INTERNATIONAL JOURNAL OF SPORTS PHYSICAL THERAPY

Featuring Abstracts from the IFSPT/BFSP Partnered International Congress
September 22-23, 2023
TO BE THE BEST, TRAIN WITH THE BEST.

WOODWAY

WOODWAY is proudly used in professional sports, Olympic training, private performance facilities, university athletics, US military, and endurance facilities.

Contact us today and begin training with the best!

WOODWAY WORLD HEADQUARTERS
W229 N591 Foster Court Waukesha, WI 53186 | PHONE (800) 966-5929 | FAX (262) 522-6235
WOODWAY.COM
Extracorporeal Shockwave Therapy (ESWT) is a versatile modality that can be used to treat a variety of musculoskeletal dysfunctions. In this free, on-demand webinar, Dr. Pavel Novak, Director of Science and Technology at Storz Medical AG in Tägerwilen, Switzerland, provides insight that will help you drive positive outcomes with this technology.

Dr. Novak has been involved with the design of shockwave equipment for the past 25 years in addition to conducting other biomedical research for almost 50 years. In that time, he has developed 90 patents and written more than 100 research papers and presentations. Dr. Novak’s extensive experience with shockwave technology, biomedical research, and product development makes for an illuminating discussion. Topics of discussion include the following:

- Identifying an inferior piece of ESWT equipment
- RPW vs FSW
- Mechanisms behind treatment outcomes
- The future of ESWT technology

Register using the QR code above or visit learn.chattanoogarehab.com/ijspt-apr-23-journal
IJSPT is an official journal of the International Federation of Sports Physical Therapy (IFSP). Countries with access to IJSPT as a member benefit:

- Argentina
- Australia
- Austria
- Belgium
- Bulgaria
- Brazil
- Cameroon
- Canada
- Chile
- Cyprus
- Denmark
- Finland
- France
- Germany
- Greece
- Hong Kong
- Indonesia
- Israel
- Italy
- Japan
- Luxembourg
- The Netherlands
- New Zealand
- Nigeria
- Norway
- Portugal
- Saudi Arabia
- Slovenia
- South Africa
- South Korea
- Spain
- Sweden
- Switzerland
- Thailand
- Turkey
- United Kingdom
- Zambia

Reach us at www.ifspt.org.
Ellen Shanley
PhD, PT, OCS
Sr. Director of Care Delivery Optimization

Charles A. Thigpen
PT, PhD, ATC
Vice President of Care Delivery

Ellen Shanley
PhD, PT, OCS
Sr. Director of Care Delivery Optimization

Thomas Denninger
DPT, OCS, FAAOMPT
Sr. Director of Learning and Development

**Topic**
Shoulder Sports & Arthroplasty & Elbow

**Program**
The Postop Rehab Protocol: Someday Maybe We Will All Agree But Not Today

**Thursday, Sept 21**
11:23 am

**Topic**
Orthopaedic Sports Rehab

**Program**
Rotator Cuff Repair Or Total Shoulder Arthroplasty: Can They Return To Sports?

**Friday, Sept 22**
3:00 pm

**Topic**
Orthopaedic Sports Rehab

**Program**
A Pain In The Neck? Keys To A Correct Differential Diagnosis

**Friday, Sept 22**
2:15 pm

To view all sessions featuring ATI’s Care Delivery team, visit orthosummit.com

ATIpt.com

IS HONORED TO JOIN
INDUSTRY EXPERTS AT OSET 2023

Sept 21 - 23, 2023 | Boston, MA
Introducing PiezoWave²T

FASTER. SMARTER. LIGHTER.
The Best Rehab Technology Just Got Better

Controlling Pain. Restoring Motion.
BACKED BY SCIENCE

Benefits to Your Patients
- Boost circulation and lymphatic return
- Decrease edema
- Clear inflammation
- Increase ROM
- Decrease muscle soreness
- Alleviate pain

Benefits to Your Practice
- Improve patient outcomes
- Increase revenue through insurance, cash based services, and retail sales
- Attract new patients
- Increase retention

Interested in finding out more?
Reach out today to Rehab@Hyperice.com

Learn more about the research and science behind Hyperice technology, how other practices are utilizing the products, and discuss the option of receiving a free demo kit to trial in your clinic.
Gold Standard Sports Medicine Solutions

From Pre-Op to Return-to-Play

Biodex™ advanced rehabilitation technology allows clinicians to quantify performance parameters – before and after an injury occurs.

Detailed reports track recovery and provide the medical team with quantitative data to help with the return-to-play decision.

Understand Test Results at a Glance

Return-to-Play Reports help to simplify the RTP decision with clear pass/fail results.
Technology designed for rehabilitation and improving movement regardless of age or level of physical capability.

IMMERSIVE REACTION-BASED ACTIVITIES
Assessments | Workouts | Injury-Specific Protocols | Drills | Games

MEASURE WHAT MATTERS
Simultaneously measures physical and cognitive function for holistic rehabilitation and improved neuromechanical performance.

ALIGNING CARE, DATA, & ROI
Aligned with CPT 97 Billing Codes
Aligned to deliver better outcomes for end-users, care professionals, and organizations.

DOWNLOAD: How TRAZER Works
Aligned at the point where biology, technology, and data intersect.

WATCH: TRAZER in Action

www.trazer.com
EDITORIAL BOARD

David Behm, PhD  
Memorial University of Newfoundland  
St. John’s, Newfoundland, Canada

Barton N. Bishop, PT, DPT, SCS, CSCS  
Kaizo Clinical Research Institute  
Rockville, Maryland, USA

Mario Bizzini, PhD, PT  
Schulthess Clinic Human Performance Lab  
Zürich, Switzerland

Joe Black, PT, DPT, SCS, ATC  
Total Rehabilitation  
Maryville, Tennessee, USA

Turner A. "Tab" Blackburn, APTA Life Member, ATC-Ret, AOSSM-Ret  
NASMI  
Lanett, AL, USA

Lori Bolgla, PT, PhD, MAcc, ATC  
Augusta University  
Augusta, Georgia, USA

Matthew Briggs  
The Ohio State University  
Columbus, OH, USA

Tony Brosky, PT, PhD  
Bellarmine University  
Louisville, KY, USA

Brian Busconi, MD  
UMass Memorial Hospital  
Boston, MA, USA

Robert J. Butler, PT, PhD  
St. Louis Cardinals  
St. Louis, MO, USA

Duane Button, PhD  
Memorial University  
St. Johns, Newfoundland, Canada

J. W. Thomas Byrd, MD  
Nashville Sports Medicine and Orthopaedic Center  
Nashville, TN, USA

Lyle Cain, MD  
Andrews Institute & Sports Medicine Center  
Birmingham, AL, USA

Gary Calabrese, PT, DPT  
Cleveland Clinic  
Cleveland, Ohio, USA

Meredith Chaput, PT, DPT, SCS  
Ohio University  
Athens, OH, USA

Rita Chorba, PT, DPT, MAT, SCS, ATC, CSCS  
United States Army Special Operations Command  
Fort Campbell, KY, USA

John Christoforetti, MD  
Texas Health  
Dallas, TX, USA

Richard Clark, PT, PhD  
Tennessee State University  
Nashville, TN, USA

Juan Colado, PT, PhD  
University of Valencia  
Valencia, Spain

Brian Cole, MD  
Midwest Orthopaedics at Rush  
Chicago, IL, USA

Ann Cools, PT, PhD  
Ghent University  
Ghent, Belgium

Andrew Contreras, DPT, SCS  
Washington, DC, USA

George Davies, PT, DPT, MEd, SCS, ATC, LAT, CSCS, PES, FAPTA  
Georgia Southern University  
Savannah, Georgia, USA

Pete Draovitch, PT  
Jacksonville Jaguars Football  
Jacksonville, FL, USA

Jeffrey Dugas, MD  
Andrews Institute & Sports Medicine Center  
Birmingham, AL, USA

Jiri Dvorak, MD  
Schulthess Clinic  
Zurich, Switzerland

Todd Ellenbecker  
Rehab Plus  
Phoenix, AZ, USA

Carolyn Emery, PT, PhD  
University of Calgary  
Calgary, Alberta, Canada

Ernest Esteve Caupena, PT, PhD  
University of Girona  
Girona, Spain

Sue Falsone, PT, MS, SCS, ATC, CSCS, COMT  
Structure and Function Education and A.T. Still University  
Phoenix, Arizona, USA

J. Craig Garrison, PhD, PT, ATC, SCS  
Texas Health Sports Medicine  
Fort Worth, Texas, USA

Maggie Gebhardt, PT  
Fit Core Physical Therapy/Myopain Seminars  
Atlanta, GA and Bethesda, MD, USA

Lance Gill, ATC  
LG Performance-TPI  
Oceanside, CA, USA

Phil Glasgow, PhD, MTh, MRes, MCSP  
Sports Institute of Northern Ireland  
Belfast, Northern Ireland, UK

Robert S. Gray, MS, AT  
Cleveland Clinic Sports Health  
Cleveland, Ohio, USA

Jay Greenstein, DC  
Kaizo Health  
Baltimore, MD, USA
EDITORIAL BOARD

Martin Hagglund, PT PhD
Linkoping University
Linkoping, Sweden

Allen Hardin, PT, SCS, ATC, CSCS
University of Texas
Austin, TX, USA

Richard Hawkins, MD
Professor of surgery, University of South Carolina
Adjunct Professor, Clemson University
Principal, Steadman Hawkins, Greenville and Denver (CU)

John D. Heick, PT, PhD, DPT, OCS, NCS, SCS
Northern Arizona University
Flagstaff, AZ, USA

Tim Hewett, PhD
Hewett Consulting
Minneapolis, Minnesota, USA

Per Holmich, MD
Copenhagen University Hospital
Copenhagen, Denmark

Kara Mae Hughes, PT, DPT, CSCS
Wolfe PT
Nashville, TN, USA

Lasse Ishoi, PT, MSc
Sports Orthopedic Research Center
Copenhagen University Hospital
Hvidovre, Denmark

Jon Karlsson, MD
Sahlgrenska University
Goteborg, Sweden

Brian Kelly, MD
Hospital for Special Surgery
New York, NY, USA

Benjamin R. Kivlan, PhD, PT, OCS, SCS
Duquesne University
Pittsburgh, PA, USA

Dave Kohlireser, PT, DPT, SCS, OCS, CSCS
Ortho One
Columbus, OH, USA

Andre Labbe PT, MOPT
Tulane Institute of Sports Medicine
New Orleans, LA USA

Henning Langberg, PT, PhD
University of Copenhagen
Copenhagen, Denmark

Robert LaPrade, MD
Twin Cities Orthopedics
Edina, MN, USA

Lance Lueldke, PT, DPT
University of Wisconsin Oshkosh
Oshkosh, WI, USA

Lenny Macrina, PT, SCS, CSCS, C-PS
Champion Physical Therapy and Performance
Boston, MA, USA

Phillip Malloy, PT, PhD
Arcadia University/Rush University Medical Center
Glenside, PA and Chicago, IL, USA

Terry Malone, PT, EdD, ATC, FAPTA
University of Kentucky
Lexington, KY, USA

Robert Mangine, PT
University of Cincinnati
Cincinnati, OH, USA

Eric McCarty, MD
University of Colorado
Boulder, CO, USA

Ryan P. McGovern, PhD, LAT, ATC
Texas Health Sports Medicine Specialists
Dallas/Fort Worth, Texas, USA

Mal McHugh, PhD
NISMAT
New York, NY, USA

Joseph Miller, PT, DSc, OCS, SCS, CSCS
Pikes Peak Community College
Colorado Springs, CO, USA

Havard Moksnes, PT PhD
Oslo Sports Trauma Research Center
Oslo, Norway

Michael J. Mullaney, PT, SCS
NISMAT
Mullaney & Associates Physical Therapy
New York, NY and Matawan, NJ, USA

Andrew Murray, MD, PhD
European PGA Tour
Edinburgh, Scotland, UK

Andrew Naylor, PT, DPT, SCS
Bellin Health
Green Bay, WI, USA

Stephen Nicholas, MD
NISMAT New York
New York, NY, USA

John O’Donnel, MD
Royal Melbourne Hospital
Melbourne, Australia

Russ Paine, PT
McGovern Medical School
Houston, TX, USA

Snehal Patel, PT, MSPT, SCD
HSS Sports Rehabilitation Institute
New York, NY, USA

Marc Philippon, MD
Steadman-Hawkins Clinic
Vail, CO, USA

Kevin Plancher, MD, MPH, FAAOS
Plancher Orthopedics and Sports Medicine
New York, NY USA
EDITORIAL BOARD

Marisa Pontillo, PT, PhD, DPT, SCS
University of Pennsylvania Health System
Philadelphia, PA, USA

Matthew Provencher, MD
Steadman Hawkins Clinic
Vail, CO, USA

Charles E. Rainey, PT, DSc, DPT, MS, OCS, SCS, CSCS, FAAOMPT
United States Public Health Service
Springfield, MO, USA

Alexandre Rambaud, PT PhD
Saint-Etienne, France

Carlo Ramponi, PT
Physiotherapist, Kinè Rehabilitation and Orthopaedic Center
Treviso, Italy

Michael Reiman, PT, PhD
Duke University
Durham, NC, USA

Mark F. Reinking, PT, PhD, SCS, ATC
Regis University
Denver, CO, USA

Mike Reinold, PT, DPT, SCS, ATC, CSCS, C-PS
Champion Physical Therapy and Performance
Boston, MA, USA

Mark Ryan, ATC
Steadman-Hawkins Clinic
Vail, CO, USA

David Sachse, PT, DPT, OCS, SCS
USAF
San Antonio, TX, USA

Marc Safran, MD
Stanford University
Palo Alto, CA, USA

Alanna Salituro, PT, DPT, SCS, CSCS
New York Mets
Port Saint Lucie, FL, USA

Mike Samukawa, PT, PhD, AT (JSPO)
Hokkaido University
Sapporo, Japan

Barbara Sanders, PT, PhD, FAPTA, Board Certified Sports Physical Therapy Emeritus
Professor and Chair, Department of Physical Therapy
Texas State University
Round Rock, TX, USA

Felix “Buddy” Savioe, MD, FAAOS
Tulane Institute of Sport Medicine
New Orleans, LA, USA

Teresa Schuemann, PT, DPT, ATC, CSCS, Board Certified Specialist in Sports Physical Therapy
Evidence in Motion
Fort Collins, CO, USA

Timothy Sell, PhD, PT, FACSM
Atrium Health Musculoskeletal Institute
Charlotte, NC, USA

Andreas Serner, PT PhD
Aspetar Orthopedic and Sports Medicine Hospital
Doha, Qatar

Ellen Shanley, PT, PhD
ATI
Spartanburg, SC, USA

Karin Silbernagel, PT, PhD
University of Delaware
Newark, DE, USA

Holly Silvers, PT, PhD
Velocity Physical Therapy
Los Angeles, CA, USA

Lynn Snyder-Mackler, PT, ScD, FAPTA
STAR University of Delaware
Newark, DE, USA

Alston Stubbs, MD
Wake Forest University
Winston-Salem, NC, USA

Amir Takla, B.Phys, Mast.Physio (Manip), A/Prof
Australian Sports Physiotherapy
The University of Melbourne
Melbourne, Australia

Charles Thigpen, PhD, PT, ATC
ATI
Spartanburg, SC, USA

Steven Tippett, PT, PhD, ATC, SCS
Bradley University
Peoria, IL, USA

Tim Tyler, PT, ATC
NISMAT
New York, NY, USA

Timothy Uhl, PT, PhD, ATC
University of Kentucky
Lexington, KY, USA

Bakare Ummukulthoum, PT
University of the Witswatersrand
Johannesburg, Gauteng, South Africa

Yuling Leo Wang, PT, PhD
Sun Yat-sen University
Guangzhou, China

Mark D. Weber, PT, PhD, SCS, ATC
Texas Women's University
Dallas, TX, USA

Richard B. Westrick, PT, DPT, DSc, OCS, SCS
US Army Research Institute
Boston, MA, USA

Chris Wolfe, PT, DPT
Belmont University
Nashville, TN, USA

Tobias Wörner, PT, MSc
Lund University
Stockholm, Sweden
# TABLE OF CONTENTS

## VOLUME 18, NUMBER 4

<table>
<thead>
<tr>
<th>Page</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>EDITORIAL</strong></td>
</tr>
<tr>
<td></td>
<td>Can Leveraging Technology “Save” Physical Therapy?</td>
</tr>
<tr>
<td></td>
<td>Reuben Gobezie, MD</td>
</tr>
</tbody>
</table>

|      | **CLINICAL VIEWPOINT**  |
|      | Lupowitz LG. |

|      | **ORIGINAL RESEARCH**  |
|      | Thoracolumbar and Lumbopelvic Spinal Alignment During the Barbell Back Squat: A Comparison Between Men and Women.  |
|      | Rasmussen RG, Jacomson JS, Blaabjerg B, et al. |

|      | Thoracolumbar and Lumbopelvic Spinal Alignment During The Barbell Back Squat: A Comparison Between Men and Women.  |
|      | Bengtsson V, Berglund L, Ohberg F, et al. |

|      | The Effect Of Volitional Preemptive Abdominal Contraction On Biomechanical Measures During A Front Versus Back Loaded Barbell Squat.  |

|      | Effect of Footwear versus Barefoot on Double-Leg Jump-Landing and Jump Height Measures: A Randomized Cross-Over Study.  |
|      | Hébert-Losier K, Boswell-Smith C, Hanzlíková A. |

|      | Concurrent Force Feedback on Load Symmetry in Total Knee Arthroplasty Patients.  |

|      | The Association of Joint Power Kinetic Variables with Running Injuries: A Case Control Study.  |
|      | Dewald ML, Dalland JL, Stockland JR |

|      | No Effect of Return to Sport Test Batteries with and without Psychological PROs on the Risk of a Second ACL Injury: A Critical Assessment of Four Different Test Batteries.  |

|      | 2D and 3D Biomechanical Factors During 90° Change of Direction Are Associated with Non-contact ACL Injury in Female Soccer Players.  |

|      | Balance Error Scoring System Performance Differences in Figure Skaters Based on Discipline.  |
|      | Mangum LC, Skibski A, Devorski L, et al. |

|      | Maximal and Explosive Muscle Strength during Hip Adduction Squeeze and Hip Abduction Press Test using a Handheld Dynamometer: An Intra- and Inter-tester Reliability Study.  |
|      | Ishii L, Thorborg K, Krohn L et al. |

|      | The Validity and Reliability of a Smartphone Application for Break-Point Angle Measurement during Nordic Hamstring Exercise.  |
|      | (Hirose) Soga T, Yamaguchi S, Inami T, et al. |
Evaluating Psychometric Properties of the International Knee Documentation Committee Subjective Knee Form in a Heterogeneous Sample of Post-Operative Patients.
(Baker) Richardson RD, Casanova MP, Reeves AJ, et al.

Novice Inter-rater Reliability on the Selective Functional Movement Assessment (SFMA) After a 4-Hour Training Session.
Harper B, Aron A.

Does High Medial Elbow Stress During Pitching Compromise the Dynamic Stabilizers of The Elbow?
McHugh MP, Mullaney MJ.


Reliability Analysis of In-person and Virtual Goniometric Measurements of the Upper Extremity.

Normative Values of Isometric Shoulder Strength among Healthy Adults.
Bradley H, Pierpoint L.

Validity, Reliability, and Efficiency of a Standard Goniometer, Medical Inclinometer, and Builder's Inclinometer.
Hanks J, Myers BA

A Novel Intrinsic Foot Muscle Strength Dynamometer Demonstrates Moderate-To-Excellent Reliability and Validity.
Xu J, Goss DJ, Saliba S.

CLINICAL COMMENTARY
A Model for Applying Situational Awareness Theory to the Return to Sport Continuum.
Porter K, Hoch M

ULTRASOUND BITES
Musculoskeletal Ultrasound: An Essential Tool in Diagnosing Patellar Tendon Injuries.
Manske RC, Page P, Voight M, Wolfe C.

INTERNATIONAL CONGRESS 2023
Abstracts for the 2023 International Symposium, Brussels, Belgium.
Stimulates the body’s natural repair mechanisms with our non-invasive current. The unique combination of Winback energy & manual treatments offers great rehabilitation results and a faster recovery of muscle and joint functions.

