2002) Surgical approach, abductor function, and total hip arthroplasty dislocation. **Clinical Orthopaedics and Related Research**, 405, pp.46-53.

209. Mayne, E., Memarzadeh, A., Raut, P., Arora, A. and Khanduja, V. (2017) Measuring hip muscle strength in patients with femoroacetabular impingement and other hip pathologies: A systematic review. **Bone & Joint Research**, 6(1) 02/10, pp.66-72.

210. McCarthy, J., Jarrett, B., Ojeifo, O., Lee, J. and Bragdon, C. (2011) What factors influence long-term survivorship after hip arthroscopy? **Clinical Orthopaedics and Related Research®**, 469(2) 2011/02/01, pp.362-371.

211. McCarthy, J. and McMillan, S. (2013) Arthroscopy of the hip: Factors affecting outcome. **Orthop Clin North Am**, 44(4) Oct, pp.489-498.

212. McCormack, H. M., Horne, D. J. and Sheather, S. (1988) Clinical applications of visual analogue scales: A critical review. **Psychol Med**, 18(4) Nov, pp.1007-1019.

213. McCormick, F., Nwachukwu, B. U., Alpaugh, K. and Martin, S. D. (2012) Predictors of hip arthroscopy outcomes for labral tears at minimum 2-year follow-up: The influence of age and arthritis. **Arthroscopy**, 28(10) Oct, pp.1359-1364.

214. Meftah, M., Rodriguez, J. A., Panagopoulos, G. and Alexiades, M. M. (2011) Long-term results of arthroscopic labral debridement: Predictors of outcomes. **Orthopedics**, 34(10) Oct, pp.e588-592.

215. Mendis, M. D., Wilson, S. J., Hayes, D. A., Watts, M. C. and Hides, J. A. (2014) Hip flexor muscle size, strength and recruitment pattern in patients with acetabular labral tears compared to healthy controls. **Manual Therapy**, 19(5) 2014/10/01/, pp.405-410.

216. Methenitis, S., Karandreas, N., Spengos, K., Zaras, N., Stasinaki, A. N. and Terzis, G. (2016) Muscle fiber conduction velocity, muscle fiber composition, and power performance. **Med Sci Sports Exerc**,Apr 29.

217. Misangyi, V. F., LePine, J. A., Algina, J. and Goeddeke, F. (2006) The adequacy of repeated-measures regression for multilevel research: Comparisons with repeated-measures anova, multivariate repeated-measures

anova, and multilevel modeling across various multilevel research designs. **Organizational Research Methods**, 9(1) January 1, 2006, pp.5-28.

218. Moore, K. L., Agure, A. M. and Dalley, A. F. (2006) Chapter 5: Lower limb. In: **Clinically oriented anatomy**. Lippincott Williams & Wilkins,

219. Muff, G., Dufour, S., Meyer, A., Masat, J., Vautravers, P., Lecocq, J. and Isner-Horobeti, M. E. (2015) Comparative study on the assessment of the strength of the knee extensor and flexor muscles by hand-held dynamometer and isokinetic dynamometer. **Annals of Physical and Rehabilitation Medicine**, 58, Supplement 19//, pp.e159-e160.

220. Nadler, S. F., DePrince, M. L., Hauesien, N., Malanga, G. A., Stitik, T. P. and Price, E. (2000) Portable dynamometer anchoring station for measuring strength of the hip extensors and abductors. **Arch Phys Med Rehabil**, 81(8) Aug, pp.1072-1076.

221. Nadler, S. F., Malanga, G. A., Bartoli, L. A., Feinberg, J. H., Prybicien, M. and Deprince, M. (2002) Hip muscle imbalance and low back pain in athletes: Influence of core strengthening. **Med Sci Sports Exerc**, 34(1) Jan, pp.9-16.