Contact us for more information:
contact@winback.us
www.winback.us
Winback America
1133 Westchester Avenue Suite N-220
White Plains, NY 10604

Relieve pain.
Free up movement.
Speed up healing.
**DYNAMIC COMPRESSION CRYOTHERAPY**

- Patented dynamic, distal to proximal cryocompression
- Dependable and durable
- Effectively alleviates pain and edema
- No water, lightweight, rechargeable
- Less than one minute setup

![SQUID Image]

**Personal system from $300**

**Professional system from $700**

---

**ISHA 2023**

Annual Scientific Meeting
Cape Town, South Africa
5 - 7 October 2023

**Scientific programme topics:**
1. Patient Assessment and Pre-operative Planning
2. Role of Biologics in Hip Preservation
3. Patient Education
4. Artificial Intelligence, Virtual Reality, Augmented Reality
5. Arthroscopic Surgery
6. Open Hip Surgery
7. Nutrition, Well-being and Rehabilitation of the Elite Athlete and Others

**Call for Papers**

ISHA invites submissions from all with an interest in treatment, rehabilitation and conditioning around injury through sport, including physiotherapists, conditioning specialists and biokineticists

Abstract submission deadline 16 April
Early registration 1 March - 30 July

**Post-Conference Physiotherapy & Return to Sport Education Day - 7 October 2023**

A fantastic opportunity for worldwide physical therapists, sports rehab and conditioning specialists to come together to learn about and discuss the latest issues, approaches and experiences in global hip management

**Provisional Programme Elements include:**
- Anterior Hip Pathology
- Lateral Hip Pathology
- Posterior Hip Pathology
- Intra-articular Hip Pathology
Welcome to the Fifth World Congress of Sports Physical Therapy, presented by the International Federation of Sports Physical Therapy and NFFs Faggruppe for Idrettsfysioterapi og Aktivitesmedisin!

Join us June 14-15, 2024 in beautiful Oslo, Norway for a fantastic experience, whether you are a novice or experienced clinician. You will meet up with colleagues from all around the world, sharing knowledge, best practices, create networks and have a lot of fun. Enjoy an exciting schedule full of interesting lectures and engaging workshops!

**Key Dates:**
- February 28, 2024: Abstract Submission
- March 28, 2024: Early Bird Pricing Deadline
- May 14, 2024: Workshops
- June 13, 2024: IFSPT General Meeting
- June 14-15, 2024: Congress

**From Research to Clinical Practice**
- 18 practical Workshops
- Oral presentations
- Posters

**REGISTRATION IS NOW OPEN AT WCSPT.ORG!**
drive revenue through remote monitoring

deliver a virtual therapy solution to your organization

geniehealth at a glance...

Founded by two prominent orthopedic surgeons and leveraged by two of the top 10 largest orthopedic groups in the country, Genie Health is managed by therapists and industry experts.

Featuring a monitored HEP using computer vision, Genie Health offers both fee-for-service and value-based-care models on the same platform.

tech platform & optional clinical staff turn-key solution

PT genie
digital physical therapy solution combining remote monitoring and telehealth

sports genie
in-clinic and remote sports/functional assessment and management

woRx genie
risk assessment and remote management tools for occupational health

improve
revenue accessibility staffing challenges

genie.health | sales@genie.health | 321-558-6855
**IMPROVED EXPERIENCES FOR PATIENTS AND YOUR TEAM**

**AWARD-WINNING DIAL-IN RESISTANCE SYSTEM.**
- Instant, dial-in system offers 10–100 pounds of resistance
- More comfortable accessories for targeted resistance from ankle to shoulder
- Ability to anchor ANYWHERE, ANYTIME, IN LESS TIME.

---

**LIMITLESS OPTIONS. FASTER SETUP TIMES.**
- Delivers widest range of isometric and dynamic exercise options in a single system
- Quickly swap in VECTORS to address stability, mobility and strength needs
- Optimal solution for smaller spaces or when traveling

---

“I UTILIZE MY VECTOR EVERY DAY AND I LOVE IT!”

*Kevin Wilk, DPT, FAPTA - Champion Sports Medicine*

Visit us at www.kayezen.com/ljspt
We are a mission-driven online community seeking to change outcomes with research while helping direct-access clinicians improve the overall awareness and quality of care for concussion patients across the globe with robust educational programming, open office hours, and non-profit partnerships.

JOIN US FOR FALL 2023

CONCUSSION
The Patient Rehabilitation Journey

DURATION  
Sep-3 thru Dec-10, 2023

CLASS  
12 Weeks Online

CONTENT  
On-Demand | Virtual

COURSE FEE  
Starting at $797

Group Pricing Available

CE  
20* or 40 Hours
PT, *OT, *ATC: All 50 United States

A donation will be made to our Non-Profit Partner, Headway Foundation, when you sign up with the IJSPT QR code below.

REGISTRATION OPEN NOW
ConcussionCorner.org
JOIN OUR ONLINE COMMUNITY

EDUCATIONAL PARTNER

IJSPT
INTERNATIONAL JOURNAL OF SPORTS PHYSICAL THERAPY
As a healthcare provider for over 20 years in the musculoskeletal (MSK) space, I acknowledge a simple reality: physical therapy is a critical part of our nation’s efforts to improve the quality of life for our citizens. PT is necessary, vital, valuable, important, and good. But despite what may seem obvious to those of us in ‘the field’, the reality is that market forces of various kinds have made the business of physical therapy…which is necessary for the provision of physical therapy…challenging. The forces shaping the business and the practice of physical therapy include the pandemic, Medicare and private payors, hospital systems, consumers, and the availability of physical therapists (PTs) and physical therapist assistants (PTAs) as well as the reality of higher inflation impacting all aspects of overhead in brick-and-mortar facilities. In general, developed societies like ours must tackle these challenges with innovation in order to “survive and thrive”. The question I want to address in this feature is: can technology and innovation “save” physical therapy?

Let me set-the-table for our discussion with a fact-check:

- Physical/Occupational Therapy is a $34 billion/year market growing at an estimated 6.2% average annual rate. The PT market is expected to be $43 billion by 2025.

- Physical therapy services reimbursements over the last 4 years have been cut from anywhere between 15-30%.

- There is a severe shortage of PTs and PTAs. The average PT clinic has a 17% vacancy in their PT staff and a 16% vacancy in their PTA staff. This problem shows no signs of slowing and is projected to worsen significantly over next 10 years.

- There are just over 42,300 physical therapy clinics with an average annual revenue of $855,000 and average net profit of 14%. The average income for a physical therapist is $91,000/year.

- A survey of physical therapy practice owners concluded that 77% of owners felt the loss of PTs and PTAs in their clinics was due to work-life balance and compensation.

- There is a tremendous growth in the number of PT's/OT's transitioning into virtual care, resulting in a more flexible work schedule and the ability to work from home...a luxury never previously possible for therapy providers.

We are very much in a state of transition within the physical therapy market and the healthcare marketplace, overall. The challenges we face will require innovations in healthcare delivery, human resources management and payment models in order to navigate to calmer waters where the enjoyment of the practice of physical therapy can be assured for the average practitioner.

I believe the pandemic and the introduction of the remote therapeutic monitoring codes (RTM) have given rise to a “Kairos” moment, a pivotal time for action and decision, in the physical therapy space.
The pandemic accelerated the consumerization of healthcare such that patients expect to have a virtual component to their healthcare journey. The RTM codes have provided the impetus and the funding to strive for innovative methods to monetize home exercise programs using insurance as a means for payment. These two factors will give rise to new healthcare delivery models that shape the future of physical therapy. Genie Health is working closely with physical therapy practices today to deliver care in a hybrid model using our technology to enable therapists to manage access for their patients, leverage technology to monetize home exercise programs and obtain valuable functional data from their patients using the most cutting-edge computer vision technology with the goal of standardizing the quality of care, demonstrating value through quantitative analysis of outcomes, and enabling more leverage for insurance contracts with an eye towards sharing risk with payors.

In subsequent issues of the IJSPT, I will be working with the Editorial Staff to introduce thought leaders within our space to share their ideas and strategies for tackling the challenges present in our profession. The present is tumultuous...but, at Genie Health, we believe there is a break in the storm ahead through innovation and technology.

Reuben Gobezie, MD  
CEO, Genie Health
Patient-reported Outcomes and Muscle Strength after a Physiotherapy-led Exercise and Support Brace Intervention in Patients with Acute Injury of the Posterior Cruciate Ligament: A Two-year Follow-up Study

Randi Gram Rasmussen, Julie Sandell Jacobsen, Birgitte Blaabjerg, Torsten Grønbæk Nielsen, Lene Lindberg Miller, Martin Lind

1 Department of Physiotherapy and Occupational Therapy, Aarhus University Hospital, 2 Research Centre for Health and Welfare Technology, Programme for Rehabilitation, VIA University College, 3 Department of Orthopaedic Surgery, Division of Sportstraumatology, Aarhus University Hospital

Keywords: Physiotherapy, Progressive exercises, Non-operative treatment, Isometric knee strength, Self-reported outcome

https://doi.org/10.26603/001c.83214

Background
While outcomes of posterior cruciate ligament (PCL) injuries treated surgically are well described, prospective studies reporting outcomes of exercise interventions are lacking.

Purpose
The purpose of this study was to investigate changes in patient-reported outcomes of a physiotherapy-led exercise and support brace intervention in patients with acute injury of the PCL over a two-year follow-up period. Furthermore, this study sought to investigate changes in isometric knee muscle strength over an eight-month follow-up period, and finally to report conversion to surgical reconstruction over a two-year follow-up period.

Study design
Case series study, prospective

Methods
Fifty patients with an acute injury of the PCL were treated with a brace and a physiotherapy-led exercise intervention and followed prospectively. Changes in patient-reported outcomes were measured with the International Knee Documentation Committee Subjective Knee Form (IKDC-SKF) and the Knee injury and Osteoarthritis Outcome Score (KOOS) from baseline (diagnosis) to two-year follow-up. Furthermore, changes in isometric knee flexion and extension strength were measured with a static strength dynamometer from 16 weeks after diagnosis to one-year follow-up. Conversion to surgery was prospectively extracted from medical records. Mean changes were analyzed with a mixed effects model with time as a fixed factor.

Results
The IKDC-SKF score improved 28 (95%CI 24-33) IKDC points from baseline to two-year follow-up. Isometric knee flexion strength of the injured knee increased 0.18 (95%CI 0.11-0.25) Nm/kg from 16 weeks after diagnosis to one-year follow-up, corresponding to

Corresponding Author:
Randi Gram Rasmussen
Department of Physiotherapy and Occupational Therapy
Aarhus University Hospital
8200 Aarhus
Denmark
Phone: +45 7845 6750
Email: ranrasmu@rm.dk
an increase of 16%. In contrast, isometric knee extension strength of the injured knee did not change (0.12 (95%CI 0.00-0.24) Nm/kg, p=0.042). Over two years, seven patients converted to PCL surgical reconstruction. One and two-year follow-up were completed by 46 and 51 patients, respectively.

Conclusions

The physiotherapy-led exercise and support brace intervention demonstrated clinically relevant improvements in patient-reported outcomes and knee flexion strength, and the risk of PCL surgical reconstruction was considered low within the first two years.

Level of evidence

5b
©The Author(s)

INTRODUCTION

The posterior cruciate ligament (PCL) resists excessive posterior tibial translation and excessive rotation of the knee at different knee flexion angles.1 PCL injuries are uncommon with an estimated annual incidence of 2 per 100,000.5,3 These injuries normally occur due to high-speed traffic and sports-related accidents,1,4-6 and the prevalence is highest among young men.5,7-9 The major deficits associated with PCL injury is abnormal tibiofemoral kinematics10,11 and early knee osteoarthritis.11,12 These deficits are experienced as discomfort and knee instability.13 Historically, PCL injuries have been treated surgically14 but in recent years, it has been accepted that the PCL has a natural ability to heal7,8,14 and non-operative interventions have become first-line treatments.7 Several studies have reported clinically relevant patient-reported improvements after non-operative interventions.5,6,8,9,12,14,15 However, results of previous studies have reported insufficient PCL stability following both non-operative interventions and surgical reconstructions.7,8 In addition, prevalence of knee osteoarthritis after PCL reconstruction has been reported to be similar or worse compared to non-operative interventions.5,16-20

Non-operative treatments normally cover supervised exercise and support brace interventions focusing on the most optimal healing of the PCL and regaining range of motion (ROM) and strength with an emphasis on quadriceps activation.1,14,21 Several authors have reported normalized knee strength following non-operative interventions.5,6,15 as well as clinically relevant long-term improvements in patient-reported outcomes.12,22 However, the majority of these studies fail to report the specific rehabilitation regimens and the support brace interventions.5,7,9,16,19,23 Therefore, the primary aim was to investigate changes in patient-reported outcomes of a physiotherapy-led exercise and support brace intervention in patients with acute injury of the PCL over a two-year follow-up period. Furthermore, secondary aims were to investigate changes in isometric knee muscle strength over an eight-month follow-up period, and finally to report conversion to surgical reconstruction over a two-year follow-up period.

MATERIAL AND METHODS

STUDY DESIGN

This study was performed in accordance with the Code of Ethics of the World Medical Association and the Danish Code of Conduct for Research Integrity. All patients gave informed consent to participate. This study was notified to the Central Jutland Region Committee on Health Research Ethics (Case no. 1-10-72-1-19). The Danish Data Protection Agency permitted handling of personal data (1-16-02-549-13). Patients who declined to participate received similar treatment except for prospective registrations.

PARTICIPANTS

From June 2015 to January 2018, patients with a PCL injury were consecutively recruited from the Department of Orthopaedic Surgery at Aarhus University Hospital, Denmark. The inclusion criteria were an acute injury to the PCL, isolated or in combination with other knee ligament injuries presenting within eight weeks after injury. Exclusion criteria were fibular head fractures, avulsion fractures of the PCL, inability to comprehend, read, or speak Danish.

Scheduled assessments took place at 12- and 16-week follow-up visits and was a part of the routine treatment, while one- and two-year follow-up was scheduled explicitly for the study. Sport traumatology orthopaedic surgeons diagnosed all patients by magnetic resonance imaging in combination with a thorough physical examination.

PHYSIOTHERAPY-LED EXERCISE INTERVENTION

Over a period of 16 weeks, the included patients were offered seven physiotherapy-led training sessions: one time every two weeks. All sessions lasted 60 minutes and were conducted in the hospital with one-to-one supervision by physiotherapists with more than 15 years clinical experience in training of musculoskeletal disorders. At the first visit a support brace was fitted to the patients, the physiotherapists counselled the patients on how to wear the brace and the patients were instructed in exercises to do at home (Additional File 1). Exercise equipment (handheld dumbbells and a balance mat) was loaned to the patients to increase load and progress stability exercises. All exercises...
were based on recommendations from previously described exercise interventions\textsuperscript{24,25} and were individually selected and tailored to each patient.

Weight bearing was allowed from the start of the treatment with or without ambulating. In the first four weeks of the exercise intervention, it was important to respect tissue healing, joint response and individual adaptions, and to produce a minimum stimulus in this post injury recovery period to preserve muscle mass, but not overload the joint. Therefore, patients were recommended to perform one set of 12 repetitions three times a day and the starting level of the exercises (e.g., use of hand support, dumbbells or Thera-band elastic bands) was adapted to each patient’s pain level and ability to perform each exercise.

Exercises focused on effusion control, ROM, proprioception, muscle activation of the knee extensors and muscle strength exercises for the knee extensors, hip abductors and calf muscles. At week three a cross-trainer and a stationary exercise bike was recommended for warm up. Progressive Strength Exercises (PSE) were introduced gradually from week five with a starting load of three sets of 12 Repetition Maximum (RM). A 12 RM was chosen because it is considered a moderate load and has been recommended for hypertrophy training by the American College of Sports Medicine in patients who are novice to intermediate in resistance training.\textsuperscript{26} Patients were recommended to perform the PSE three times per week with a minimum of one restitution day between the sessions. The PSE training modality is documented in Additional File 1, using the strength training descriptors suggested by Toigo and Boutelle.\textsuperscript{26} PSE included leg extension (ROM 90-0°), heel raise, squat (ROM 0-90°), deadlift, modified leg press and hip abduction. The absolute training load (handheld dumbbells lifted or resistance applied with Thera-band elastic bands) was adjusted on a set-by-set basis for all muscular strength exercises. Patients with moderate swelling and acceptable pain (Numeric Rating Scale [NRS] < five) and who were motivated, were allowed to go to a fitness center from week four. This was to ensure that patients did not overload with the homebased exercises. They were given instructions from the physiotherapists on which exercises to do, and the exercises resembled the exercises from the PSE protocol e.g., leg extension with elastic bands was done in the leg extension machine and hip abduction was done in the pulley tower.

Criteria for load increase in each specific exercise was that patients could perform 1-2 extra repetitions over the desired number.\textsuperscript{27} This was a simple way for the patients to determine load on their own, making sure that load was progressed when indicated and not only at physiotherapy sessions.

In terms of regression and progression, three principles were applied\textsuperscript{28} (1) The patients could perform the exercises correctly (i.e. adequate hip, knee and foot alignment); (2) the patients could perform 12 repetitions; (3) patients experienced no knee joint pain above five out of 10 on a NRS for pain anchored by "no pain" (score of 0) and "unbearable pain" (score of 10) during and after the training session.\textsuperscript{29} The criteria were implemented to ensure that the applied training stimulus was not excessive and causing tissue overload. Patients were also told that the day after training, pain should subside to "pain as usual." If the pain did not subside, the level of training was reduced.\textsuperscript{29} The NRS for pain was chosen because it is a well-known scale and easy to administer for patients when rating pain. At the end of the 16-week period, the patients were encouraged to continue exercising at home or at a fitness centre and gradually return to usual activities. The patients aiming to return to knee-strenuous sport were advised not to return before they had recovered full range of knee motion and achieved a Leg Symmetry Index (LSI) of ≥ 90 % in knee extensor strength.\textsuperscript{2} At the end of the 16 weeks exercise intervention, these patients were recommended to seek supervision from a physiotherapist and complete an intense sport-specific rehabilitation programme including evaluation of return to sport readiness.

**SUPPORT BRACE INTERVENTION**

In addition to exercises, the patients wore one of two support braces for the first 12 weeks, either the PCL-Jack brace (Albrecht GmbH, Stephanskirchen, Germany) or the Rebound PCL brace (Össur Inc., Foothill Ranch, CA, USA). Both braces provided an anteriorly directed force to the posterior proximal tibia. Throughout the 12-week period, physiotherapists (BB, RGR) with in-depth knowledge of adapting and adjusting the brace performed individual brace sizing and fitting with the ROM set to 0-90° and the force loading as high as tolerable. In cases with injury of both cruciate ligaments, the load on the brace was reduced to avoid possible anterior subluxation of the knee.

**OUTCOME MEASURES**

Patient-reported outcomes were recorded at baseline (immediately after the diagnosis was confirmed), at one-year and at two-year follow-up. Isometric knee strength was assessed after 16 weeks and at one-year follow-up. Posterior translation of the tibia was measured by stress radiography at one-year follow-up, and by KT-1000 arthrometry and a posterior drawer test at three-month and one-year follow-up.

**PRIMARY OUTCOME**

The primary outcome was measured by the International Knee Documentation Committee Subjective Knee Form (IKDC-SKF).\textsuperscript{30,31} The IKDC-SKF is designed to measure symptoms and function in patients undergoing knee surgery or non-operative interventions.\textsuperscript{30,31} The Danish translation of the IKDC-SKF has demonstrated excellent test-retest reliability at group and individual level and adequate responsiveness with an intra class correlation of 0.94. Standard error of measurement of 2.6 points and a minimal clinically important change of 7.0 points.\textsuperscript{32}

The IKDC-SKF consists of 10 questions, which can be converted to a total score ranging from 0 to 100 points, where 100 points indicate the best possible outcome.
Tests were carried out by a physiotherapist (RGR) trained and experienced in measuring knee strength with the static strength dynamometer and tests were performed as a make test. The make test assesses the patient’s full maximum voluntary isometric contraction possible. However, in this study pull force was assessed using a load cell. This type of dynamometers has shown to be reliable with an ICC of 0.953, and MDC of 11.4 kilogram-force. The patient was seated with 90° of knee flexion. An external strap was placed around the ankle five cm proximal to the prominence of the lateral malleolus, and the load cell was fixed between the strap and a wire attached perpendicular to the wall. Isometric knee extension was measured first followed by measurement of isometric knee flexion. For practice, patients were instructed to exert one submaximal contraction into the strap. After practice, the patients exerted four maximum voluntary contractions against the strap with a break of 30 seconds between each contraction. The patients were instructed to push as hard as possible for five seconds against the strap. If the fourth contraction was higher than the previous, additional trials were performed, until no higher measurements were recorded. The best result of each test (knee extension and flexion) was recorded. The moment arm was calculated by measuring the distance from the knee joint lateral space to five cm proximal from the prominence of the lateral malleolar tip - the site of the external strap fixation including the load cell.

Strength values were normalized to moment arms and weight and reported as Nm/kg bodyweight. Finally, the limb symmetry index (LSI) was calculated as knee extensor and flexor strength performance of the involved limb/knee extensor and flexor strength of the uninvolved limb x 100%

The physiotherapist (RGR, BB) extracted information about conversion to surgery from the patients’ medical records.

OTHER OUTCOMES

Other outcomes included the Knee injury and Osteoarthritis Outcome Score (KOOS), PCL stress radiography, objective PCL laxity measured by KT-1000 arthrometry (instrumented drawer testing) and tibial offset using the posterior drawer test. The same radiographer specialized in the kneeling technique for posterior cruciate ligament stress radiography did all the PCL stress examinations. The KT-1000 and tibial offset examinations were performed by an experienced and trained physiotherapist (LM).

The KOOS is a valid, reliable and responsive outcome measure designed to measure symptoms, function and quality of life in patients with knee injuries and knee conditions following operative and non-operative interventions. The KOOS includes 42 questions divided into five separate subscales: Pain, symptoms, function of daily living, function in sport and recreation, and knee-related quality of life. Each sub score can be converted to a total score ranging from 0 to 100 points, where 100 points indicate the best possible outcome. The KOOS has a high test-retest reliability with an ICC across subscales ranging from
0.61 - 0.95 and the MDC across subscales ranges from 5-12 points.\textsuperscript{37}

Objective quantification of knee laxity was done by PCL stress radiography \textsuperscript{38,39} measuring total posterior displacement (Additional File 2). PCL stress radiography using the kneeling stress method provides a reproducible method to quantify posterior laxity in patients with PCL lesions with intra- and interobserver reliability ICC of 0.973 and 0.955 respectively.\textsuperscript{40,41} The degree of tibial displacement in millimeters for the injured and the non-injured knee was measured, and the mean side-to-side difference (SSD) was calculated.

The posterior tibial translation of the knee was additionally measured with the KT-1000 arthrometer (MEDmetric, San Diego, CA) using the method described by Daniel et al.\textsuperscript{42} The KT-1000 arthrometer has been shown to have an intra- and intertester reliability of 0.79 and 0.62 respectively and a SEM of ±2.53 in patients with a PCL injury.\textsuperscript{43} Two tests on each leg were performed; the first test at 30° of flexion and the second test at 70° of flexion. KT-1000 arthrometry was done with a 30 lb force, equivalent to 135 Newton. The mean SSD of the second test was calculated and used in the analysis.

Evaluation of tibial offset using the posterior drawer test\textsuperscript{39} was classified by the tibial grading system; In grade A injuries, the plateau remains anterior to the medial femoral condyle. In grade B injuries, the plateau is flush with the medial femoral condyle and in grade C injuries the plateau is displaced posteriorly to the medial femoral condyle. Evaluation of tibial off set using the posterior drawer test provides a 96% accuracy for detecting a posterior cruciate ligament tear, with a 90% sensitivity and a 99% specificity.\textsuperscript{44} Patient characteristics were recorded at baseline, including age, gender, height, weight, dominant leg (i.e., preferred leg to kick a ball), mechanisms of injury, date of injury and previous knee injuries.

**STATISTICAL ANALYSIS**

Continuous data were reported as means with standard deviations (SD) if normally distributed, otherwise reported as medians with interquartile ranges. Histograms and Q-Q plots were used to test for normality. Categorical data were reported as numbers and percentages. Changes from diagnosis to follow-up after 12 weeks, 16 weeks, one and two years were analyzed with a mixed effects model with patients as a random factor and time as a fixed factor. Model assumptions were based on inspection of plots of standardized residuals versus fitted values and Q-Q plots of the standardized residuals. The STATA 14.2 (StataCorp, College Station, TX, USA) software package was used for data analysis, and results were considered statistically significant if \( p < 0.05 \).

All data were reported as one group only (covering both isolated and combined PCL injuries) as the study population was considered representative for the isolated and combined PCL injury distribution at Aarhus University Hospital. However, outcomes divided into isolated PCL injury and multi-ligament injury were additionally calculated applying the same mixed effects model as aforementioned.

**SAMPLE SIZE CONSIDERATIONS**

The aim of this study was to investigate changes in the IKDC-SKF over a two-year follow-up period and changes in isometric knee muscle strength during eight-month follow-up. Annually, approximately 20 patients are treated non-surgically at Aarhus University Hospital. Based on a clinical and research judgment, a convenience sample size of 45 patients was considered appropriate to represent the target population and large enough to provide data on changes from before to two years after treatment. To consider dropout, this study aimed to recruit 50 patients.

**RESULTS**

In the study period, 50 patients were included out of 52 eligible patients (Figure 2).

In total, four patients (one isolated PCL injury) were lost for the one-year follow-up (baseline characteristics; three males, median age 38 (range 20-41), median BMI 29 (range

**Figure 2. Patient flow diagram**

One patient was unable to complete this specific follow-up due to knee pain

* Three patients did not complete muscle strength tests and PCL stress radiography due to knee pain and one patient did not return patient-reported outcome
Table 1. Baseline patient characteristics of 50 patients with acute PCL injury treated with a physiotherapy-led exercise and support brace intervention

<table>
<thead>
<tr>
<th>Patients with PCL injury</th>
<th>Gender, males, n (%)</th>
<th>Age, years*</th>
<th>BMI*</th>
<th>Diagnosis, n (%)</th>
<th>Isolated PCL injury</th>
<th>Multi-ligament injury</th>
<th>Mechanisms of injury, n (%)</th>
<th>Sport</th>
<th>Traffic</th>
<th>Daily activity</th>
<th>Work</th>
<th>Time in days from injury to initiation of treatment*</th>
</tr>
</thead>
<tbody>
<tr>
<td>N=45</td>
<td>37 (74)</td>
<td>33 (15-61)</td>
<td>27 (19-41)</td>
<td>28 (56)</td>
<td></td>
<td>227 (44)</td>
<td>37 (74)</td>
<td>6 (12)</td>
<td>5 (10)</td>
<td>2 (4)</td>
<td></td>
<td>23 (3-55)</td>
</tr>
</tbody>
</table>

21-31), mean IKDC 36 (range 22-53) and a total of 15 patients (11 isolated PCL injuries) were lost for the two-year follow-up (baseline characteristics; 14 males, median age 33 (range 17-47), median BMI 28 (range 19-37), mean IKDC 36 (range 22-53). The characteristics of the included patients are reported in Table 1.

Results divided into groups of isolated PCL injury and multi-ligament injury are available in the Additional File 3.

**PRIMARY OUTCOME**

The IKDC score increased statistically significantly from 35 IKDC points at baseline to 65 IKDC points two years after injury (p < 0.001) (Table 2 and Figure 3).

### SECON DARY OUTCOMES

**ISOMETRIC KNEE STRENGTH**

There was no change in isometric knee extension strength. In contrast, isometric knee flexion strength of the injured knee increased from 0.93 Nm/kg at 16 weeks to 1.1 Nm/kg after one year, corresponding to an increase of 16% (p < 0.001) (Table 3).

**KOOS SCORES**

All KOOS subscores increased statistically significantly, with the highest increase in the subscore function in sport and recreation.

**KNEE STABILITY MEASURES**

PCL stress radiography showed a mean of 6 mm tibial displacement in the injured knee whereas the healthy knee showed a mean of 10 mm tibial displacement at one-year follow-up. The mean SSD was 4 mm at one-year follow-up. For the KT-1000 arthrometer assessment, baseline posterior laxity in the injured knee was unchanged from baseline 9 mm (SD 3.5) to one-year follow-up 9 mm (SD 3.3) p=0.804, whereas the healthy knee was 6 mm (SD 2.7) at baseline and 5 mm (SD 2.5) at one-year follow-up and the overall mean SSD was 5 mm (range 0-14). Tibial offset was measured in 39 patients: Five patients improved from grade B to grade A, one patient improved from grade C to B, one patient improved from grade A to grade 0, and no changes were observed in 24 patients (62%). Finally, eight patients worsened from grade A to grade B or C.

**CONVERSION TO SURGERY**

During the study period, seven patients converted to PCL surgical reconstruction. Out of these patients, two patients had an isolated PCL injury and five patients had knee dislocation injuries. Injuries at baseline and conversion to surgery are reported in Table 4. Median time from initiation of non-operative intervention to surgery was 13 months (range 10-14).

Table 2. Mean changes in patient-reported outcomes from baseline to two-year follow-up in patients with acute PCL injury treated with a physiotherapy-led exercise and support brace intervention

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Baseline</th>
<th>1 year</th>
<th>2 years</th>
<th>Change baseline vs. 1 year</th>
<th>Change baseline vs. 2 years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=50</td>
<td>N=45</td>
<td>N=31</td>
<td>Mean (95% CI) p-value</td>
<td>Mean (95% CI) p-value</td>
</tr>
<tr>
<td>IKDC</td>
<td>35 (9.7)</td>
<td>61 (13)</td>
<td>65 (13)</td>
<td>26 (22-30) &lt; 0.001</td>
<td>28 (24-33) &lt; 0.001</td>
</tr>
<tr>
<td>KOOS pain</td>
<td>56 (24)</td>
<td>79 (17)</td>
<td>87 (16)</td>
<td>23 (17-29) &lt; 0.001</td>
<td>25 (18-32) &lt; 0.001</td>
</tr>
<tr>
<td>KOOS symp</td>
<td>52 (20)</td>
<td>78 (17)</td>
<td>84 (15)</td>
<td>25 (20-31) &lt; 0.001</td>
<td>29 (22-36) &lt; 0.001</td>
</tr>
<tr>
<td>KOOS ADL</td>
<td>58 (21)</td>
<td>84 (16)</td>
<td>90 (15)</td>
<td>26 (21-32) &lt; 0.001</td>
<td>26 (20-32) &lt; 0.001</td>
</tr>
<tr>
<td>KOOS sport/rec</td>
<td>17 (22)</td>
<td>58 (28)</td>
<td>71 (26)</td>
<td>42 (34-50) &lt; 0.001</td>
<td>48 (38-58) &lt; 0.001</td>
</tr>
<tr>
<td>KOOS QOL</td>
<td>23 (16)</td>
<td>56 (23)</td>
<td>70 (25)</td>
<td>34 (27-40) &lt; 0.001</td>
<td>40 (31-49) &lt; 0.001</td>
</tr>
</tbody>
</table>

Data are provided as mean (SD) with mean changes (95% CI) from baseline to one- and two-year follow-up. Abbreviations: CI: confidence interval; SD: standard deviation; IKDC: International Knee Documentation Committee Subjective Knee Form; KOOS: Knee injury and Osteoarthritis Outcome Score; symp: symptoms; ADL: function in daily living; sport/rec: function in sport and recreation; QOL: quality of life.
The primary findings of the present study were that patients with acute isolated and multi-ligament PCL injuries treated with the physiotherapy-led exercise and support brace intervention reported clinically relevant improvements from baseline to two-year follow-up, isometric knee flexion strength improved clinically relevant from 16 weeks after injury to one-year follow-up and only 15% of patients needed conversion to surgical reconstruction.

Several previous studies have reported patient-reported outcome after non-operative intervention for PCL injury. However, no studies have investigated changes from injury to one- and two-year follow-up. Compared to previous studies on non-operative treatment, the scores of this study are relatively low. Two studies by Shelbourne et al.8,15 reported IKDC-SKF scores of 83 points in 85 patients at nine-year follow-up and 73 points in 68 patients at 18-year follow-up in patients with an isolated PCL injury. Patel et al.6 reported an IKDC-SKF score of 84 points in 57 patients with an isolated PCL injury seven years after injury. Jacobi et al.7 reported an IKDC-SKF score of 95 points two years after injury in 17 patients with an isolated PCL injury. The lower patient-reported outcome score in the present study can probably be explained by a shorter follow-up and since this

DISCUSSION

Figure 3. IKDC and KOOS scores at baseline and one- and two-year follow-up

IKDC: The International Knee Documentation Committee; KOOS: The Knee injury and Osteoarthritis Outcome Score; Symp: Symptoms; ADL: Activities of Daily Living; Rec: Recreational; QoL: Quality of Life

Table 3. Mean changes in isometric knee strength from 16 weeks to one year after injury in patients with acute PCL injury treated with a physiotherapy-led exercise and support brace intervention

<table>
<thead>
<tr>
<th>Outcome</th>
<th>16 weeks (n = 48†)</th>
<th>1 year (n = 43‡)</th>
<th>Change 16 weeks vs. 1 year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Injured knee</td>
<td>Healthy knee</td>
<td>Injured knee</td>
</tr>
<tr>
<td>Knee ext, Nm/kg</td>
<td>1.6 (0.59)</td>
<td>1.8 (0.47)</td>
<td>1.7 (0.56)</td>
</tr>
<tr>
<td>Knee ext, LSI %</td>
<td>87 (23)</td>
<td>91 (72)</td>
<td></td>
</tr>
<tr>
<td>Knee flex, Nm/kg</td>
<td>0.93 (0.36)</td>
<td>1.2 (0.45)</td>
<td>1.10 (0.36)</td>
</tr>
<tr>
<td>Knee flex, LSI %</td>
<td>71 (21)</td>
<td>82 (22)</td>
<td></td>
</tr>
</tbody>
</table>

Data are provided as mean (SD) with mean changes (95% CI) from 16 weeks follow-up to 1-year follow-up. *One patient was unable to complete due to knee pain †Three patients were unable to complete due to knee pain. Abbreviations: CI: confidence interval; SD: standard deviation; ext: extension; flex: flexion; Nm/kg: Newton meter/kilogram. LSI: leg symmetry index.
study included 22 patients with multi-ligament injuries, who were more severely injured compared to the patients in the previous studies with isolated PCL injuries,\(^5,8,15\) as shown in Additional File 3. Based on one study by Shelbourne et al.\(^5\) reporting subjective improvement in 23% of patients after two years, the authors expect the patients with an isolated PCL injury to improve to a level similar to the follow-up score reported in the previous studies.\(^7,8,15,23\) (also indicated in Additional File 3). In the Additional File 3, the authors report baseline patient characteristics and results divided into two groups e.g., isolated and multi-ligament injury. These data indicate that patients with multi-ligament injury report lower baseline scores compared to the patients with isolated injuries. Nonetheless, at one- and two-years follow-up the scores of the two groups seem to be similar. However, no formal statistical tests for group differences were done.

Knee extensor strength is associated with patient-reported outcomes following a PCL injury.\(^45,46\) In this study, knee extensor strength in the injured knee did not change from 16 weeks to one-year follow-up, but the isometric knee flexor strength improved by 16%. Previous studies on knee extensor strength following non-operative intervention showed that the majority of patients treated non-operatively regained almost normal knee extensor strength after isolated PCL injury.\(^5,12,15,23\) One study by Shelbourne et al.\(^5\) reported an LSI of 97% in mean knee extensor strength and 93% in mean knee flexor strength in 44 patients 14 years after isolated PCL injury. In another study by Mygind-Klavsøen et al.,\(^47\) only minor differences in knee extensor strength between 77 patients with isolated PCL injuries compared with 119 patients with combined PCL injuries were reported six years after PCL reconstruction, and both groups were classified as normal or nearly normal. Follow-up strength outcomes were slightly lower in the present study compared to the previous studies, and this can be explained due to several factors. First, this study included patients with isolated PCL and multi-ligament injuries, typically experiencing more residual knee instability and symptoms negatively affecting rehabilitation. Second, as opposed to the study by Shelbourne et al.\(^5\) only a few of the patients in the present study were semi-professional athletes, possibly impacting negatively on rehabilitation conditions due to lower motivation and physical functioning. Third, the start of progressive strength exercises was delayed as the mean time from injury to start of treatment was 25 days. Fourth, different muscle strength tests were used making comparison difficult. Fifth, in the present study knee strength was measured 13 years earlier than reported in the previous studies. Finally, all patients underwent prospective scheduled follow-ups. On the contrary, the previous studies were typically based on retrospective data, possibly negatively influencing risk of information and selection bias.

Currently, non-operative intervention is considered first-line treatment for patients with an isolated PCL injury followed by PCL reconstruction if needed.\(^48\) The results in the present study support this treatment strategy, showing little need for PCL reconstruction (9%) in case of isolated PCL injury. This is supported by a previous study,\(^6\) reporting that two out of 58 patients required PCL reconstruction after non-operative intervention in patients with an isolated PCL injury. In recent years, non-operative interventions have played an increasing role with the use of support braces to ensure optimal healing of the PCL. In this study, both progressive exercises and a support brace were applied. However, the residual PCL laxity may indicate that the brace treatment failed to ensure sufficient anatomical PCL healing to a degree where the PCL is fully stable. Nevertheless, no previous clinical studies have described the non-operative intervention as thoroughly as in this study.

The authors acknowledge that there are several limitations to the present study. First, the sample size was based on a convenience sample, and therefore lack of statistically significant changes because of to low power may exist. Moreover, the authors expected a dropout of five out of 50 patients. However, at the two-year follow-up, only 31 patients completed the outcomes, and therefore, the risk of too low power is even higher for this time point. In general, PCL injuries are relatively rare resulting in a low sample

### Table 4. Ligament injuries at baseline and conversion to surgery in 50 patients with acute posterior cruciate ligament injury treated with a physiotherapy-led exercise and support brace intervention

<table>
<thead>
<tr>
<th>Ligament injury classification</th>
<th>Injuries at baseline based on MRI scan N=50</th>
<th>Conversion to surgery N=8*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isolated PCL injury</td>
<td>28</td>
<td>2</td>
</tr>
<tr>
<td>Schenck KD-I</td>
<td>6 PCL+MCL</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>1 PCL+LCL+PLC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 PCL+MCL+LCL</td>
<td></td>
</tr>
<tr>
<td>Schenck KD-II</td>
<td>1 ACL+PCL</td>
<td>0</td>
</tr>
<tr>
<td>Schenck KD-III Medial</td>
<td>9 ACL+PCL+MCL</td>
<td>1 PCL+ACL+MCL</td>
</tr>
<tr>
<td></td>
<td>1 ACL+PCL+MCL+LCL</td>
<td>1 ACL+MCL</td>
</tr>
<tr>
<td>Schenck KD-IV</td>
<td>3 ACL+PCL+MCL+LCL</td>
<td>1 ACL+PCL+LCL+PLC</td>
</tr>
<tr>
<td></td>
<td>1 ACL+PCL+MCL+LCL+PLC</td>
<td>1 ACL+MCL+LCL+PLC</td>
</tr>
</tbody>
</table>

Abbreviations: ACL: anterior cruciate ligament; LCL: lateral collateral ligament; MCL: medial collateral ligament; PCL: posterior cruciate ligament; PLC: posterior lateral corner; KD: knee dislocation; MRI: magnetic resonance imaging. *One did not include PCL.
size in most studies on the management and outcomes of PCL injuries. Nevertheless, the authors still believe that this study adds valuable knowledge about non-operative treatment intervention of this patient group. The patients lost to follow-up were comparable to patients who completed follow-ups regarding baseline characteristics and thus, the loss to follow-up could result in either under- or overestimation of the results. Second, knee muscle strength and PCL laxity were not measured at baseline for ethical reasons (risk of injuring a healing ligament). Third, the study population was heterogeneous covering patients with both isolated PCL injuries and multi-ligament injuries. However, dividing patients into two groups, would also be problematic, because of the consequence of additional concerns for low power. The heterogeneous group causes a large variety of data, meaning that results may be underestimated if compared to patients with isolated PCL injuries whereas changes would be overestimated if results are compared to patients with multi-ligament injuries. This study consecutively included all patients who met the inclusion criteria and thus, the study population represents the distribution of isolated and multi-ligament PCL injuries at Aarhus University Hospital. Fourth, the physiotherapists who tested the patients clinically were not blinded, and therefore, the authors cannot rule out that lack of blinding may have an impact on the results. Fifth, the results were not compared to results of a control group and no training or health technologies were used to document rehabilitation adherence. Consequently, the authors do not know if the changes occurred due to time or if the changes were because of the intervention. However, offering no treatment to a control group would not be ethically acceptable. Sixth, a delay in treatment initiation of up to eight weeks in some patients may have resulted in worse outcomes. Nevertheless, delay in treatment initiation at our institution do occur, and therefore, the results in this study will describe what patients and clinicians should expect following treatment.