222. Nepple, J. J., Goljan, P., Briggs, K. K., Garvey, S. E., Ryan, M. and Philippon, M. J. (2015) Hip strength deficits in patients with symptomatic femoroacetabular impingement and labral tears. **Arthroscopy**, 31(11) Nov, pp.2106-2111.

223. Neumann, D. A. (2009a) Chapter 1: Kinematics. In: **Kinesiology of the musculoskeletal system: Foundations for rehabilitation**. Mosby, pp.5-7

224. Neumann, D. A. (2009b) Chapter 12: Hip. In: **Kinesiology of the musculoskeletal system: Foundations for rehabilitation**. Mosby,

225. Neumann, D. A. (2009c) Osteology, arthrology, muscle and joint interaction. In: **Kinesiology of the musculoskeletal system: Foundations for rehabilitation**. Mosby, pp.465-519

226. Neumann, D. A. (2010) Kinesiology of the hip: A focus on muscular actions. **J Orthop Sports Phys Ther**, 40(2) Feb, pp.82-94.

227. Neumann, D. A., Soderberg, G. L. and Cook, T. M. (1988) Comparison of maximal isometric hip abductor muscle torques between hip sides. **Physical Therapy**, 68(4) April 1, 1988, pp.496-502.

228. Newman, D. A. (2014) Missing data: Five practical guidelines. **Organizational Research Methods**, 17(4), pp.372-411.

229. Nguyen, A.-D., Boling, M. C., Levine, B. and Shultz, S. J. (2009) Relationships between lower extremity alignment and the quadriceps angle. **Clin J Sport Med**, 19(3), pp.201-206.

230. Nho, S. J., Magennis, E. M., Singh, C. K. and Kelly, B. T. (2011) Outcomes after the arthroscopic treatment of femoroacetabular impingement in a mixed group of high-level athletes. **The American Journal of Sports Medicine**, 39(1 suppl) July 1, 2011, pp.14S-19S.

231. Nielsen, T. G., Miller, L. L., Lund, B., Christiansen, S. E. and Lind, M. (2014) Outcome of arthroscopic treatment for symptomatic femoroacetabular impingement. **BMC Musculoskeletal Disorders**, 15(1) 11/23

232. 05/26/received

233. 11/11/accepted, p.394.

234. Oberfeld, D. and Franke, T. (2013) Evaluating the robustness of repeated measures analyses: The case of small sample sizes and nonnormal data. **Behavior Research Methods**, 45(3) 2013/09/01, pp.792-812.

235. Olson, V. L., Smidt, G. L. and Johnston, R. C. (1972) The maximum torque generated by the eccentric, isometric, and concentric contractions of the hip abductor muscles. **Phys Ther**, 52(2) Feb, pp.149-158.

236. OpenStax. (2015) **The pelvic girdle and pelvis** [Internet]. Available from: <https://cnx.org/contents/FPtK1zmh@7.30:U6HolX6c@3/The-Pelvic-Girdle-and-Pelvis> [Accessed].

237. Pheasant, S. and Haslegrave, C. M. (2006) **Bodyspace : Anthropometry, ergonomics, and the design of work**. 3rd ed. Boca raton, Taylor & Francis.

238. Philippon, M., Schenker, M., Briggs, K. and Kuppersmith, D. (2007a) Femoroacetabular impingement in 45 professional athletes: Associated pathologies and return to sport following arthroscopic decompression. **Knee Surgery, Sports Traumatology, Arthroscopy**, 15(7), pp.908-914.

239. Philippon, M. J., Christensen, J. C. and Wahoff, M. S. (2009) Rehabilitation after arthroscopic repair of intra-articular disorders of the hip in a professional football athlete. **J Sport Rehabil**, 18(1) Feb, pp.118-134.

240. Philippon, M. J., Ejnisman, L., Ellis, H. B. and Briggs, K. K. (2012a) Outcomes 2 to 5 years following hip arthroscopy for femoroacetabular impingement in the patient aged 11 to 16 years. **Arthroscopy**, 28(9) Sep, pp.1255-1261.

241. Philippon, M. J., Maxwell, R. B., Johnston, T. L., Schenker, M. and Briggs, K. K. (2007b) Clinical presentation of femoroacetabular impingement. **Knee Surgery, Sports Traumatology, Arthroscopy**, 15(8) 2007//, pp.1041-1047.

242. Philippon, M. J., Schroder, E. S. B. G. and Briggs, K. K. (2012b) Hip arthroscopy for femoroacetabular impingement in patients aged 50 years or older. **Arthroscopy**, 28(1) Jan, pp.59-65.

243. Philippon, M. J., Weiss, D. R., Kuppersmith, D. A., Briggs, K. K. and Hay, C. J. (2010) Arthroscopic labral repair and treatment of femoroacetabular impingement in professional hockey players. **The American Journal of Sports Medicine**, 38(1) January 1, 2010, pp.99-104.

244. Philippon, M. J., Yen, Y. M., Briggs, K. K., Kuppersmith, D. A. and Maxwell, R. B. (2008) Early outcomes after hip arthroscopy for femoroacetabular impingement in the athletic adolescent patient: A preliminary report. **J Pediatr Orthop**, 28(7) Oct-Nov, pp.705-710.

245. Polesello, G. C., Lima, F. R., Guimaraes, R. P., Ricioli, J. W. and Queiroz, M. C. (2014) Arthroscopic treatment of femoroacetabular impingement: Minimum five-year follow-up. **Hip International**, 24(4), pp.381-386.

246. Pollard, H., Lakay, B., Tucker, F., Watson, B. and Bablis, P. (2005) Interexaminer reliability of the deltoid and psoas muscle test. **Journal of Manipulative and Physiological Therapeutics**, 28(1) 1//, pp.52-56.

247. Pollard, T. C. B., Villar, R. N., Norton, M. R., Fern, E. D., Williams, M. R., Murray, D. W. and Carr, A. J. (2010) Genetic influences in the aetiology of femoroacetabular impingement. **Journal of Bone & Joint Surgery, British Volume**, 92-B(2), p.209.

248. Quene, H. and van den Bergh, H. (2004) On multi-level modeling of data from repeated measures designs: A tutorial. **Speech Communication**, 43(1-2) Jun, pp.103-121.

249. Quinn, A. (2010) Hip and groin pain: Physiotherapy and rehabilitation issues. **The Open Sports Medicine Journal**, 4, pp.93-107.

250. Rasch, A., Dalén, N. and Berg, H. E. (2010) Muscle strength, gait, and balance in 20 patients with hip osteoarthritis followed for 2 years after tha. **Acta Orthopaedica**, 81(2) 2010/04/01, pp.183-188.

251. Redmond, J. M., Gupta, A., Cregar, W. M., Hammarstedt, J. E., Gui, C. and Domb, B. G. (2015a) Arthroscopic treatment of labral tears in patients aged 60 years or older. **Arthroscopy**, 31(10) Oct, pp.1921-1927.

252. Redmond, J. M., Gupta, A., Dunne, K. F., Stake, C. E. and Domb, B. G. (2015b) Risk factors for early conversion to total hip arthroplasty following hip arthroscopy. **Orthopaedic Journal of Sports Medicine**, 3(7 suppl2) 07/17, p.2325967115S2325900145.

253. Reese, N. and Bandy, W. (2016) **Joint range of motion and muscle length testing**. 3rd edition ed., Saunders.

254. Reichenbach, S., Juni, P., Werlen, S., Nuesch, E., Pfirrmann, C. W., Trelle, S., Odermatt, A., Hofstetter, W., Ganz, R. and Leunig, M. (2010)

Prevalence of cam-type deformity on hip magnetic resonance imaging in young males: A cross-sectional study. **Arthritis Care Res (Hoboken)**, 62(9) Sep, pp.1319-1327.