CONCLUSION

Patients with an acute PCL injury treated with the physiotherapy-led exercise and support brace intervention can expect clinically relevant improvements in patient-reported outcomes and knee flexion strength, and the risk of PCL reconstruction is considered low within the first two years. However, further studies are needed to establish the effect of different exercise and support brace interventions.

CONFLICT OF INTEREST/COMPETING INTERESTS

The authors declare that they have no financial or non-financial interests to disclose.

Submitted: October 28, 2022 CDT, Accepted: May 16, 2023 CDT

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-NC-4.0). View this license's legal deed at https://creativecommons.org/licenses/by-nc/4.0 and legal code at https://creativecommons.org/licenses/by-nc/4.0/legalcode for more information.
REFERENCES


---

*International Journal of Sports Physical Therapy*
SUPPLEMENTARY MATERIALS

Additional File 1

Additional File 2

Additional File 3
Original Research

Thoracolumbar and Lumbopelvic Spinal Alignment During the Barbell Back Squat: A Comparison Between Men and Women

Victor Bengtsson¹, Lars Berglund², Fredrik Öhberg³, Ulrika Aasa³

¹ Department of Women's and Children's Health, Uppsala university, ² Community medicine and rehabilitation, Umeå university, ³ Department of Radiation Sciences, Umeå university

Keywords: Inertial sensors, Lifting technique, Posture, Powerlifting, Weightlifting

https://doi.org/10.26603/001c.83942

International Journal of Sports Physical Therapy

Background

Maintaining neutral spinal alignment is considered important when performing the barbell back squat exercise. Since male and female lifters may differ in injury location it is important to examine whether they differ in spinal alignment during the back squat.

Objectives

The study aimed to quantify the spinal alignment in the upper and lower lumbar spine during the barbell back squat exercise in male and female lifters. Secondary aims were to compare alignment during the back squat to standing habitual lumbar spine alignment and determine whether male and female lifters differ in these aspects.

Study Design

Observational, Cross-sectional.

Methods

Competitive power- and weightlifters were recruited and performed three repetitions of the barbell back squat exercise using a load equivalent to 70% of their one-repetition maximum. Spinal alignment and range of motion were measured using inertial measurement units placed on the thoracic, lumbar and sacral spine. Data was presented descriptively and comparisons between men and women as well as spinal alignment in four different positions were done with a factorial repeated measures analysis of variance.

Results

Twenty-three (14 males, 9 females) were included. During execution of the squat, spinal alignment adjustments in the lumbar spine were made in all three planes of movement, compared to the start position, in both male and female lifters. Compared to their standing habitual posture, all lifters adjusted their upper lumbar spine to a less lordotic position when in the start position of the back squat (standing upright with the barbell on their back). Only male lifters assumed a less lordotic alignment in their lower lumbar spine in the start position compared to their habitual posture.

Conclusions

Adjustments of spinal alignment, predominantly in the sagittal plane, are made during execution of the back squat in both male and female lifters. Further, lifters adopt a less lordotic alignment with a heavy barbell on their upper back, more so in male than female
INTRODUCTION

The barbell back squat is considered a safe strength training exercise, provided it is performed and progressed in a controlled manner. However, the squat appears to be associated with injuries in people participating in powerlifting, weightlifting, bodybuilding, and strongman sports. In Sweden, 70% of sub-elite powerlifters reported a current injury, and 87% had experienced an injury within the prior 12 months. The lumbar region was one of the most frequently injured areas for both sexes, although men appeared to have a slightly higher frequency of lumbo-pelvic injuries, 42%, compared to 23% for women. Also notable was the difference between men and women regarding injuries to other parts of the spine where women had a significantly higher frequency of neck (20%) and thoracic injuries (29%) than men which reported zero (0%) injuries to these parts of the spine.

Numerous authors have discussed the importance of optimal technique for prevention of injuries during the barbell back squat. With respect to the spine, a correct technique is described as maintaining the spine in an upright position with preservation of its neutral posture while the trunk is held stable without any observation of waverings or displacement in all planes of movement, throughout the entire squat movement. In a neutral posture, where the spine usually has a gentle “S” shape where the lower back has a slight lordotic curve and the thoracic spine a slight kyphotic curve, it is said that axial load is distributed in a balanced way in all movement segments. It seems, however, that there is an ongoing discussion about what constitutes correct squat technique for different parts of the body, though robust evidence in the matter is somewhat lacking. Nevertheless, experts have stated that improper technique in combination with heavy loads may cause back injuries while squatting and that flexion (rounding), twisting (rotating), and side bending of the spine are movements that should be avoided to reduce risk of injury risk and improve performance. In support of this, previous studies have reported that the ability to maintain the lumbar spine in a neutral position modifies the forces exerted on the body structures and that an inability to maintain a neutral position increases the potential to overload spine and soft tissues to the point of injury, especially when repeated over time.

Previous authors have described the kinetics and kinematics of the squat using video analysis, magnetic tracking devices, or motion capture systems. However, a growing body of scientific literature has investigated the utility of inertial measurement units (IMUs) to assess exercise technique and for monitoring of the spinal alignment during resistance training. Inertial measurement units have been validated to electromagnetic based system for measuring 3D spinal ranges of movement and spinal coupled motion measurement and are reliable for measuring joint angles during physical activities. Further, IMU systems have been shown to provide data that can distinguish between acceptable and aberrant squat techniques and have been shown to be able to identify technique deviations.

Results from previous studies have shown that the spinal alignment is adjusted during the performance of the squat and that the adjustments may be influenced by sex, load, lifting phase (i.e. concentric or eccentric phase) and stance width. However, there are several more variables that might influence spinal alignment, for example, one is how proficient the lifter is in the specific movement and so far, no study has evaluated the kinematics in a group of experienced lifters.

Considering that the lumbo-pelvic region appears to be one of the most injured anatomical locations among powerlifters and weightlifters and that men and women who powerlift appear to have different injury rates to different part of the spine, it was hypothesized that men and women move, and thereby load, their spine differently when performing the barbell back squat. Therefore, the aims of the present study were to quantify the spinal alignment in the upper (thoracolumbar, T11-L2) and lower (lumbo-pelvic, L2-S2) lumbar spine during the barbell back squat exercise in male and female lifters. Secondary aims were to compare alignment during the barbell back squat to standing habitual lumbar spine alignment and determine whether male and female lifters differ in these aspects.

METHODS

To quantify lumbar spine alignment IMUs where used to measure three-dimensional angles of the upper lumbar spine (i.e. thoracolumbar) and lower lumbar spine (i.e. lumbo-pelvic). Spinal alignment was measured in habitual standing and during performance of three repetitions of the barbell back squat with a high load (70% of self-estimated one repetition maximum (RM)).

PARTICIPANTS

Powerlifters and weightlifters were recruited through invitations to powerlifting and weightlifting clubs, respectively, in Umeå, Sweden. Only lifters with ≥2 years of strength training experience, without present injuries that could affect squat performance, and with the intent of competing in powerlifting or weightlifting were included. Also, to avoid the risk of IMUs touching each other during the data collection, only lifters whose height was >150 cm were
included. In order to ensure that eligibility criteria was met, all participants completed a questionnaire. No participant had any recent or previous medical issues which prevented them from participating in the study. They also signed written informed consent form prior to participation. The study was approved by the Regional Ethical Review Board of Umeå, Sweden (Dnr 2014-285-31M).

PROCEDURES

At the day of data collection, participants first answered a questionnaire and thereafter they completed a self-administered warm-up typically consisting of squats with an unloaded barbell progressing to heavy squats. Thereafter, three calibrated IMUs (MPU-9150, InvenSense, San Jose, USA) were affixed with double-sided tape and elastic self-adhesive bandage wraps to their back by the test leader at processus spinous T11 and L2, and Sacrum (S2). Thereafter, the lifters completed one further set of warm-up bodyweight squats before data collection while the test leader ensured that the IMUs were set firmly and did not hinder squat performance. The lifters, who were wearing their preferred shoes, were then instructed to assume their habitual standing posture with their arms at their sides while looking straight ahead. Habitual spinal posture (habitual posture) was measured using the IMUs in this position. Thereafter, spinal alignment was registered during the squat. The lifters were asked to perform one set of three repetitions at 70% 1RM. Weight plates of official measures were attached to each end of a powerlifting barbell and the weight was adjusted to the nearest 2.5 kg. After receiving a start signal, the lifters were instructed to descend by flexing at the hip, knee and ankle joints until the crease of the hip was lower than the top of the knee. From the bottom position, the lifters ascended to the start position by extending the same joints. They stayed in this erect position and waited for a new signal before beginning the descent of the next repetition. No equipment other than wrist wraps was allowed.

INSTRUMENTS AND MEASUREMENTS

The spinal alignment (degrees) was measured during habitual posture in standing and during execution of the squat exercise. The IMUs recorded their position in all three planes of movement relative to each other and thus measured the three-dimensional angles of the upper lumbar spine (i.e. thoracolumbar spine, T11-L2) and lower lumbar spine (i.e. lumbopelvic spine, L2-S2). A positive sagittal plane value indicated a lordotic spinal alignment and negative sagittal plane value indicated a kyphotic spinal alignment. A positive value in the frontal and horizontal plane indicated a right lateral flexion or rotation, respectively. A negative value in the frontal and horizontal plane indicated a left lateral flexion or rotation, respectively.

During the squat, four measures were selected to quantify spinal alignment: First, start position (the lifters standing erect with the barbell held across the back), second, min angle (the minimum angle, at any time during the squat, of the respective lumbar region in degrees), third, max angle (the maximum angle, at any time during the squat, of the respective lumbar region in degrees), and fourth, range of motion (ROM) between the minimum and maximum angles of the respective lumbar regions. The mean value of the three repetitions were used for data analysis for all variables.

The IMUs have a size of length 60 x width 45 x height 10 mm and weigh 14 g each and the IMUs communicated with a laptop via WiFi and had a sampling frequency of 100 Hz, a 16-bit resolution and an anti-aliasing low pass filter set at 50 Hz. The full-scale range was ±1000 °/s for the gyroscopes, ±8 g for the accelerometers and ±4800 μT for the magnetometers. Using three axis gyro and three axis accelerometers, the IMUs detected three-dimensional spinal alignment and real-time orientation was calculated using a customised system MoLabTM POSE (AnyMo AB, Umeå, Sweden).

DATA HANDLING AND STATISTICAL ANALYSES

Orientation data (i.e., segment angles) from the IMUs were processed in Matlab (version 7.10.0 (R2010a), The MathWorks, Inc., USA). The Euler sequence used for the segment angles were X (rotations in the sagittal plane), Y (rotations in the frontal plane), and Z (rotations in the transverse plane). To each segment, a caudally and adjacent segment was selected as reference in the calculation of joint angles. All orientation data was low-pass filtered offline with a second order Butterworth filter at a cut-off frequency of 10 Hz. The filter coefficient was set to 10 Hz as human gross motion seldom contains frequencies above 10 Hz. A more detailed description of the used algorithms can be found in Ohberg et al. Inter- and intra-tester reliability has been estimated in a previous study showing higher intra-tester reliability.

Statistical analysis was performed using Statistical Package for the Social Sciences (SPSS) version 23 (IBM Corp., Armonk, NY, USA). A factorial repeated measures analysis of variance (mixed ANOVA) was conducted to compare the influence of the independent variables (group: 1=men and 2=women) and the effect of the dependent variable (segment angle at four different positions (1=habitual posture, 2=start position, 3=min min angle at any timepoint, and 4=max angle at any timepoint)) using the mean values for the three repetitions. Sphericity was calculated using Mauchly’s test of Sphericity. If sphericity was not assumed, a correction was made using the Greenhouse-Geisser estimation. If significant position x group effects were found, the results were also presented separately for male and female lifters. If significant within-subjects effects were found, post-hoc pairwise comparisons were calculated. Partial eta squared ($\eta_p^2$) was calculated for effect size, using 0.01, 0.06 and 0.14 to indicate small, medium and large effects respectively. Significance level was set at 0.05 and Bonferroni corrections were performed for multiple comparisons.
RESULTS

Background characteristics for participants are presented in Table 1. Comparisons between male and female lifters showed that male lifters were significantly taller and heavier, had more resistance training experience and were stronger (Independent samples t-test, p=0.05) than the female lifters. There was no significant difference between male and female lifters for age. The lifters (n=23) reported current injuries (i.e., pain and impaired ability to perform the squat or deadlift exercises) of the lumbar region (n=2), hip (n=5), knee (n=2) and shoulder (n=2).

The spinal alignment of the upper lumbar spine during standing habitual posture, and during the squats for the start position, minimum and maximum angle, and range of motion are presented in Table 2. For the upper lumbar spine, there were no differences between male and female lifters in spinal alignment (group x position interaction in the sagittal plane (F[1.6, 32.8] = 1.2, p = 0.309), frontal plane (F[1.7, 35.4] = 1.61, p = 0.217), or horizontal plane (F[1.9, 39.3] = 0.14, p = 0.860). In all participants their sagittal plane spinal alignment in standing habitual posture differed from the alignment at the start position and spinal adjustments were made in all three movement planes during the squat (significant main effect for position in the sagittal plane (F[1.6, 32.8] = 51.6, p < 0.001, η²p = 0.711), frontal plane (F[1.7, 35.4] = 21.27, p < 0.001, η²p = 0.503), and horizontal plane (F[1.9, 39.5] = 22.10, p < 0.001, η²p = 0.513). All three statistically significant comparisons for the upper lumbar spine had large effect sizes (η²p > 0.14).

For the lower lumbar spine (Table 3), only the male lifters decreased their lumbar lordosis during the start position compared to during standing habitual posture (group x position interaction in sagittal plane spinal alignment (F[1.9, 38.9] = 8.59, p = 0.001, η²p = 0.290)). There were no differences between men and women in spinal alignment in the frontal (F[1.6, 34.5] = 1.16, p = 0.516) or horizontal planes (F[1.7, 34.7] = 0.45, p = 0.605) spinal alignment. In all participants their sagittal plane spinal alignment in standing habitual posture differed from the alignment at the start position and spinal adjustments were made in all three movement planes during the squat (significant main effect for position in the sagittal plane (F[1.9, 38.9] = 131.52, p < 0.001, η²p = 0.862), frontal plane (F[1.6, 34.5] = 19.15, p < 0.001, η²p = 0.477), and horizontal plane (F[1.7, 34.7] = 50.92, p < 0.001, η²p = 0.596). All three statistically significant comparisons for the lower lumbar spine had large effect sizes (η²p > 0.14).

The factorial repeated measures ANOVA simple effects for position in upper and lower lumbar sagittal plane spinal alignment are presented in Figures 1 and 2.

The full kinematic tracings of the angle between the IMU at L2 and axis of gravity, in the sagittal plane, for the full sample (n=25), during the squat is visualized in Figure 3.

DISCUSSION

This is the first study to investigate spinal alignment during the barbell back squat in experienced male and female power- and weightlifters. During the squat, spinal alignment was adjusted in all three planes of movement, especially in the sagittal plane. Lordosis in the upper lumbar spine (T11-L2) decreased among all participants from habitual posture to start position. In the lower lumbar spine (L2-S2) lordosis also decreased, but only among the male lifters. The difference in spinal alignment are consistent with the results of a study by McKeon et al.,15 showing that with both a wide and narrow stance width the lumbar spine adjusts to a less lordotic alignment when a loaded barbell is placed on the shoulders.15 It is reasonable to hypothesize that the participants adjusted their spinal alignment in order to cope with the load and keep the combined body and barbell center of gravity within the base of support.15 However, the reason why male but not female lifters decreased their lordosis in the lower lumbar spine in their habitual posture compared to the start position is somewhat unclear. One possible explanation could be the female lifters, on average, used a load of 125% of their bodyweight and male lifters 192% of their bodyweight, and thereby not creating the same need to adjust their spinal alignment in order to keep their center of gravity within their base of support. Also, whether these differences between men and women in spinal kinematics could account for differences in injury localizations are unknown.

The results showing that lifters decrease their lumbar lordosis during the squat, is consistent with previous studies.13,15,16,29,30 Hebling Campos et al.29 found that the lumbar curvature is more flexed in the deepest position of a squat compared to the habitual standing. Further, the lumbar flexion is more evident when restricting the anterior translation of the knees.13,29 The less lordotic spinal alignment could, at least in part, be explained by the inherent hip flexion performed during the squat. Powerlifters in particular tend to perform a so called powerlifting style squat by "sitting back" into the squat. A greater forward lean during lifting tasks increases the likelihood of a less lordotic spinal alignment31 and it has further been argued that it is easier to maintain lumbar lordosis with an upright trunk

### Table 1. Participant characteristics (mean ± SD).

<table>
<thead>
<tr>
<th>Participants</th>
<th>Age (y)</th>
<th>Weight (kg)</th>
<th>Height (cm)</th>
<th>Experience (y)*</th>
<th>Squat 1RM (kg)†</th>
</tr>
</thead>
<tbody>
<tr>
<td>All (n=23)</td>
<td>25.5 ± 5.5</td>
<td>80.4 ± 11.5</td>
<td>171.5 ± 7.2</td>
<td>7.9 ± 6.3</td>
<td>134.5 ± 42.1</td>
</tr>
<tr>
<td>Men (n=14)</td>
<td>26.7 ± 6.3</td>
<td>85.2 ± 10.8</td>
<td>174.9 ± 5.3</td>
<td>9.8 ± 7.5</td>
<td>162.9 ± 25.8</td>
</tr>
<tr>
<td>Women (n=9)</td>
<td>23.3 ± 3.0</td>
<td>72.9 ± 8.5</td>
<td>166.1 ± 9.8</td>
<td>5.0 ± 2.0</td>
<td>90.5 ± 14.1</td>
</tr>
</tbody>
</table>

*Strength training experience; †Self-estimated squat 1 repetition maximum.
Table 2. Three-dimensional angles of the upper lumbar spine (thoracolumbar region) during standing Habitual posture, and during the squat for the Start position, Minimum (Min) angle, Maximum (Max) angle and range of motion (ROM) in degrees (°) as well as results of the factorial repeated measures ANOVA (within-subjects effect).

<table>
<thead>
<tr>
<th>Measure</th>
<th>Habitual posture (°)</th>
<th>Start position (°)</th>
<th>Min angle (°)</th>
<th>Max angle (°)</th>
<th>ROM (°)</th>
<th>Within-subjects effect Time*group</th>
<th>p</th>
<th>Partial Eta Squared</th>
<th>Within-subjects effect Time</th>
<th>p</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>All (n=23)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sagittal plane</td>
<td>17.5 ± 12.5</td>
<td>7.8 ± 9.7†</td>
<td>4.0 ± 7.7†‡</td>
<td>13.7 ± 9.5†‡</td>
<td>9.7 ± 3.1</td>
<td>0.309</td>
<td>0.053</td>
<td>&lt;0.001</td>
<td>0.711</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frontal plane</td>
<td>0.9 ± 2.4</td>
<td>0.3 ± 3.1</td>
<td>-1.9 ± 4.2†‡</td>
<td>1.7 ± 3.3‡</td>
<td>3.6 ± 2.0</td>
<td>0.217</td>
<td>0.071</td>
<td>&lt;0.001</td>
<td>0.503</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horizontal plane</td>
<td>-0.4 ± 0.8</td>
<td>-0.3 ± 2.8</td>
<td>-2.2 ± 2.6†‡</td>
<td>1.5 ± 2.2†‡</td>
<td>3.7 ± 1.5</td>
<td>0.860</td>
<td>0.006</td>
<td>&lt;0.001</td>
<td>0.513</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men (n=14)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sagittal plane</td>
<td>12.1 ± 6.1</td>
<td>2.7 ± 6.0</td>
<td>0.3 ± 4.6</td>
<td>8.8 ± 5.3</td>
<td>8.5 ± 2.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frontal plane</td>
<td>0.4 ± 2.4</td>
<td>-0.7 ± 3.0</td>
<td>-3.2 ± 4.4</td>
<td>0.4 ± 3.1</td>
<td>3.6 ± 2.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horizontal plane</td>
<td>-0.1 ± 0.3</td>
<td>-0.3 ± 3.2</td>
<td>-2.1 ± 2.8</td>
<td>1.7 ± 2.1</td>
<td>3.8 ± 1.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women (n=9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sagittal plane</td>
<td>25.8 ± 15.6</td>
<td>15.8 ± 9.2</td>
<td>9.8 ± 8.2</td>
<td>21.4 ± 9.6</td>
<td>11.6 ± 2.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frontal plane</td>
<td>1.8 ± 2.4</td>
<td>1.8 ± 2.6</td>
<td>0.0 ± 3.0</td>
<td>3.7 ± 2.7</td>
<td>3.6 ± 1.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horizontal plane</td>
<td>-0.4 ± 0.8</td>
<td>-0.3 ± 2.8</td>
<td>-2.2 ± 2.6</td>
<td>1.3 ± 2.6</td>
<td>3.6 ± 1.3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*A positive sagittal plane angle indicated a lordotic spinal alignment and negative sagittal plane angle indicated a kyphotic spinal alignment. A positive value in the frontal and horizontal plane indicated a right lateral flexion or rotation, respectively. A negative value in the frontal and horizontal plane indicated a left lateral flexion or rotation, respectively.