255. Reiman, M. P. and Thorborg, K. (2014) Clinical examination and physical assessment of hip joint-related pain in athletes. **International Journal of Sports Physical Therapy**, 9(6), pp.737-755.

256. Retchford, T. H., Crossley, K. M., Grimaldi, A., Kemp, J. L. and Cowan, S. M. (2013) Can local muscles augment stability in the hip? A narrative literature review. **J Musculoskelet Neuronal Interact**, 13(1) Mar, pp.1-12.

257. Roach, K. E. and Miles, T. P. (1991) Normal hip and knee active range of motion: The relationship to age. **Physical Therapy**, 71(9), pp.656-665.

258. Rodriguez-Ruiz, D., Diez-Vega, I., Rodriguez-Matoso, D., Fernandez-del-Valle, M., Sagastume, R. and Molina, J. J. (2014) Analysis of the response speed of musculature of the knee in professional male and female volleyball players. **Biomed Res Int**, 2014, p.239708.

259. Röling, M. A., Mathijssen, N. M. C. and Bloem, R. M. (2016) Incidence of symptomatic femoroacetabular impingement in the general population: A prospective registration study. **J Hip Preserv Surg**, 3(3) 03/25

260. 08/11/received

261. 01/19/revised

262. 02/05/accepted, pp.203-207.

263. Ross, M. D., Nordeen, M. H. and Barido, M. (2003) Test-retest reliability of patrick's hip range of motion test in healthy college-aged men. **Journal of Strength & Conditioning Research (Allen Press Publishing Services Inc.)**, 17(1), pp.156-161.

264. Rylander, J. H., Shu, B., Andriacchi, T. P. and Safran, M. R. (2011) Preoperative and postoperative sagittal plane hip kinematics in patients with femoroacetabular impingement during level walking. **The American Journal of Sports Medicine**, 39(1 suppl) July 1, 2011, pp.36S-42S.

265. Saadat, E., Martin, S. D., Thornhill, T. S., Brownlee, S. A., Losina, E. and Katz, J. N. (2014) Factors associated with the failure of surgical treatment for femoroacetabular impingement: Review of the literature. **The American Journal of Sports Medicine**, 42(6) June 1, 2014, pp.1487-1495.

266. Safran, M. R. (2005) Evaluation of the hip: History, physical examination, and imaging. **Operative Techniques in Sports Medicine**, 13(1) Jan, pp.2-12.

267. Safran, M. R. and Hariri, S. (2010) Hip arthroscopy assessment tools and outcomes. **Operative Techniques in Orthopaedics**, 20(4) 12//, pp.264-277.

268. Samora, J. B., Ng, V. Y. and Ellis, T. J. (2011) Femoroacetabular impingement: A common cause of hip pain in young adults. **Clin J Sport Med**, 21(1) Jan, pp.51-56.

269. Sansone, M., Ahldén, M., Jonasson, P., Thomeé, C., Swärd, L., Baranto, A., Karlsson, J. and Thomeé, R. (2015) Good results after hip arthroscopy for femoroacetabular impingement in top-level athletes. **Orthopaedic Journal of Sports Medicine**, 3(2) 02/10, p.2325967115569691.

270. Sansone, M., Ahldén, M., Jonasson, P., Thomeé, C., Swärd, L., Collin, D., Baranto, A., Karlsson, J. and Thomeé, R. (2016a) Outcome of hip arthroscopy in patients with mild to moderate osteoarthritis—a prospective study. **J Hip Preserv Surg**, 3(1) 12/26

271. 08/12/received

272. 09/21/revised

273. 11/20/accepted, pp.61-67.

274. Sansone, M., Ahlden, M., Jonasson, P., Thomee, C., Sward, L., Ohlin, A., Baranto, A., Karlsson, J. and Thomee, R. (2016b) Outcome after hip arthroscopy for femoroacetabular impingement in 289 patients with minimum 2-year follow-up. **Scand J Med Sci Sports**,Jan 21.

275. Sapega, A. A. (1990) Muscle performance evaluation in orthopaedic practice. **J Bone Joint Surg Am**, 72(10) Dec, pp.1562-1574.

276. Savva, C., Giakas, G., Efstathiou, M. and Karagiannis, C. (2014) Test-retest reliability of handgrip strength measurement using a hydraulic hand dynamometer in patients with cervical radiculopathy. **Journal of Manipulative and Physiological Therapeutics**, 37(3) 3//, pp.206-210.

277. Schilders, E., Dimitrakopoulou, A., Bismil, Q., Marchant, P. and Cooke, C. (2011) Arthroscopic treatment of labral tears in femoroacetabular impingement: A comparative study of refixation and resection with a minimum two-year follow-up. **J Bone Joint Surg Br**, 93(8) Aug, pp.1027-1032.

278. Scott, D. A., Bond, E. Q., Sisto, S. A. and Nadler, S. F. (2004) The intra- and interrater reliability of hip muscle strength assessments using a handheld versus a portable dynamometer anchoring station. **Arch Phys Med Rehabil**, 85(4), pp.598-603.

279. Seijas, R., Alentorn-Geli, E., Álvarez-Díaz, P., Marín, M., Ares, O., Sallent, A., Cuscó, X. and Cugat, R. (2016) Gluteus maximus impairment in femoroacetabular impingement: A tensiomyographic evaluation of a clinical fact. **Archives of Orthopaedic and Trauma Surgery**, 136(6) 2016//, pp.785-789.

280. Simoneau, G. G., Hoenig, K. J., Lepley, J. E. and Papanek, P. E. (1998) Influence of hip position and gender on active hip internal and external rotation. **J Orthop Sports Phys Ther**, 28(3) Sep, pp.158-164.

281. Simpson, J., Sadri, H. and Villar, R. (2010) Hip arthroscopy technique and complications. **Orthopaedics & Traumatology: Surgery & Research**, 96(8, Supplement) 12//, pp.S68-S76.

282. Singh, P. J. and O'Donnell, J. M. (2010a) The outcome of hip arthroscopy in australian football league players: A review of 27 hips. **Arthroscopy**, 26(6) Jun, pp.743-749.

283. Singh, P. J. and O'Donnell, J. M. (2010b) The outcome of hip arthroscopy in australian football league players: A review of 27 hips. **Arthroscopy: The Journal of Arthroscopic & Related Surgery**, 26(6) 6//, pp.743-749.

284. Siqueira, C. M., Pelegrini, F. R. M. M., Fontana, M. F. and Greve, J. M. D. (2002) Isokinetic dynamometry of knee flexors and extensors: Comparative study among non-athletes, jumper athletes and runner athletes. **Revista do Hospital das Clínicas**, 57, pp.19-24.

285. Sisto, S. A. and Dyson-Hudson, T. (2007) Dynamometry testing in spinal cord injury. **J Rehabil Res Dev**, 44(1), pp.123-136.

286. Sole, G., Hamrén, J., Milosavljevic, S., Nicholson, H. and Sullivan, S. J. (2007) Test-retest reliability of isokinetic knee extension and flexion. **Arch Phys Med Rehabil**, 88(5), pp.626-631.

287. Steele, F. (2008). **Module 5: Introduction to multilevel modelling concepts**, LEMMA VLE.

288. Steele, F. (2014). **Module 15: Multilevel modelling of repeated measures data**, LEMMA VLE: p. 1-62.

289. Sterne, J. A. C., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., Wood, A. M. and Carpenter, J. R. (2009) Multiple imputation for missing data in epidemiological and clinical research: Potential and pitfalls. **BMJ**, 338.