†Significant difference to Habitual posture after adjustment for multiple comparisons using the Bonferroni correction.

‡Significant difference to Start position after adjustment for multiple comparisons using the Bonferroni correction.

International Journal of Sports Physical Therapy
<table>
<thead>
<tr>
<th>Measure</th>
<th>Habitual posture (°)</th>
<th>Start position (°)</th>
<th>Min angle (°)</th>
<th>Max angle (°)</th>
<th>ROM (°)</th>
<th>Within-subjects effect Time*group P</th>
<th>Partial Eta Squared</th>
<th>Within-subjects effect Time P</th>
<th>Partial Eta Squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>All (n=23)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sagittal plane</td>
<td>16.5 ± 10.7</td>
<td>12.2 ± 8.9†</td>
<td>-4.5 ± 8.0†‡</td>
<td>13.6 ± 8.4</td>
<td>18.1 ± 4.7</td>
<td>0.001</td>
<td>0.290</td>
<td>&lt;0.001</td>
<td>0.862</td>
</tr>
<tr>
<td>Frontal plane</td>
<td>-0.3 ± 3.5</td>
<td>0.2 ± 4.8</td>
<td>-1.8 ± 4.7†‡</td>
<td>1.3 ± 4.8‡</td>
<td>3.1 ± 1.1</td>
<td>0.316</td>
<td>0.052</td>
<td>&lt;0.001</td>
<td>0.477</td>
</tr>
<tr>
<td>Horizontal plane</td>
<td>0.1 ± 0.7</td>
<td>-0.8 ± 2.3</td>
<td>-2.6 ± 2.4†‡</td>
<td>0.8 ± 2.0†</td>
<td>3.5 ± 1.3</td>
<td>0.605</td>
<td>0.021</td>
<td>&lt;0.001</td>
<td>0.596</td>
</tr>
<tr>
<td>Men (n=14)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sagittal plane</td>
<td>17.9 ± 7.7</td>
<td>9.9 ± 6.0†</td>
<td>-7.1 ± 5.6†‡</td>
<td>11.5 ± 5.0†</td>
<td>18.6 ± 3.3</td>
<td>&lt;0.001</td>
<td>0.933</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frontal plane</td>
<td>-0.3 ± 3.5</td>
<td>0.8 ± 5.5</td>
<td>-1.3 ± 4.9</td>
<td>1.9 ± 5.3</td>
<td>3.2 ± 1.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horizontal plane</td>
<td>-0.1 ± 0.6</td>
<td>-1.2 ± 2.6</td>
<td>-2.8 ± 2.9</td>
<td>0.4 ± 2.2</td>
<td>3.2 ± 1.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women (n=9)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sagittal plane</td>
<td>14.4 ± 14.5</td>
<td>15.7 ± 11.7</td>
<td>-0.5 ± 9.8†‡</td>
<td>16.9 ± 11.5</td>
<td>17.4 ± 6.6</td>
<td>&lt;0.001</td>
<td>0.767</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frontal plane</td>
<td>-0.4 ± 3.5</td>
<td>-0.7 ± 3.5</td>
<td>-2.6 ± 4.3</td>
<td>0.4 ± 4.0</td>
<td>3.1 ± 1.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Horizontal plane</td>
<td>0.1 ± 0.7</td>
<td>-0.1 ± 1.6</td>
<td>-2.4 ± 1.5</td>
<td>1.4 ± 1.6</td>
<td>3.9 ± 1.4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*A positive sagittal plane angle indicated a lordotic spinal alignment and negative sagittal plane angle indicated a kyphotic spinal alignment. A positive value in the frontal and horizontal plane indicated a right lateral flexion or rotation, respectively. A negative value in the frontal and horizontal plane indicated a left lateral flexion or rotation, respectively.
†Significant difference to Habitual posture after adjustment for multiple comparisons using the Bonferroni correction.
‡Significant difference to Start position after adjustment for multiple comparisons using the Bonferroni correction.
Habitual posture and Start position in degrees (°) presented in mean values and 95% CI. LL = Lower lumbar spine (L2-S2), UL = Upper lumbar spine (T11-L2).

In contrast, Walsh et al.16 found that athletes extended their lumbar spines to a significant degree when squatting with heavier loads (60 and 80% of 1RM) but not with light loads (40% of 1RM). However, regardless of load, the ath-
Figure 3. Angle between IMU at L2 and axis of gravity during the squat movement, (eccentric phase approx. 0-50 %, concentric phase approx. 50-100 %). Presented as mean angle (degrees) for men (blue line, n=14) and women (red line, n=9) and standard deviation (blue/red shade).

...letes flexed their spines while descending to the bottom position. Walsh et al. noted that as a result, the athletes brought the weight in front of the spinal column and shifted their center of gravity anteriorly towards the forefoot. It was therefore speculated that lifters hyperextended their lumbar spine to shorten the moment arm length and to keep the center of gravity within the base of support. However, a shorter moment arm length could be achieved both by flexing and extending the spine and the reason for choosing one over the other direction is yet to be determined and may be dependent on the torso inclination. It should also be noted that in the study by Walsh et al. the subjects did not squat to IPF approved squat depth, but rather reversed the movement at approximately 90 degrees of hip flexion, which could explain the opposing results compared to the present study and the study by McKean et al.15

Minor adjustments of spinal alignment were made in the frontal and horizontal planes. This is in agreement with previous research showing less than 5° of lateral flexion and axial rotation when squatting with moderate to heavy loads.16 These minor adjustments in the frontal and horizontal planes are most likely not clinically relevant and implies that analyzing and correcting movements in these planes may be of less importance for the overall lifting performance.

Position statements and guidelines have proposed that neutral posture should be preserved during the squat, and that the trunk should be held stable without waveriing or displacements throughout the movement.6-8 The observed adjustments of the lumbar spine alignment would imply that the lifters in the present study completed the squat with a technique proposed to increase risk of back injuries.6-8 However, the causative relationship between back injury and spinal alignment when squatting is yet to be confirmed.33 Further, the magnitude of the adjustments into flexion during the squat were 5.8 ± 4.3 and 16.7 ± 5.6 degrees in upper and lower lumbar spine, respectively, and should therefore be well within the lumbar spine range of motion in flexion/extension.

Some methodological considerations should be considered. Firstly, the lifters were asked to perform a squat with the only instruction being to, at their own pace, descend to a depth where the crease of the hip was lower than the top of the knee. Regarding depth, it was ascertained that they descended until the hip crease was lower than the top of the knee but some of them might have reached a greater depth. Therefore it is possible that these variations could, at least in part, explain the adjustments in spinal alignment.

Secondly, the lifters were instructed to perform three repetitions with a load equivalent to 70% of the lifters self-estimated 1RM. This repetition range and load are commonly used by both powerlifters and weightlifters and previous studies quantifying spinal alignment when squatting have used loads ranging from 40% to 80% of 1RM. It has to be noted though that load might influence on spinal alignment among recreational lifters.10 It is not yet known whether this also is true for competitive lifters, who are more used to lifting heavy loads.

Thirdly, when performing the ANOVA analysis the mean value of the three repetitions was used. This could have resulted in a regression to the mean. However, this was chosen over using a single repetition since using the mean values reduces the risk of random variations between repetitions.

Fourthly, angles measured using sensors on the skin might differ from the actual skeletal spinal alignment when musculature underneath is contracted. Skin tissue artefacts is a problem that cannot, however, be eliminated when using measurement systems that are based on mounting markers or sensors on the skin and must be considered when interpreting the results.
Lastly, a few lifters reported current injuries to the lumbopelvic, hip, knee and shoulder regions which impaired, but did not hinder, their ability to perform the squat and/or deadlift. It is also possible that some lifters had previously experienced pain and injury to the lumbopelvic region and lower extremities. Since it is well known that pain conditions can cause short term adaptations to lumbar spine kinematics and possibly also long term adaptations, the injury prevalence in the current sample could have affected the results. However, there were only two/three participants in total who reported a current lumbopelvic or hip injury and therefore no statistical analysis of how pain could be associated with differences in lumbopelvic kinematics was included.

CONCLUSIONS

The results of the present study show that spinal adjustments are made by experienced male and female power- and weightlifters in all planes of movement during the barbell back squat. Most significantly, all lifters reduced their lordosis in the start position of the squat compared to their habitual posture, and all lifters reduced their lordosis even further during the squat. Male lifters also seem to reduce their lower lumbar spine lordosis significantly more than females in the start position compared to their habitual posture. The results can impact practice with relevance to how the barbell back squat is assessed, instructed, and executed in regard to the notion that the lumbar spine needs to maintain its neutral position during lifting. For example, the results imply that coaches and lifters might not need to be over vigilant in correction of spinal movements as long as the amplitude of movements do not place the individual in their outer ranges of motion. In all, spinal adjustments during heavy resistance training need further study to reveal their importance for performance and/or injury risk and possible explanations to why they occur.

CONFLICTS OF INTEREST

The authors report no conflicts of interest.

ACKNOWLEDGMENTS

The authors wish to thank Jimmy Falk for his help with the data collection and to Boxen Umeå, Sweden, for allowing us to collect data in their facilities.

FUNDING

None declared.

Submitted: September 28, 2022 CDT, Accepted: June 16, 2023 CDT
REFERENCES


Original Research

The Effect of Volitional Preemptive Abdominal Contraction on Biomechanical Measures During A Front Versus Back Loaded Barbell Squat


1 Physical Therapy Department, Hardin-Simmons University, 2 Center for Rehabilitation Research, School of Health Professions, Texas Tech University Health Sciences Center, 3 Nvision Biomedical Technologies, 4 Southern California PT Fellowship Program, Kaiser Permanente, 5 University Hospital for Orthopaedics and Trauma Surgery Pius-Hospital, University of Oldenburg, 6 Department of Orthopedics, University of Groningen, University Medical Center Groningen, 7 Physical Therapy Department, Mary Baldwin University, 8 Facultad de Medicina Clinica Alemana, Universidad del Desarrollo

Keywords: biomechanics, abdominal bracing, electromyography, spine, squat

https://doi.org/10.26603/001c.84306

International Journal of Sports Physical Therapy

Background

Weightlifting is growing in popularity among recreational and competitive athletes. The barbell back squat (BackS) is commonly included in these training programs, while the barbell front squat (FrontS) is commonly performed as a component of other lifts such as the power clean or clean and jerk, it is less commonly practiced in isolation.

Hypothesis/Purpose

The purpose of this study was to examine the effects of VPAC performance on trunk muscle and LE biomechanical responses during loaded BackS versus FrontS in healthy subjects.

Study Design

Controlled Laboratory Study

Methods

Healthy male subjects with the ability to perform a sub-maximal loaded barbell squat lift were recruited. Subjects completed informed consent, demographic/medical history questionnaires and an instructional video. Subjects practiced VPAC and received feedback. Surface electromyography (sEMG) electrodes and kinematic markers were applied. Muscles included were the internal oblique (IO), external oblique (EO), rectus abdominis, iliocostalis lumborum (ICL), superficial multifidi, rectus femoris, vastus lateralis, biceps femoris, and gluteus maximus. Maximal voluntary isometric contractions established reference sEMG values. A squat one-rep-max (1RM) was predicted by researchers using a three to five repetition maximum (3RM, 5RM) load protocol. Subjects performed BackS trials at 75% 1RM while FrontS trials were performed at 75% BackS weight, both with and without VPAC. Subjects performed three repetitions of each condition with feet positioned on two adjacent force plates. Significant interactions and main effects were tested using a 2(VPAC strategy) x 2(squat variation) and 2(VPAC strategy) x 2(direction) within-subject repeated measures ANOVAs. Tukey’s Post-Hoc tests identified the location of significant differences.

Corresponding author:
Joseph McCormick, PT, ScD
Physical Therapy Department, Hardin-Simmons University, Abilene, TX, USA
Fax: 325-670-5868
Phone: 325-670-1927
Email: joe.mccormick@hsutx.edu
Results
Trunk muscle activity was significantly higher during FrontS versus BackS regardless of VPAC condition. (IO: \( p<0.018 \), EO: \( p<0.001 \), ICL: \( p<0.001 \)) VPAC increased performance time for both squat variations (\( p=0.0011 \)), which may be associated with decreased detrimental force potential on the lumbar spine and knees. VPAC led to improved ability to maintain a neutral lumbar spine during both squat variations. This finding is associated with decreased detrimental force potential on the lumbar spine.

Conclusions
Findings could help guide practitioners and coaches to choose squat variations and incorporate VPAC strategies during their treatments and/or training programs.

Level of Evidence
Level 3
©The Author(s)

INTRODUCTION
The popularity of weightlifting among recreational and competitive athletes is growing. The barbell back squat (BackS) is commonly included in training programs for these athletes. The BackS is defined as positioning a barbell across the shoulders, on the trapezius and slightly above the posterior deltoids, while allowing the hips and knees to flex until the thighs are parallel to the floor. The maneuver is completed by extending the hips and knees to return to a standing position.\(^1\) The BackS is a common strengthening exercise for enhancing sports performance due to the movement’s capacity to improve strength, induce muscle hypertrophy, and mimic common sport-related movements.\(^2,3\)

The barbell front squat (FrontS) is less commonly performed in isolation for lower extremity (LE) strengthening and is considered comparable in overall muscle recruitment and performance.\(^1\) The FrontS is defined by “positioning the barbell across the anterior deltoids and clavicles and fully flexing the elbows to position the upper arms parallel to the floor.”\(^1\) The maneuver is completed in a similar fashion the BackS. Though both squat variations demonstrate the ability to provide strength improvements, it is unknown which variation best optimizes training while minimizing injury risk.

Injury incidence among weightlifters is between 1.0-4.4 injuries/1,000 training hours.\(^4\) Powerlifting movements such as BackS and deadlifts are commonly associated with self-reported low back injuries in recreational weightlifters.\(^5\) Snug\(^6\) reported low back pain sufferers demonstrated increased hip flexion and lumbar extension when squatting compared to normal subjects. The FrontS induces less forward trunk lean, which has been associated with decreased likelihood of lumbar injury.\(^7\) The BackS, however, has been associated with increased forward trunk lean, potentially increasing lumbar spine shear forces. Similarly, detrimental LE forces are noted during the BackS versus the FrontS. Gullet et al\(^8\) reported higher knee shear and compressive forces during the BackS compared to the FrontS.

Many investigators have examined neuromuscular control responses during both the BackS and FrontS. Various authors reported minimal erector spinae (ES) activation differences between the two squat variations.\(^1,2\) Comfort et al\(^8\) however, reported the FrontS resulted in higher ES activation compared to the BackS. Based on their findings, these authors suggest the FrontS produces greater spinal stabilizer activation versus the BackS.\(^8\)

Spinal stabilization occurs through muscle activation and a concurrent intra-abdominal pressure increase.\(^9\) Volitional pre-emptive abdominal contraction (VPAC) has been used to stabilize the trunk during dynamic activities. For example, when VPAC is performed during a lifting task, it increases trunk and hip extensor force, internal oblique muscle thickness, and intra-abdominal pressure.\(^9,10\) These events are associated with increased power during hip and trunk extension movements.\(^11\) In addition, VPAC may reduce low back pain recurrence.\(^12\) Research suggests VPAC produces spinal stabilization and a decreased ability to produce a VPAC may place individuals at higher injury risk.\(^13,14\)

Two commonly described VPAC strategies are the abdominal drawing-in maneuver (ADIM) and the abdominal bracing maneuver (ABM). The ADIM involves volitional transversus abdominis contraction, whereas the ABM involves concurrent muscle contraction around the entire trunk. Where ADIM is effective in activating the transversus abdominus, ABM is more effective than ADIM in activating the transverse abdominus as well as the superficial multifidi and the internal oblique.\(^15-17\) Maeo et al\(^18\) reported that the ABM is commonly used in the health and fitness industry and is very effective in providing spinal stability. The ABM is more effective than the ADIM for reducing lumbar displacement and increasing trunk stability during external perturbations.\(^19\)

No study has examined the impact of VPAC on different squat variations regarding trunk and LE kinetics, kinematics, and muscle activation. The purpose of this study was to examine the effects of VPAC performance on trunk muscle and LE biomechanical responses during loaded BackS versus FrontS in healthy subjects. The first hypothesis was that performing VPAC during BackS and FrontS trials would result in higher trunk muscle activation, regardless of squat variation. Next, the authors hypothesized the addition of VPAC will produce a significant change in LE muscle acti-
vation during a BackS and FrontS. For our third hypothesis, the authors speculated VPAC performance will result in significant differences in trunk and LE kinetics and kinematics. The fourth hypothesis surmised there would be significant differences in trunk and LE kinetics, kinematics, and muscle activity between the BackS and FrontS when performed without VPAC. For the fifth and final hypothesis, we postulated there would be significant differences in trunk and LE kinetics, kinematics, and muscle activity between the BackS and FrontS ascending and descending phases.

MATERIALS AND METHODS

This within-subject investigation examined the effect of VPAC on trunk and LE kinetics, kinematics, and muscle activation during two squat variations. All study-related measures followed the Helsinki Declaration's ethical principles. In conformity, the study protocol was reviewed and approved by the local university's Institutional Review Board for the Protection of Human Subjects (L19-046) before study initiation. Prior to enrollment, eligible subjects were informed about potential study-related risks and benefits and signed an approved informed consent.

SAMPLE

Based on a moderate effect size (f = 0.2), a desired 80% power, and α = 0.05, a convenience sample of 26 healthy male subjects (18–55 years old) was recruited from a university population.20

INCLUSION AND EXCLUSION CRITERIA

Eligible subjects had to be able to (1) perform a sub-maximal loaded barbell squat lift; (2) perform VPAC on command; and (3) follow English language instructions. Exclusion criteria were: (1) Existing active spinal, upper extremity or LE pain meriting healthcare attention based on clinical judgement; (2) Upper or lower quarter injury requiring healthcare attention within 12 months prior to testing; (3) Any underlying neuromuscular or joint disease; (4) Any diagnosed and presently active abdominal, respiratory, or gastrointestinal condition; (5) Any significant spinal condition (including but not limited to scoliosis, spina bifida, tumors, present fractures, rheumatologic disorders) requiring healthcare attention; (6) Any blood clotting disorder or anticoagulant therapy; (7) History of abdominal or spinal surgery; (8) Any skin allergy preventing the use of electrode pads; and (9) BMI > 30.

INSTRUMENTATION

Thirty-six reflective markers were used to obtain 3-D trunk and LE kinematic data using an 8-camera motion capture system recorded at 100 Hz (VICON Nexus 2.5, Denver, CO). Changes in marker position were used to quantify range of motion elicited at each joint/plane, peak angles, and associated time ranges. In-ground force plates (Bertec, Columbus, OH) collected 3-D vertical (vGRF) and horizontal (hGRF) ground reaction force data at 1000 Hz. A freestanding EMG system (Noraxon USA, Scottsdale, AZ; 2000 Hz) gathered EMG data from the following muscles: internal oblique (IO), external oblique (EO), rectus abdominis (RA), ilio-costalis lumborum (ICL), superficial multifidi (SM), rectus femoris (RF), vastus lateralis (VL), biceps femoris (BF), and gluteus maximus (GM). Surface EMG data were collected bilaterally using dual Ag/AgCl EMG electrodes positioned on the subject according to the recommendations of previous investigators.21-26 The EMG impedance were >10⁹ MΩ, with a common mode rejection ratio > -92 dB and baseline noise <1.2 µV root-mean-square. Manual isometric testing appropriate to each respective muscle was used for proper electrode placement confirmation.

PREPARATORY PROCEDURES

Upon arrival for testing, each subject completed a medical history questionnaire and an investigator measured subjects’ height (m) and weight (kg) to verify individual enrollment eligibility. Following, subjects watched an instructional video explaining the study purpose, potential risks and benefits, and all experimental procedures. Next, an investigator took anthropometric measurements relevant to the study. Next, subjects were instructed on performing the two squat variations. They were allowed to practice each squat with a 45-pound bar and received feedback. Following squat familiarization, subjects were instructed on how to perform VPAC. For the No-VPAC condition, subjects were taught to maintain a relaxed state. For the Yes-VPAC condition, subjects were cued to “gently inhale, then exhale, now stiffen your trunk as though you will be hit in the belly. Hold this contraction”27. The tester palpatated for proper and consistent abdominal contraction and visually confirmed absence of Valsalva maneuver and aberrant movement of the ribcage, shoulders, or pelvis.

Maximal voluntary isometric contraction (MVIC) was performed to normalize EMG signal amplitudes and allow measurement comparison among different muscles and between subjects.28 For the RA, ICL, and dominant side EO and IO, subjects sat in a stable chair with the with arms crossed in front of their body and hips and knees flexed to 70° and 90°, respectively. Two straps were positioned at chest-level and around the subject’s waist. The investigator stood behind the subject and instructed them to take in a breath and blow out while performing a maximal isometric axial spine rotation, while the investigator applied manual resistance to upper trunk. The knee extensors were tested in the seated position with the knee flexed to 90° and the lower leg stabilized by an adjustable strap located at the subject’s ankle and fixed to the base of the stable chair.

Hip extensor and knee flexor MVICs were performed with the subject positioned prone. For the hip extensors, the knees were flexed to 90°. An adjustable strap was secured around the iliac crests and buckled around the table while a second adjustable strap was secured on the distal posterior thigh. The same procedure was followed for the knee flexors at 45° of knee flexion and the strap secured around the ankle. For trunk and LE muscle normalization, subjects were asked to perform three consecutive 5-second trials of maximum contraction for each respective muscle.
The average was taken from the three trials for further analysis. Each MVIC was separated by a 1- to 2-minute rest period.

Following, subjects performed a five-minute warmup on a cycle ergometer. Then, they performed a squat according to the subject’s own technique to a self-selected depth, where the thighs were near parallel to the ground. They repeated this procedure for two sets of five repetitions with 15 seconds of rest between sets. No attempt was made to control for speed, as that may have altered the lifting technique. Subjects were then allowed to stretch according to their prior lifting experience.29

Next, a researcher tested the subjects for their BackS predicted one repetition maximum (1RM) load using a three to five repetition maximum (3RM, 5RM) load protocol. This method uses submaximal loads that better represent those managed during routine training sessions.30 Each subject was allowed two trials of the 3RM or 5RM in an attempt to establish their predicted 1RM. The weight of the first attempt was the subject’s preference. Weight could be increased by a minimum of 5lb in subsequent trials. The subject performed repetitions to fatigue with a target of three to five repetitions achieved. Spotters stood on both ends of the bar to assist in unloading the weight as needed. Additionally, the spotters ensured that proper form was achieved during all repetitions. During all testing, subjects squatted with feet hip-width apart to a depth that achieved thighs parallel to the floor.29,31,32 The squat depth was not strictly controlled so as to not alter the natural kinematics. A standardized rest period of two minutes was allowed after each attempt for adequate recovery.29,32 The 1RM was predicted utilizing either the Epley equation (3RM equation) or the Brzycki equation (5RM equation) depending on each subject’s performance.50 The following formulae describe the prediction from 3RM or 5RM, accordingly:

\[
\begin{align*}
3RM \text{ prediction equation: } 1RM &= [0.053 \times (\text{reps}) + \text{rep wt}] \\
5RM \text{ prediction equation: } 1RM &= \frac{\text{rep wt}}{102.78 - 2.78\times(\text{reps})}
\end{align*}
\]

Finally, researchers equipped subjects with retroreflective markers. A four-marker plate (i.e., quadratus marker) was placed at T10 and 36 kinematic markers were placed bilaterally at the following sites: iliocost, ASIS, PSIS, upper and lower posterior thigh, lower anterior thigh, lateral and medial epicondyles of the knee, head of fibula, tibial tuberosity, lateral and medial malleoli of the ankle, the heel, head of the first, second and fifth metatarsals as established by previous investigators.33,34 Subjects wore comfortable clothing during data collection and standardized lab shoes.

DATA COLLECTION

The squat rack was assembled over the force plates. Subjects were positioned with a foot on each force plate during the squat procedure. Subjects performed three loaded squats under each condition (BackS with VPAC, Backs without VPAC, FrontS with VPAC, FrontS without VPAC). The BackS was performed at a load level of 75% of the predicted 1RM (75%1RM).2 Based on subjects’ established capability during instrument testing, the FrontS was performed at a more conservative load level of 75% of the BackS 75%1RM. Though no established protocol exists in the literature to establish FrontS load based on a 1RM BackS calculation, this conservative load was chosen to ensure the safety of all subjects as well as to not fatigue subjects by performing two separate 1RM tests on the testing day. Squats were performed according to the same position and depth requirements as in 3-5RM predictive testing. All conditions were randomized for each subject to decrease the influence of fatigue and learning effect. A standardized rest period of 1–5 minutes (depending on subject preference) was allotted between trials to allow full recovery.29,32

STATISTICAL ANALYSIS

Reflective markers were tracked, labeled, and reconstructed using the Vicon Nexus software. Force and position data were filtered using a 4th order Butterworth digital filter (10 Hz cutoff frequency). These data were exported to Matlab (Version 9.5, R2018b, Mathworks, Inc, USA) for further processing using custom algorithms. A six-degrees-of-freedom link segment model was applied to the marker position data. A static standing calibration trial was used to define trunk and LE segmental coordinate systems and to calculate joint axes locations. Joint kinematics were calculated using an Xy’z Euler rotation sequence in an order of flexion/extension, abduction/adduction, and internal/external rotation. Trunk and pelvis segment kinematics were calculated as the orientation of the respective segment relative to the laboratory (global) coordinate system. Hip, knee, and ankle joint kinematics were calculated as the orientation of the distal segment relative to the proximal segment and mass location were estimated using previously published data.35 The biomechanical dependent variables of interest were calculated from the processed time series data during the descent and ascent phase of the loaded squat lift.

Similarly, all EMG data were imported into a custom Matlab program. For every MVIC and submaximal MVIC trial raw EMG data was sampled at 1000 Hz, followed by a full-wave rectification and 4th order, 2-pass, no phase shift Butterworth filter with a 20–400 Hz bandpass. The EMG signals’ average root mean square (RMS) for each MVIC and submaximal normalization trial was calculated for each muscle from the respective trial’s final three-seconds of the contraction. Each five–second MVIC was trimmed to the desired three-second contraction followed by the RMS value calculation. Furthermore, each squat trial’s raw EMG was sampled at 2000 Hz, followed by full-wave rectification and 4th order, 2-pass, no phase shift Butterworth filter with a 20–450 Hz bandpass. The RMS value was calculated for each squat repetition during the eccentric downward and concentric upward portions of the squat. All EMG data quality was checked based on power density spectrum observation and individual EMG graph inspection for artifacts and excessive noise prior to inclusion into the data set.

The trunk and LE muscle EMG data were reported as a percentage of the reference contraction values (or RMS-EMG%). Statistical analyses were conducted using the Sta-
tical Package for Social Sciences for Windows. Descriptive data analyses established values for central tendency (means) and dispersion (standard deviation and 95% confidence intervals [CI]). Data normality was established using the Shapiro-Wilk test (p-value > 0.05), as well as skewness and kurtosis (between -2.0 and +2.0). Data sphericity was assessed using a Mauchly's test for Sphericity (p > 0.05).

A 2 (VPAC strategy) x 2 (squat variation) as well as a 2 (squat variation) x 2 (direction - descend/ascend phase of each squat) within-subject, repeated measures ANOVA was used to test for interactions and significant main effects of: (1) VPAC on trunk muscle activity during BackS and FrontS; (2) VPAC on LE muscle activation during BackS and FrontS; (3) VPAC on trunk and LE kinetics and kinematics during BackS and FrontS; (4) BackS versus FrontS squat variations on trunk and LE kinetics, kinematics, and muscle activation; and (5) ascending and descending phases of both BackS and FrontS on trunk and LE kinetics, kinematics, and muscle activity, respectively. Post-hoc comparisons were used to locate significant differences. Bonferroni corrections to an alpha level of .05 were implemented in order to reduce the chance of a type 1 error.

RESULTS

Descriptive data were established for the 26 male subjects’ age (22.8 ± 3.1 years), height (182.4 ± 7.4 cm), weight (84.3 ± 11 kg) and BMI (25.3 ± 2.9 kg/m²). Moreover, descriptive data were established for subjects’ 5RM BackS weight (240.4 ± 44.0 lbs), predicted 1RM BackS weight (270.5 ± 49.6 lbs), BackS working weight (202.8 ± 37.2 lbs), and FrontS working weight (152.1 ± 27.9 lbs).

Considering the first hypothesis, VPAC use did not result in significant trunk muscle activity changes as measured by surface EMG in either squat variation (Table 1). Regarding the second hypothesis, no significant main effect was observed for LE muscle activation based on VPAC condition or squat variation (Table 2). Regarding the third hypothesis, the use of VPAC resulted in significant increased time to reach sagittal plane peak hip and knee angles for both squat variations (Table 3, Table 4). Furthermore, a significant main effect was observed for spinal position, where the addition of VPAC resulted in significantly decreased lumbar extension in both squat variations (Table 5). The addition of VPAC resulted in increased performance time during descent, ascent, and total time for both squat variations. While the descent of both squat variations demonstrated increased performance duration, the BackS demonstrated a longer descent performance time versus FrontS (Table 6).

Regarding the fourth hypothesis, there was a significant main effect for trunk muscle activity based on squat variation. FrontS resulted in significantly higher trunk muscle activity in the IO and EO muscles as well as the ICL (Table 1). At the same time, FrontS demonstrated less lumbar extension than BackS. Finally, a significant main effect was noted regarding moments at the right hip and ankle (Table 7). In the BackS, the moment was greater at both the hip and ankle joints in the sagittal plane.

Regarding the fifth hypothesis, a significant main effect for trunk muscle activity during the descending versus ascending phases of both squat variations was observed regardless of VPAC condition. Here, the IO and EO exhibited greater activity during the descending phase of both squat variations (Table 1).

DISCUSSION

The purpose of this study was to assess the effect of VPAC performance on trunk muscle and LE biomechanical responses during loaded BackS versus FrontS in healthy subjects. This study was the first to demonstrate VPAC performance resulted in slower movement performance time and a more neutral lumbar spine position during weighted barbell squats. In addition, this study supports the findings of previous researchers, showing that FrontS resulted in increased erector spinae muscle activity when compared to BackS, regardless of VPAC condition. Furthermore, this study is the first to our knowledge to demonstrate increased IO and EO muscle activation during FrontS versus BackS.

ADIM has been shown to be very useful for onset activation re-education at the beginning of rehabilitation but not as useful as ABM for more functional/highly demanding tasks. This is likely due to the fact that ABM creates a co-contraction of several trunk muscles providing spinal stability throughout the duration of functional/highly demanding activity. Regarding the first hypothesis, there was no detectable change in abdominal muscle activity in response to different VPAC conditions. This lack of change is likely due to a ceiling effect secondary to the load placed on the tested muscles. At a load of 75% IRM, it is possible trunk muscle bracing occurred automatically, regardless of VPAC condition. Additionally, subjects may not have been given sufficient time between VPAC instruction and trial completion for abdominal muscle deactivation during the no-VPAC condition. Subjects were instructed on VPAC activation shortly before performing the squat trials. It is possible the subjects were not able to suspend use of this new skill between trials. In addition, based on previous experience, subjects may be accustomed to bracing when performing loaded barbell squats and experienced a challenge when attempting to consciously not brace.

Similarly, there was no significant difference in LE muscle activity based on VPAC condition or squat variation. These findings are consistent with those of Gullett in which they observed no significant difference in overall LE muscle activation between the FrontS and BackS. Though participants lifted less weight with the FrontS, the overall muscle activity was not significantly different between squat variations. Therefore, the same benefits from the workout may be achieved with the added benefit of decreasing potentially detrimental forces on the knees by performing FrontS rather than BackS.

Regarding the third hypothesis, VPAC resulted in decreased lumbar extension during both squat variations. Shoenfeld proposed squatting with a flexed lumbar spine decreases erector spinae’s ability to accommodate compressive loads and potentially increases injury risk. The authors
suggest the more neutral spine achieved during VPAC conditions may optimize trunk muscle alignment by placing subjects in a more mechanically advantageous position, especially during the FrontS. The FrontS inherently results in less lumbar extension, however, the addition of VPAC demonstrated a significant decrease in lumbar extension in both squat variations. When VPAC was incorporated, the spine was able to remain in a more "neutral" alignment, potentially resulting in decreased shear forces and rendering the lumbar spine at less injury risk.

Previous investigators have reported individuals must lean forward to maintain balance during squats, leading to increased hip flexion and moving the center of gravity further away from the lumbar spine.\(^1\)\(^7\) This leads to increased torque at the lumbar spine, ultimately increasing shear force potential.\(^1\)\(^7\) The forward lean can also result in a decreased tissue tolerance to compressive load as well as a load transfer from muscles to passive tissues such as the discs, increasing the likelihood of disc injury.\(^4\)\(^1\) Future research should further investigate the impact of VPAC on trunk position and injury risk in light of these findings.

As observed in this study, VPAC performance decreased performance speed in both squat variations. This has the potential to benefit weightlifters with respect to reducing injury risk. Hattin et al\(^4\) reported increased squat performance speed results in significantly higher tibiofemoral joint anteroposterior shear and compressive forces. In addition, a common finding with increased squatting speed is a
Table 2. One-way ANOVA tests of Within-Subjects effects for LE EMG

<table>
<thead>
<tr>
<th>Analyses</th>
<th>Muscle</th>
<th>df</th>
<th>F</th>
<th>Sig</th>
<th>PES</th>
<th>PWR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barbell•VPAC•LiftPhase</td>
<td>GM</td>
<td>1.25</td>
<td>0.469</td>
<td>0.5</td>
<td>0.018</td>
<td>0.101</td>
</tr>
<tr>
<td>VPAC•LiftPhase</td>
<td>GM</td>
<td>1.25</td>
<td>0.635</td>
<td>0.433</td>
<td>0.025</td>
<td>0.12</td>
</tr>
<tr>
<td>Barbell•LiftPhase</td>
<td>GM</td>
<td>1.25</td>
<td>0.503</td>
<td>0.485</td>
<td>0.02</td>
<td>0.105</td>
</tr>
<tr>
<td>Barbell•VPAC</td>
<td>GM</td>
<td>1.25</td>
<td>4.165</td>
<td>0.052</td>
<td>0.143</td>
<td>0.501</td>
</tr>
<tr>
<td>LiftPhase</td>
<td>GM</td>
<td>1.25</td>
<td>3.251</td>
<td>0.083</td>
<td>0.115</td>
<td>0.411</td>
</tr>
<tr>
<td>VPAC</td>
<td>GM</td>
<td>1.25</td>
<td>4.868</td>
<td>0.037</td>
<td>0.163</td>
<td>0.564</td>
</tr>
<tr>
<td>Barbell</td>
<td>GM</td>
<td>1.25</td>
<td>0.001</td>
<td>0.974</td>
<td>&lt; .001</td>
<td>0.05</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Analyses</th>
<th>Muscle</th>
<th>df</th>
<th>F</th>
<th>Sig</th>
<th>PES</th>
<th>PWR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barbell•VPAC•LiftPhase</td>
<td>BF</td>
<td>1.25</td>
<td>5.108</td>
<td>0.033</td>
<td>0.17</td>
<td>0.584</td>
</tr>
<tr>
<td>VPAC•LiftPhase</td>
<td>BF</td>
<td>1.25</td>
<td>0.77</td>
<td>0.389</td>
<td>0.03</td>
<td>0.135</td>
</tr>
<tr>
<td>Barbell•LiftPhase</td>
<td>BF</td>
<td>1.25</td>
<td>4.35</td>
<td>0.047</td>
<td>0.148</td>
<td>0.518</td>
</tr>
<tr>
<td>Barbell•VPAC</td>
<td>BF</td>
<td>1.25</td>
<td>0.41</td>
<td>0.528</td>
<td>0.016</td>
<td>0.095</td>
</tr>
<tr>
<td>LiftPhase</td>
<td>BF</td>
<td>1.25</td>
<td>8.073</td>
<td>0.009*</td>
<td>0.244</td>
<td>0.78</td>
</tr>
<tr>
<td>VPAC</td>
<td>BF</td>
<td>1.25</td>
<td>0.193</td>
<td>0.664</td>
<td>0.008</td>
<td>0.071</td>
</tr>
<tr>
<td>Barbell</td>
<td>BF</td>
<td>1.25</td>
<td>2.748</td>
<td>0.11</td>
<td>0.099</td>
<td>0.357</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Analyses</th>
<th>Muscle</th>
<th>df</th>
<th>F</th>
<th>Sig</th>
<th>PES</th>
<th>PWR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barbell•VPAC•LiftPhase</td>
<td>RF</td>
<td>1.25</td>
<td>2.513</td>
<td>0.126</td>
<td>0.091</td>
<td>0.332</td>
</tr>
<tr>
<td>VPAC•LiftPhase</td>
<td>RF</td>
<td>1.25</td>
<td>0.445</td>
<td>0.511</td>
<td>0.017</td>
<td>0.998</td>
</tr>
<tr>
<td>Barbell•LiftPhase</td>
<td>RF</td>
<td>1.25</td>
<td>0.157</td>
<td>0.695</td>
<td>0.006</td>
<td>0.067</td>
</tr>
<tr>
<td>Barbell•VPAC</td>
<td>RF</td>
<td>1.25</td>
<td>9.617</td>
<td>0.005*</td>
<td>0.278</td>
<td>0.846</td>
</tr>
<tr>
<td>LiftPhase</td>
<td>RF</td>
<td>1.25</td>
<td>0.008</td>
<td>0.928</td>
<td>&lt; .001</td>
<td>0.051</td>
</tr>
<tr>
<td>VPAC</td>
<td>RF</td>
<td>1.25</td>
<td>0.567</td>
<td>0.458</td>
<td>0.022</td>
<td>0.112</td>
</tr>
<tr>
<td>Barbell</td>
<td>RF</td>
<td>1.25</td>
<td>0.214</td>
<td>0.648</td>
<td>0.008</td>
<td>0.073</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Analyses</th>
<th>Muscle</th>
<th>df</th>
<th>F</th>
<th>Sig</th>
<th>PES</th>
<th>PWR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barbell•VPAC•LiftPhase</td>
<td>VL</td>
<td>1.25</td>
<td>0.238</td>
<td>0.63</td>
<td>0.009</td>
<td>0.076</td>
</tr>
<tr>
<td>VPAC•LiftPhase</td>
<td>VL</td>
<td>1.25</td>
<td>1.097</td>
<td>0.305</td>
<td>0.042</td>
<td>0.172</td>
</tr>
<tr>
<td>Barbell•LiftPhase</td>
<td>VL</td>
<td>1.25</td>
<td>0.574</td>
<td>0.456</td>
<td>0.022</td>
<td>0.113</td>
</tr>
<tr>
<td>Barbell•VPAC</td>
<td>VL</td>
<td>1.25</td>
<td>0.059</td>
<td>0.811</td>
<td>0.002</td>
<td>0.056</td>
</tr>
<tr>
<td>LiftPhase</td>
<td>VL</td>
<td>1.25</td>
<td>0.005</td>
<td>0.943</td>
<td>&lt; .001</td>
<td>0.051</td>
</tr>
<tr>
<td>VPAC</td>
<td>VL</td>
<td>1.25</td>
<td>1.038</td>
<td>0.318</td>
<td>0.04</td>
<td>0.165</td>
</tr>
<tr>
<td>Barbell</td>
<td>VL</td>
<td>1.25</td>
<td>0.504</td>
<td>0.484</td>
<td>0.02</td>
<td>0.105</td>
</tr>
</tbody>
</table>

Barbell = front or back barbell position, VPAC = yes/AB or no/NB, LiftPhase = ascend or descend, Muscle: GM = gluteus maximus, BF = bipec femoris, RF = rectus femoris, VL = vastus lateralis (all muscles refer to right side); df = degrees of freedom, F = f-statistic, Sig = significance, PES = partial eta squared effect size, PWR = power. (significance was familywise adjusted to α = .025 for GM, BF pairing and α = .025 for RF, VL pair); *significant result.

concurrent “bounce” at the bottom of the squat, potentially increasing these compressive forces up to 33%. Lavender et al.[14] found faster lifting speed resulted in a greater lumbar spine flexion moment. Similarly, Greenland et al.[15] reported peak lumbar compressive forces occurred at higher speeds. The authors propose the act of attending to VPAC performance improved task vigilance and changed the central nervous system programming during the lifting task. Thus, future research should explore both lifting speed optimization and the effects of VPAC.