290. Sturridge, S. and Bankes, M. (2015) Hip: Focus on acetabular dysplasia in adults. **Bone and Joint**.

291. Suzuki, H., Omori, G., Uematsu, D., Nishino, K. and Endo, N. (2015) The influence of hip strength on knee kinematics during a single-legged medial drop landing among competitive collegiate basketball players. **International Journal of Sports Physical Therapy**, 10(5), pp.592-601.

292. Taylor-Haas, J. A., Hugentobler, J. A., DiCesare, C. A., Hickey Lucas, K. C., Bates, N. A., Myer, G. D. and Ford, K. R. (2014) Reduced hip strength is associated with increased hip motion during running in young adult and adolescent male long-distance runner. **International Journal of Sports Physical Therapy**, 9(4), pp.456-467.

293. Thorborg, K., Bandholm, T. and Holmich, P. (2013a) Hip- and knee-strength assessments using a hand-held dynamometer with external belt-fixation are inter-tester reliable. **Knee Surg Sports Traumatol Arthrosc**, 21(3) Mar, pp.550-555.

294. Thorborg, K., Bandholm, T., Schick, M., Jensen, J. and Hölmich, P. (2013b) Hip strength assessment using handheld dynamometry is subject to intertester bias when testers are of different sex and strength. **Scand J Med Sci Sports**, 23(4), pp.487-493.

295. Thorborg, K., Petersen, J., Magnusson, S. P. and Holmich, P. (2010a) Clinical assessment of hip strength using a hand-held dynamometer is reliable. **Scand J Med Sci Sports**, 20(3) Jun, pp.493-501.

296. Thorborg, K., Roos, E. M., Bartels, E. M., Petersen, J. and Holmich, P. (2010b) Validity, reliability and responsiveness of patient-reported outcome questionnaires when assessing hip and groin disability: A systematic review. **Br J Sports Med**, 44(16) Dec, pp.1186-1196.

297. Thorborg, K., Serner, A., Petersen, J., Madsen, T. M., Magnusson, P. and Holmich, P. (2011) Hip adduction and abduction strength profiles in elite soccer players: Implications for clinical evaluation of hip adductor muscle recovery after injury. **Am J Sports Med**, 39(1) Jan, pp.121-126.

298. Tijssen, M., van Cingel, R., de Visser, E. and Nijhuis-van der Sanden, M. (2016) A clinical observational study on patient-reported outcomes, hip functional performance and return to sports activities in hip arthroscopy patients. **Physical Therapy in Sport**, 207//, pp.45-55.

299. Toonstra, J. and Mattacola, C. G. (2013) Test-retest reliability and validity of isometric knee-flexion and -extension measurement using 3 methods of assessing muscle strength. **J Sport Rehabil**, Technical Notes(7).

300. Tourville, T. W., Smith, H. C., Shultz, S. J., Vacek, P. M., Slauterbeck, J. R., Johnson, R. J. and Beynnon, B. D. (2013) Reliability of a new stabilized dynamometer system for the evaluation of hip strength. **Sports Health**, 5(2), pp.129-136.

301. Traina, F., Cristofolini, L., De Fine, M., Brunetti, L., Frakulli, I., Abati, C. N., D'Uva, L. and Toni, A. (2010) A reproducible and inexpensive method of measuring hip abductor strength. **Hip International: The Journal Of Clinical And Experimental Research On Hip Pathology And Therapy**, 20(4), pp.512-517.

302. Tran, P., Pritchard, M. and O'Donnell, J. (2013) Outcome of arthroscopic treatment for cam type femoroacetabular impingement in adolescents. **ANZ Journal of Surgery**, 83(5), pp.382-386.

303. Trudelle-Jackson, E., Emerson, R. and Smith, S. (2002) Outcomes of total hip arthroplasty: A study of patients one year postsurgery. **Journal of Orthopaedic & Sports Physical Therapy**, 32(6) 2002/06/01, pp.260-267.