Though the authors did not see differences in trunk muscle activation related to VPAC, they did see a difference in trunk muscle activity based on squat variation. With reference to the fourth hypothesis, IO and EO muscles as well as the ICL demonstrated greater activity during the FrontS. The authors propose this is related to the more neutral lumbar spine alignment seen in the FrontS. Similarly, Comfort et al.[16] found the FrontS demonstrated greater ES activity. However, Clark et al.[17] and Gullet et al.[18] found no difference in ES activity comparing FrontS to BackS.

The authors also noted a significant main effect regarding right hip and ankle joint moments in which both moments were greater in the BackS. This finding is expected due to barbell position inducing a forward trunk lean during the BackS versus the FrontS. The authors propose that in addition to reducing detrimental lumbar spine forces, use of the FrontS should be considered to also minimize hip and ankle forces. These findings are consistent with those
of Gullet in which they observed no significant difference in LE muscle activation between the FrontS and BackS. In addition, they also found less potentially detrimental forces on the knees during the FrontS. Future research should further explore these findings.

Regarding the final hypothesis, the authors noted a difference in trunk muscle activity during the descending versus ascending phases of both squat variations, regardless of VPAC condition. In this study the IO and EO muscles exhibited greater activity during the descending phase of both squat variations, with exception to FrontS without bracing. In addition, the authors did not see a difference in ES activity between descending versus ascending phases in either squat variation. This finding is not consistent with previous literature that found increased ES activity during the squat ascent.

One major difference in the current study was the inclusion of IO and EO EMG measures, whereas previous studies focused on the ES. These findings point to the possibility that the oblique muscles play a valuable role in providing stability during the squat descent. The oblique muscles, when contracted bilaterally, flex the trunk and posteriorly tilt the pelvis. During a squat, this action is necessary to counteract the strong erector spinae activity that serves to extend the spine, creating a relative anterior tilt of the pelvis. Haddas et al. found that VPAC increased EO activity during uniplanar drop landing and uniplanar symmetrical box lifting tasks. Additionally, they did not note a difference between descending or ascending phases in the box lift.

The authors propose the key difference in findings is due to the significantly greater load the current participants were under while performing both the FrontS and BackS. With a greater load, subjects must accommodate for the increased stress applied to their trunks. This is especially likely during the decent-to-ascent-transition where lifters may struggle to maintain a neutral posture. In this transition, lifters often lean further forward to maintain bal-

Table 3. One-way ANOVA tests of Within-Subjects effects for kinematic times in the sagittal plane

<table>
<thead>
<tr>
<th>Analyses</th>
<th>Joint</th>
<th>df</th>
<th>F</th>
<th>Sig</th>
<th>PES</th>
<th>PWR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barbell+VPAC</td>
<td>Lumbar Spine</td>
<td>1.25</td>
<td>0</td>
<td>0.999</td>
<td>&lt; .001</td>
<td>0.05</td>
</tr>
<tr>
<td>VPAC</td>
<td>Lumbar Spine</td>
<td>1.25</td>
<td>4.787</td>
<td>0.038</td>
<td>0.161</td>
<td>0.557</td>
</tr>
<tr>
<td>Barbell</td>
<td>Lumbar Spine</td>
<td>1.25</td>
<td>0.158</td>
<td>0.694</td>
<td>0.006</td>
<td>0.067</td>
</tr>
<tr>
<td>Analyses</td>
<td>Joint</td>
<td>df</td>
<td>F</td>
<td>Sig</td>
<td>PES</td>
<td>PWR</td>
</tr>
<tr>
<td>Barbell+VPAC</td>
<td>Right Hip</td>
<td>1.25</td>
<td>0.132</td>
<td>0.719</td>
<td>0.005</td>
<td>0.064</td>
</tr>
<tr>
<td>VPAC</td>
<td>Right Hip</td>
<td>1.25</td>
<td>8.79</td>
<td>0.007*</td>
<td>0.26</td>
<td>0.813</td>
</tr>
<tr>
<td>Barbell</td>
<td>Right Hip</td>
<td>1.25</td>
<td>2.522</td>
<td>0.125</td>
<td>0.092</td>
<td>0.333</td>
</tr>
<tr>
<td>Analyses</td>
<td>Joint</td>
<td>df</td>
<td>F</td>
<td>Sig</td>
<td>PES</td>
<td>PWR</td>
</tr>
<tr>
<td>Barbell+VPAC</td>
<td>Right Knee</td>
<td>1.25</td>
<td>0.058</td>
<td>0.811</td>
<td>0.002</td>
<td>0.056</td>
</tr>
<tr>
<td>VPAC</td>
<td>Right Knee</td>
<td>1.25</td>
<td>9.749</td>
<td>0.004*</td>
<td>0.281</td>
<td>0.851</td>
</tr>
<tr>
<td>Barbell</td>
<td>Right Knee</td>
<td>1.25</td>
<td>1.247</td>
<td>0.275</td>
<td>0.048</td>
<td>0.189</td>
</tr>
<tr>
<td>Analyses</td>
<td>Joint</td>
<td>df</td>
<td>F</td>
<td>Sig</td>
<td>PES</td>
<td>PWR</td>
</tr>
<tr>
<td>Barbell+VPAC</td>
<td>Right Ankle</td>
<td>1.25</td>
<td>0.101</td>
<td>0.753</td>
<td>0.004</td>
<td>0.061</td>
</tr>
<tr>
<td>VPAC</td>
<td>Right Ankle</td>
<td>1.25</td>
<td>5.753</td>
<td>0.024</td>
<td>0.187</td>
<td>0.635</td>
</tr>
<tr>
<td>Barbell</td>
<td>Right Ankle</td>
<td>1.25</td>
<td>1.744</td>
<td>0.199</td>
<td>0.065</td>
<td>0.246</td>
</tr>
</tbody>
</table>

Barbell = front or back barbell position, VPAC = yes/AB or no/NB; df = degrees of freedom, F = f-statistic; Sig = significance; PES = partial eta squared effect size, PWR = power. (Significance was familywise adjusted to α = .0125 for lumbar, R hip, R knee and R ankle); *significant result.

Table 4. One-way ANOVA tests of Within-Subjects effects for kinematic times in the frontal plane

<table>
<thead>
<tr>
<th>Analyses</th>
<th>Joint</th>
<th>df</th>
<th>F</th>
<th>Sig</th>
<th>PES</th>
<th>PWR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barbell+VPAC</td>
<td>Right Knee</td>
<td>1.25</td>
<td>0.636</td>
<td>0.433</td>
<td>0.025</td>
<td>0.12</td>
</tr>
<tr>
<td>VPAC</td>
<td>Right Knee</td>
<td>1.25</td>
<td>1.008</td>
<td>0.325</td>
<td>0.039</td>
<td>0.162</td>
</tr>
<tr>
<td>Barbell</td>
<td>Right Knee</td>
<td>1.25</td>
<td>1.886</td>
<td>0.182</td>
<td>0.07</td>
<td>0.262</td>
</tr>
<tr>
<td>Analyses</td>
<td>Joint</td>
<td>df</td>
<td>F</td>
<td>Sig</td>
<td>PES</td>
<td>PWR</td>
</tr>
<tr>
<td>Barbell+VPAC</td>
<td>Left Knee</td>
<td>1.25</td>
<td>0.709</td>
<td>0.408</td>
<td>0.028</td>
<td>0.128</td>
</tr>
<tr>
<td>VPAC</td>
<td>Left Knee</td>
<td>1.25</td>
<td>2.705</td>
<td>0.113</td>
<td>0.098</td>
<td>0.353</td>
</tr>
<tr>
<td>Barbell</td>
<td>Left Knee</td>
<td>1.25</td>
<td>0.053</td>
<td>0.821</td>
<td>0.002</td>
<td>0.056</td>
</tr>
</tbody>
</table>

Barbell = front or back barbell position, VPAC = yes/AB or no/NB; Joint = left or right knee; df = degrees of freedom, F = f-statistic; Sig = significance; PES = partial eta squared effect size, PWR = power. (Significance was familywise adjusted to α = .0125 for R knee, L knee pairing); *significant result.
Table 5. One-way ANOVA tests of Within-Subjects effects for kinematic angles in the sagittal plane

<table>
<thead>
<tr>
<th>Analyses</th>
<th>Joint</th>
<th>df</th>
<th>F</th>
<th>Sig</th>
<th>PES</th>
<th>PWR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barbell•VPAC</td>
<td>Lumbar Spine</td>
<td>1.25</td>
<td>0.007</td>
<td>0.935</td>
<td>&lt;.001</td>
<td>0.051</td>
</tr>
<tr>
<td>VPAC</td>
<td>Lumbar Spine</td>
<td>1.25</td>
<td>7.388</td>
<td>0.012*</td>
<td>0.228</td>
<td>0.743</td>
</tr>
<tr>
<td>Barbell</td>
<td>Lumbar Spine</td>
<td>1.25</td>
<td>0.172</td>
<td>0.682</td>
<td>0.007</td>
<td>0.068</td>
</tr>
<tr>
<td>Analyses</td>
<td>Joint</td>
<td>df</td>
<td>F</td>
<td>Sig</td>
<td>PES</td>
<td>PWR</td>
</tr>
<tr>
<td>Barbell•VPAC</td>
<td>Right Hip</td>
<td>1.25</td>
<td>3.683</td>
<td>0.066</td>
<td>0.128</td>
<td>0.454</td>
</tr>
<tr>
<td>VPAC</td>
<td>Right Hip</td>
<td>1.25</td>
<td>1.12</td>
<td>0.3</td>
<td>0.043</td>
<td>0.175</td>
</tr>
<tr>
<td>Barbell</td>
<td>Right Hip</td>
<td>1.25</td>
<td>8.366</td>
<td>0.008*</td>
<td>0.251</td>
<td>0.794</td>
</tr>
<tr>
<td>Analyses</td>
<td>Joint</td>
<td>df</td>
<td>F</td>
<td>Sig</td>
<td>PES</td>
<td>PWR</td>
</tr>
<tr>
<td>Barbell•VPAC</td>
<td>Right Knee</td>
<td>1.25</td>
<td>4.406</td>
<td>0.046</td>
<td>0.15</td>
<td>0.523</td>
</tr>
<tr>
<td>VPAC</td>
<td>Right Knee</td>
<td>1.25</td>
<td>0.102</td>
<td>0.752</td>
<td>0.004</td>
<td>0.061</td>
</tr>
<tr>
<td>Barbell</td>
<td>Right Knee</td>
<td>1.25</td>
<td>51.555</td>
<td>&lt;.001*</td>
<td>&lt;.001</td>
<td>1</td>
</tr>
<tr>
<td>Analyses</td>
<td>Joint</td>
<td>df</td>
<td>F</td>
<td>Sig</td>
<td>PES</td>
<td>PWR</td>
</tr>
<tr>
<td>Barbell•VPAC</td>
<td>Right Ankle</td>
<td>1.25</td>
<td>1.393</td>
<td>0.248</td>
<td>0.053</td>
<td>0.206</td>
</tr>
<tr>
<td>VPAC</td>
<td>Right Ankle</td>
<td>1.25</td>
<td>4.554</td>
<td>0.043</td>
<td>0.154</td>
<td>0.537</td>
</tr>
<tr>
<td>Barbell</td>
<td>Right Ankle</td>
<td>1.25</td>
<td>4.147</td>
<td>0.052</td>
<td>0.142</td>
<td>0.499</td>
</tr>
</tbody>
</table>

Barbell = front or back barbell position, VPAC = yes/AB or no/BB, Joint; df = degrees of freedom, F = f-statistic, Sig = significance, PES = partial eta squared effect size, PWR = power. (significance was familywise adjusted to α = .0125 for lumbar, R hip, R knee and R ankle); *significant result.

ance, compensating with increased lumbar extension. As a result, the IO and EO may be elongated into a mechanically disadvantageous position and therefore demonstrate decreased EMG activity. As previously stated, slower performance of these movements decreases the likelihood of detrimental forces occurring at the lumbar spine and LE. Future studies should further investigate the oblique muscles’ role in trunk stabilization during squatting activities and the influence of VPAC.

STUDY LIMITATIONS, DELIMITATIONS, AND FUTURE RESEARCH

The results of this study must be interpreted considering some limitations. As discussed above, it is likely our subjects experienced a ceiling effect with respect to trunk muscle activity during the squat lifting sequence. It is possible there was no significant muscle activity change in response to VPAC because the load subjects were under resulted in a trunk muscle contraction regardless of VPAC condition.

Another limitation centers on the brief time subjects were given for executing a new skill. Subjects were introduced to VPAC performance shortly before performing both squat variations. Though provided with clear instructions for VPAC performance and verified by an expert, it is possible the subjects had not mastered this skill prior to squat performance. Hall et al. reported one VPAC training session may not be sufficient to improve EO, IO, and ES muscle EMG outputs. In future studies, two separate robust practice sessions may be beneficial for enhancing VPAC performance.

This study presents with three delimitations. First, subjects included a male convenience sample, which is consistent with other investigators. However, inclusion of females would improve generalizability of findings. Moreover, the use of a convenience sample risks having an overly homogenous sample, or one with significant variability, impeding true representation of the greater population. Future studies should attempt other sampling methods for improved results generalizability.

Secondly, weighted barbell squats were a fairly specialized movement involving a lifting activity at a predicted 75% of IRM. Comfort et al. discussed the need for testing subjects at higher training loads to adequately represent the athletic population. However, such higher loads may not be representative of the weight level used by the general population. Future studies should examine similar parameters at multiple different weight levels, including higher levels associated with competitive athletes.

Finally, inferences cannot be drawn between these findings and other functional lifting activities, such as those found in the industrial setting. For example, most lifting tasks in industrial settings involve loads that are lifted and/or carried in front of the body. While, the barbell BackS has little application to industrial lifting tasks, the barbell FrontS exhibits potential applicability. To support generalizing to the industrial population, future research should examine the effects of VPAC on functional FrontS lifting activities, deadlifting, box lifting, and farmer’s carry lifts that better represent industrial lifting.

CONCLUSION

The results of this study indicate that trunk muscle activity was higher during FrontS versus BackS regardless of VPAC condition. Regarding VPAC, this study demonstrated increased time of performance for both squat variations and improved ability to maintain a neutral lumbar spine. Both adaptations have been associated with decreased detrimen-
tal lumbar spine and knee forces. These findings can help guide clinicians and coaches to incorporate weighted FrontS and VPAC strategies into treatments and/or training programs.

CONFLICT OF INTEREST

Dr. Sizer is the co-founder of TKQuant LLC. This relationship/patent has nothing to do with this submitted work. All other authors declare no conflicts of interest.

Submitted: February 27, 2025 CDT, Accepted: June 16, 2023

CDT
Table 7. One-way ANOVA tests of Within-Subjects effects for sagittal plane joint moments

<table>
<thead>
<tr>
<th>Analyses</th>
<th>Joint</th>
<th>df</th>
<th>F</th>
<th>Sig</th>
<th>PES</th>
<th>PWR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barbell</td>
<td>Right Hip</td>
<td>1, 25</td>
<td>0.361</td>
<td>0.554</td>
<td>0.014</td>
<td>0.089</td>
</tr>
<tr>
<td>VPAC</td>
<td>Right Hip</td>
<td>1, 25</td>
<td>0.078</td>
<td>0.782</td>
<td>0.003</td>
<td>0.058</td>
</tr>
<tr>
<td>Barbell</td>
<td>Right Hip</td>
<td>1, 25</td>
<td>86.734</td>
<td>&lt; .001*</td>
<td>0.776</td>
<td>1</td>
</tr>
<tr>
<td>Analyses</td>
<td>Joint</td>
<td>df</td>
<td>F</td>
<td>Sig</td>
<td>PES</td>
<td>PWR</td>
</tr>
<tr>
<td>Barbell</td>
<td>Right Knee</td>
<td>1, 25</td>
<td>0.23</td>
<td>0.636</td>
<td>0.009</td>
<td>0.075</td>
</tr>
<tr>
<td>VPAC</td>
<td>Right Knee</td>
<td>1, 25</td>
<td>3.27</td>
<td>0.083</td>
<td>0.116</td>
<td>0.413</td>
</tr>
<tr>
<td>Barbell</td>
<td>Right Knee</td>
<td>1, 25</td>
<td>2.231</td>
<td>0.148</td>
<td>0.082</td>
<td>0.301</td>
</tr>
<tr>
<td>Analyses</td>
<td>Joint</td>
<td>df</td>
<td>F</td>
<td>Sig</td>
<td>PES</td>
<td>PWR</td>
</tr>
<tr>
<td>Barbell</td>
<td>Right Ankle</td>
<td>1, 25</td>
<td>1.058</td>
<td>0.314</td>
<td>0.041</td>
<td>0.167</td>
</tr>
<tr>
<td>VPAC</td>
<td>Right Ankle</td>
<td>1, 25</td>
<td>2.839</td>
<td>0.104</td>
<td>0.102</td>
<td>0.367</td>
</tr>
<tr>
<td>Barbell</td>
<td>Right Ankle</td>
<td>1, 25</td>
<td>76.201</td>
<td>&lt; .001*</td>
<td>0.753</td>
<td>1</td>
</tr>
</tbody>
</table>

Barbell = front or hack barbell position, VPAC = yes/AB or no/NB; df = degrees of freedom, F = f-statistic, Sig = significance, PES = partial eta squared effect size, PWR = power. (significance was familywise adjusted to α = .0167 for R hip, R knee and R ankle); *significant result.
REFERENCES


Effect of Footwear Versus Barefoot on Double-Leg Jump-Landing and Jump Height Measures: A Randomized Cross-Over Study

Kim Hébert-Losier1, Caleb Boswell-Smith1, Ivana Hanzlíková2

1 Te Huataki Waiora School of Health, University of Waikato, 2 Department of Physiotherapy, Palacký University Olomouc

Keywords: Anterior Cruciate Ligament, biomechanics, knee, screening

Background
Assessing individuals in their own athletic footwear in clinics is common, but can affect movement, performance, and clinical measures.

Purpose
The aim was to compare overall Landing Error Scoring System (LESS) scores, injury risk categorization, specific LESS errors, and jump heights between habitual athletic footwear and barefoot conditions.

Study design
Randomized cross-over laboratory study.

Methods
Eighty healthy individuals (55% male) completed the LESS following standard procedures (i.e., land from a 30-cm box to a distance of 50% of body height and then jump upwards maximally). Participants performed the LESS three times in two randomized conditions: footwear and barefoot. LESS data were extracted from 2D videos to compare group-level mean LESS scores, group-level and individual-level injury risk categorization (5-error threshold), specific landing errors, and jump heights between conditions.

Results
LESS scores were significantly greater (0.3 errors, p=0.022) and jump heights were significantly lower (0.6 cm, p=0.029) in footwear than barefoot, but differences were trivial (d = 0.18 and -0.07, respectively) and not clinically meaningful. Although the number of high injury-risk participants was not statistically different at a group level (p=1.000); 27 individuals (33.8%) exhibited a clinically meaningful difference between conditions of one error or more in LESS score, categorization was inconsistent for 16.3% of individuals, and four of the 17 landing errors significantly differed between conditions.

Conclusion
At a group level, habitual athletic footwear does not meaningfully influence LESS scores, risk categorization, or jump height. At an individual level, footwear can meaningfully affect LESS scores, risk categorization, and alter landing strategies. Use of consistent protocol and footwear is advised for assessing movement patterns and injury risk from the LESS given the unknown predictive value of this test barefoot.