304. Tyler, T. F., Nicholas, S. J., Campbell, R. J. and McHugh, M. P. (2001) The association of hip strength and flexibility with the incidence of adductor muscle strains in professional ice hockey players. **The American Journal of Sports Medicine**, 29(2) March 1, 2001, pp.124-128.

305. Tyler, T. F., Silvers, H. J., Gerhardt, M. B. and Nicholas, S. J. (2010) Groin injuries in sports medicine. **Sports Health**, 2(3), pp.231-236.

306. Vaz, S., Falkmer, T., Passmore, A. E., Parsons, R. and Andreou, P. (2013) The case for using the repeatability coefficient when calculating test–retest reliability. **PLoS ONE**, 8(9), p.e73990.

307. Wadsworth, C. T., Krishnan, R., Sear, M., Harrold, J. and Nielsen, D. H. (1987) Intrarater reliability of manual muscle testing and hand-held dynametric muscle testing. **Physical Therapy**, 67(9) September 1, 1987, pp.1342-1347.

308. Wall, P. D. H., Brown, J. S., Karthikeyan, S. and Griffin, D. (2012) An introduction to hip arthroscopy. Part two: Indications, outcomes and complications. **Orthopaedics and Trauma**, 26(1) 2//, pp.38-43.

309. Wall, P. D. H., Brown, J. S., Karthikeyan, S., Wyse, M. and Griffin, D. (2011) An introduction to hip arthroscopy. Part one: Surgical anatomy and technique. **Orthopaedics and Trauma**, 25(6) 12//, pp.441-447.

310. Wang, R., Bhandari, M. and J. Lachowski, R. (2017) **A systematic approach to adult hip pain, part 1**.

311. Whiteley, R., Jacobsen, P., Prior, S., Skazalski, C., Otten, R. and Johnson, A. (2012) Correlation of isokinetic and novel hand-held dynamometry measures of knee flexion and extension strength testing. **Journal of Science and Medicine in Sport**, 15(5) 9//, pp.444-450.

312. WHO (2008). **Waist circumference and waist–hip ratio: Report of a who expert consultation**.

313. WHO. (2016) **Women's health** [Internet]. Available from: <http://www.who.int/mediacentre/factsheets/fs334/en/> [Accessed 22nd July, 2016].

314. Widler, K. S., Glatthorn, J. F., Bizzini, M., Impellizzeri, F. M., Munzinger, U., Leunig, M. and Maffiuletti, N. A. (2009) Assessment of hip abductor muscle

strength. A validity and reliability study. **J Bone Joint Surg Am**, 91(11) 2009-11-01 00:00:00, pp.2666-2672.

315. Willson, J. D., Petrowitz, I., Butler, R. J. and Kernozek, T. W. (2012) Male and female gluteal muscle activity and lower extremity kinematics during running. **Clinical Biomechanics**, 27(10) 12//, pp.1052-1057.

316. Wilson, J. J. and Furukawa, M. (2014) Evaluation of the patient with hip pain. **Am Fam Physician**, 89(1) Jan 1, pp.27-34.

317. Wollin, M., Purdam, C. and Drew, M. K. (2016) Reliability of externally fixed dynamometry hamstring strength testing in elite youth football players. **Journal of Science and Medicine in Sport**, 19(1) 1//, pp.93-96.

318. Zapparoli, F. Y. and Riberto, M. (2016) Isokinetic evaluation of the hip flexor and extensor muscles: A systematic review. **Journal of Sport Rehabilitation**, pp.1-20.

319. Zimny, N. J. (1998) Clinical reasoning in the evaluation and management of undiagnosed chronic hip pain in a young adult. **Physical Therapy**, 78(1) January 1, 1998, pp.62-73.

320. Zuša, A., Lanka, J., Čupriks, L. and Dravniece, I. (2015) A descriptive profile of isometric muscle strength and muscle strength imbalance in young tennis players. **Baltic Journal of Sport & Health Sciences**, 99(4), pp.54-61.