Level of Evidence
Level 3.
©The Author(s)
INTRODUCTION

Anterior cruciate ligament (ACL) injury is one of the most common sport injuries and has a devastating influence on the activity levels and quality of life of individuals.\(^1\) ACL injuries can occur without physical contact, and thus, are considered preventable.\(^2\) The most common situation for noncontact ACL injuries appears to be deceleration, which is when the athlete cuts, changes direction, or lands from a jump.\(^3\) The Landing Error Scoring System (LESS) is a screening tool used to identify athletes presenting with high injury-risk movement patterns during a double-leg jump-landing (DLJL) task.\(^5\) Clinicians score 17 items based on movements during the DLJL task. The overall LESS score ranges from 0 to 17 errors where lower scores reflect fewer landing errors and thus fewer movement patterns linked with noncontact ACL injuries. Scores of five or more errors indicate poor jump-landing technique\(^4\) and have been linked to higher risk of ACL injury.\(^5\) Specifically, the risk ratio for sustaining a noncontact or indirect contact ACL injury was 10.7 in individuals scoring five errors or more compared to less than five errors.\(^5\) Compared to healthy controls, LESS scores are greater after an unilateral ACL reconstruction despite being cleared to return to physical activity, indicating that the LESS may provide useful information to guide rehabilitation and return-to-sport decisions.\(^6\)

Across the literature, the overall LESS score demonstrates good-to-excellent reliability, and moderate-to-excellent validity versus 3-dimensional (3D) motion capture data for the items linked with risk factors for knee injury.\(^7\) Overall LESS scores are, however, sensitive to various factors, such as gender, previous injury, and intervention programs.\(^8\) Research has also identified that jump landing distance\(^9\) and final LESS score computational method\(^10\) can affect LESS scores and individual-level risk categorization. Altogether, these studies highlight that several factors and procedural methods need to be considered when administering and interpreting LESS outcomes.

Footwear plays a central role in sport and is typically designed to enhance performance and protect the body from injury. However, footwear can influence the human-ground interaction and result in different multi-joint landing strategies to moderate impact forces.\(^11\) Indeed, initial contact from a 30-cm DLJL task similar to the LESS has been associated with a more plantar-flexed ankle,\(^12\) greater foot-ground angle,\(^12\) and smaller knee range of motion\(^13\) when performed barefoot compared to with shoes. Together, these studies indicate that LESS scoring might differ between barefoot and footwear conditions, although this topic has not yet been explicitly examined. Although the LESS is traditionally performed wearing a person’s own athletic shoes,\(^6,10\) it has also been conducted barefoot.\(^14\) Therefore, the aim was to compare overall LESS scores, injury risk categorization, and specific LESS errors between habitual athletic footwear and barefoot conditions. The hypothesis was that wearing footwear would result in higher overall LESS scores, lead to a greater number of individuals classified at high risk of injuries, and influence specific LESS errors compared to barefoot. Given how footwear can influence jump performance,\(^15\) jump heights from flight times were also compared between conditions.

MATERIALS AND METHODS

EXPERIMENTAL APPROACH

A randomized cross-over experimental design was used to explore the influence of footwear on LESS scores, LESS injury risk categorization, specific LESS errors, and jump height. Sample size calculations were performed \textit{a priori} using G*Power 3.1.9.7, and applied a standard two-tailed hypothesis, 90% power ($\beta = 0.10$), 5% significance level ($\alpha = 0.05$), one error LESS difference in paired means defining a clinically-meaningful change,\(^4,8\) and 2.47 standard deviation of the difference in paired means based on previous work implementing similar testing procedures and comparing LESS scores between two experimental conditions.\(^9\) Based on these assumptions, 67 participants were required and would be sufficient to detect a small effect size difference (Cohen $d = 0.40$) between conditions. A sample size of 80 participants was targeted to account for a 20% drop out rate.

PARTICIPANTS

Eighty participants were recruited and tested within one month in 2021 from a convenience sample of healthy university students. All volunteers were free of injury, illness, or conditions that may have affected their movements or landing mechanics. Participants with a lower extremity, back, or pelvis injury in the prior three months were excluded. LESS testing was performed in individuals’ own athletic footwear, as is typical in research and clinical settings.\(^6,10\) Participants were excluded when their footwear scored 70% or more on the Minimalist Index\(^16\) (described under Procedures) as deemed to represent minimal shoes\(^17\) that could potentially mimic barefoot.\(^18\) It was deemed inappropriate to merge data from trials performed in conventional athletic footwear to those from minimal footwear given the reported effect of these different footwear types on the biomechanics of dynamic tasks.\(^18-20\) All participants signed an informed consent document that explained the potential risks of participation (e.g., chance of injury due to physical activity). The University of Waikato Human Research Ethics Committee (HREC(Health)#2017-41) approved the protocol before data collection, which adhered to the Declaration of Helsinki. This project was retrospectively registered with the Australian New Zealand Clinical Trials Registry (ACTRN12622001558730).

PROCEDURES

Following informed consent, baseline characteristics of participants were collected, which included measuring body height using a stadiometer (seca model 0123, Medical Measuring Systems and Scales, Mount Pleasant, South Carolina) and mass on an electronic scale (seca model ESE813, Medical Measuring Systems and Scales, Mount Pleasant, South Carolina). Participants also completed a short sport
with sagittal frontal to Sony the tested order formed corresponding ately. body distance high were 1

five sured footwear of参与 the testing. All participants were pre-informed of the study aims and asked to bring their own athletic footwear for testing. Footwear characteristics were measured for all participants and included the use of the Minimalist Index alongside more traditional characteristics. In summary, the Minimal Index measures five shoe features to quantify the level of minimalism of footwear, where 100% represents the highest degree of minimalism. The five characteristics are footwear mass, longitudinal and torsional flexibility, stack height, heel-to-toe drop, and the presence/absence of technologies. Minimal Index scores of participants’ own shoes ranged from 4 to 64%. The hardness of the midsole material in the center of the heel region was assessed using an Asker-C durometer (Supertech Precision Supply Co., LTD, Osaka, Japan) with an accuracy of 1 unit. The average of three consecutive durometer measurements was recorded and used to quantify Asker-C heel hardness.

All experimentation took place in a biomechanics laboratory. The original LESS testing and scoring procedures were used, except in the barefoot condition when no shoes were worn. Participants jumped horizontally from a 30-cm box to 50% of their body height and jumped vertically as high as possible upon landing. The horizontal landing distance was indicated on the floor using tape. Trials were disregarded when participants did not land at 50% of their body height or did not perform the task in one fluid motion. Feedback on performance was not given to avoid influencing outcomes unless the task was performed inappropriately. Before the formal tests, participants were allowed up to three familiarization trials in both the footwear and barefoot conditions immediately before testing for each corresponding condition. For testing, each participant performed three trials in each condition with 30 seconds rest between trials and 15 minutes rest between conditions. The order of conditions was block randomized prior to study commencement by a third party to ensure an equal number of participants starting in each condition. The condition tested first (barefoot or footwear) was allocated sequentially and announced to participants upon study enrollment. It was not possible to blind the participants and examiners to the condition examined.

Two cameras with a focal length of 8.8 to 75.3 mm (35-mm equivalent focal length of 24-200 mm) captured the DJLJ trials at 120 frames per second (Sony RX10 II, Sony Corporation, Tokyo, Japan). These videos were used to derive LESS scores post testing. One camera captured frontal plane movement and the other captured right-side sagittal plane movement. Each camera was placed 3.5 meters away from the landing area and mounted on tripods with a 1.5 m lens-to-ground distance. The videos were analyzed using Kinovea (version 0.9.4, www.kinovea.org). The time from take-off from the ground to the final landing was extracted from the sagittal plane videos to compute jump heights from flight times as:

$$h = \frac{1}{8} \cdot g \cdot t^2 \cdot 100$$

where $h$ is jump height (cm), $g$ is gravitational acceleration constant (9.81 m/s²), and $t$ is flight time (s).

DATA PROCESSING

A single rater (CBS) with over three years of experience analyzing human movement conducted all data processing after receiving four training sessions from an expert LESS rater (IH) who had completed over 400 LESS evaluations. The single rater completed more than 20 LESS assessments before analyzing the current dataset. After analyzing all videos for this study (i.e., 80 participants x 3 trials x 2 conditions = 480 videos), the rater re-analyzed the first 20 to ensure consistency in ratings. The rater was blinded to the randomization sequence and LESS scores of individuals from the other experimental condition, as trials were presented in a random order for rating.

To ensure rater reliability of the videos collected, two raters (CBS and DB) with similar experience and LESS training participated in an inter-rater and intra-rater reliability study of the overall LESS score using a subset of videos from 10 participants. Inter-rater reliability was excellent based on intra-class correlation coefficient (ICC) and 95% confidence interval (lower, upper) values for both footwear (ICC(2,1) = 0.957 [0.815, 0.990]) and barefoot (ICC(2,1) = 0.957 [0.847, 0.989]) conditions. Intra-rater reliability was also excellent for both footwear (ICC(3,1) = 0.974 [0.903, 0.993]) and barefoot (ICC(3,1) = 0.970 [0.815, 0.993]) conditions.

STATISTICAL ANALYSIS

The effect of footwear on group mean LESS scores, injury risk categorization (high risk, LESS > 5 errors; low risk, LESS < 5 errors), individual-level risk categorization, and jump height was examined. The average of participants’ three trials was used for analysis. Taking the average of three trials is consistent with the original LESS protocol and is the most common approach used to interpret LESS data. Differences in group mean LESS scores and jump heights between conditions were assessed using mean differences, two tailed paired $t$-tests, and Cohen’s $d$ effect sizes for paired samples using an average variance with 95% confidence intervals. Cohen’s $d$ effect sizes were considered small, medium, and large when reaching 0.20, 0.50, and 0.80, respectively, and trivial when less than 0.20.

Differences in the number of participants categorized at high and low risk of injury based on the 5-error LESS threshold between conditions were assessed using McNemar’s tests and odds ratio with 95% confidence intervals. The odds ratio reflects the number of participants exclusively at high risk in the footwear condition versus those exclusively at high risk in the barefoot condition. Hence, odds ratios > 1 reflect a higher proportion of at-risk individuals in the footwear condition. The number of participants demonstrating a clinically meaningful change in LESS scores (i.e., one error or more difference) between conditions was also examined. Finally, differences in the
occurrence of specific LESS errors between conditions were explored using McNemar’s tests. For each participant, an error was considered present when present in two of the three trials for Items 1-15. For Items 16-17, an error was considered present when the ‘average’ rating was present in two of three trials or when the ‘poor/stiff’ rating was present in one of three trials.4,9 The significance level was set at p≤0.05 for all analyses, which were conducted using Microsoft Excel for Microsoft 365 (version 2109), Microsoft Corp, Redmond, WA, USA) and RStudio® version 1.1.463 with R version 4.0.5 (R Core Team, 2021). There were no missing data, and all participants completed the experimentation without harm.

RESULTS

Eighty participants (44 males and 36 females) completed the study. Their demographic and footwear characteristics are presented in Table 1. Approximately half of participants (52.5%) participated in court or field sports (e.g., basketball, football, netball, rugby), with most of the others (42.5%) participating in another sporting activity (e.g., running, cycling, rowing).

The group mean LESS scores in the footwear condition (range: 2.7-10.0 errors) was significantly greater (0.5 errors, p=0.022) than barefoot (range: 2.3-10.0 errors), as shown in Table 2. However, the magnitude of the difference was trivial (Cohen’s d=0.18 [0.03, 0.33]). The number of individual classified at high risk was not significantly different between conditions (62 participants footwear vs 61 participants barefoot, p=1.000), with seven participants categorized at high risk exclusively in footwear and six barefoot (Figure 1). At an individual level, 27 participants (33.8%) demonstrated a clinically meaningful difference of one error or more in LESS scores between conditions. The risk categorization was conflicting between conditions for 13 participants (16.3%, Figure 2). Six participants changed from being categorized as low risk in footwear to high risk barefoot, and seven from high risk in footwear to low risk barefoot. The difference in mean LESS score was one or more in all but one of these participants (92.3%). Jump height in footwear (range: 8.5-56.4 cm) was significantly lower (<0.6 cm, p=0.029) than barefoot (range: 11.4-55.1 cm), but the difference was trivial (d = 0.07 [-0.13, -0.01], Table 2).

The occurrence of specific LESS errors significantly differed between conditions for four of the 17 items. Specifically, there were more errors for Item 4 (ankle plantar flexion at initial contact) and Item 5 (knee valgus at initial contact) in footwear, and more errors for Item 8 (stance width-narrow) and Item 10 (foot position-toe out) barefoot (Table 3).

DISCUSSION

In agreement with the hypothesis, footwear led to significantly higher LESS scores than barefoot; however, the difference was trivial and not clinically meaningful as it was less than one error.6,8 Footwear led to significantly lower jump heights than barefoot, but the difference was also trivial and not clinically meaningful as it was less than the 2 cm typical error associated with this measure.25 A greater number of participants at high risk of injury when wearing footwear was hypothesized; however, the number of high injury-risk participants was not significantly different to barefoot. Despite the similarities in LESS scores and high injury-risk categorization at a group level, differences in LESS scores were clinically meaningful (i.e., one error or more) for approximately one third of participants, and individual-level risk categorization was incon-
Table 2. Comparison of Landing Error Scoring System (LESS) mean scores and group-level injury risk categorization between footwear and barefoot conditions. Data are reported as means ± standard deviations and differences with 95% confidence intervals [lower, upper].

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Footwear</th>
<th>Barefoot</th>
<th>Difference</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>LESS score (errors)</td>
<td>6.2 ± 1.5</td>
<td>5.9 ± 1.6</td>
<td>0.3 [0.05 to 0.52]</td>
<td>0.022†</td>
</tr>
<tr>
<td>High injury risk (%)</td>
<td>77.5%</td>
<td>76.3%</td>
<td>1.17 [0.39 to 3.47]</td>
<td>1.000‡</td>
</tr>
<tr>
<td>Jump height (cm)</td>
<td>32.1 ± 9.2</td>
<td>32.8 ± 8.9</td>
<td>-0.6 [-0.1, -1.2]</td>
<td>0.029†</td>
</tr>
</tbody>
</table>

Note. †Significant difference between conditions (p ≤ 0.05) are in bold. ‡Difference in means with paired t-test. ‡Odds ratio significance with McNemar’s test.

Figure 1. Venn diagrams representing participants at high (>5 errors) and low (<5 errors) injury risk for both footwear and barefoot conditions. The number in the circle represents the sum of participants categorized at low or high risk for each condition. The overlapping area represents the number of participants at low or high risk in both conditions. "Significant difference in the proportion of individuals at high and low risk based on McNemar’s tests (p ≤ 0.05).

consistent for approximately a sixth of participants between conditions. Furthermore, differences in specific landing errors were noted, with greater odds of knee valgus and heel-to-toe or flat foot landing at initial contact in footwear, and lesser odds of landing with a narrow stance width and toe-out foot position. Overall, performing the LESS with compared to without footwear led to comparable mean LESS scores, group-level injury risk categorization, and jump heights, but influenced specific LESS errors, individual-level risk categorization (i.e., 16.3% of individuals inconsistently categorized between conditions), and LESS scores of some participants in a clinically meaningful manner (i.e., change of one error or more for 33.8% of individuals).

The mean LESS scores in footwear in this study are similar to means reported elsewhere for similar cohorts of young active individuals.9,10 The current findings also reflect previous ones where altering the jump landing distance of the LESS did not meaningfully affect group-level LESS scores and risk categorization, but significantly influenced the odds of individual LESS errors and individual-level injury risk categorization.9 The comparable outcomes imply that studies can implement the LESS either with shoes or barefoot when the main outcome is the group mean LESS score or group-level injury risk categorization. Implementing the LESS barefoot can be easier to standardize across participants as guarantees no effect of footwear or footwear type on landing mechanics. Nonetheless, it would be inappropriate to compare specific LESS errors between studies or infer similar risk of injury at an individual level between conditions. For instance, O’Malley, Murphy performed the LESS barefoot.14 Their results would likely be comparable if performed with shoes in terms of the group mean LESS score and proportion of high injury-risk individuals, but the individual-level risk categorization might differ. Furthermore, the predictive value of the LESS performed barefoot for noncontact ACL injury has not been researched. Hence, when using the LESS in a clinical setting, test parameters should be kept constant for a given
Figure 2. Landing Error Scoring System (LESS) score plots for both footwear and barefoot conditions for all 80 participants. The dashed grey dotted line represents the identity line. The dashed black lines represent the 5-error threshold that defines high (>5 errors) and low (<5 errors) injury risk.

Table 3. Landing Error Scoring System (LESS) specific errors for 80 participants.

<table>
<thead>
<tr>
<th>No.</th>
<th>Items</th>
<th>Number of errors (% participants)</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Footwear</td>
<td>Barefoot</td>
</tr>
<tr>
<td>1</td>
<td>Knee flexion at initial contact</td>
<td>57 (71.3%)</td>
<td>50 (62.5%)</td>
</tr>
<tr>
<td>2</td>
<td>Hip flexion at initial contact</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>3</td>
<td>Trunk flexion at initial contact</td>
<td>1 (1.3%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>4</td>
<td>Ankle plantar flexion at initial contact</td>
<td>14 (17.5%)</td>
<td>4 (5.0%)</td>
</tr>
<tr>
<td>5</td>
<td>Knee valgus at initial contact</td>
<td>69 (86.3%)</td>
<td>59 (73.8%)</td>
</tr>
<tr>
<td>6</td>
<td>Lateral trunk flexion at initial contact</td>
<td>6 (7.5%)</td>
<td>8 (10.0%)</td>
</tr>
<tr>
<td>7</td>
<td>Stance width (wide) at initial contact</td>
<td>12 (15.0%)</td>
<td>11 (13.8%)</td>
</tr>
<tr>
<td>8</td>
<td>Stance width (narrow) at initial contact</td>
<td>25 (31.3%)</td>
<td>32 (40.0%)</td>
</tr>
<tr>
<td>9</td>
<td>Foot position (toe-in)</td>
<td>0 (0%)</td>
<td>1 (1.3%)</td>
</tr>
<tr>
<td>10</td>
<td>Foot position (toe-out)</td>
<td>12 (15.0%)</td>
<td>26 (32.5%)</td>
</tr>
<tr>
<td>11</td>
<td>Symmetric foot contact at initial contact</td>
<td>53 (66.3%)</td>
<td>52 (65.0%)</td>
</tr>
<tr>
<td>12</td>
<td>Knee flexion at maximal knee flexion</td>
<td>5 (6.3%)</td>
<td>5 (6.3%)</td>
</tr>
<tr>
<td>13</td>
<td>Hip flexion at maximal knee flexion</td>
<td>1 (1.3%)</td>
<td>1 (1.3%)</td>
</tr>
<tr>
<td>14</td>
<td>Trunk flexion at maximal knee flexion</td>
<td>3 (3.8%)</td>
<td>2 (2.5%)</td>
</tr>
<tr>
<td>15</td>
<td>Knee valgus displacement</td>
<td>60 (75.0%)</td>
<td>56 (70.0%)</td>
</tr>
<tr>
<td>16</td>
<td>Joint displacement</td>
<td>55 (68.8%)</td>
<td>58 (72.5%)</td>
</tr>
<tr>
<td>17</td>
<td>Overall impression</td>
<td>76 (95.0%)</td>
<td>70 (87.5%)</td>
</tr>
</tbody>
</table>

Note. *Significant difference between conditions (p ≤ 0.05) are in bold. †McNemar’s test p-values for differences between conditions.

Individual on separate occasions and the use of footwear is recommended given that it has demonstrated predictive value in youth. Most ACL injuries that occur during sports and recreational activities probably involve individuals wearing their own athletic footwear; hence, performing the LESS with shoes is arguably more ecologically valid. In biomechanics research, relying solely on null hypothesis significance testing without use of appropriate effect sizes or consideration of the magnitude of the difference is
discouraged. It has been proposed that a one error change in LESS score is clinically meaningful. In fact, the inter-session standard error of measurement for the LESS is 0.81 error, which exceeds the observed difference of 0.3 errors between footwear and barefoot conditions. Hence, although the difference between conditions reached statistical significance, the effect of footwear on overall mean LESS scores is not clinically meaningful. Despite this, differences of one error or more were observed in 33.8% of individuals and changes in injury risk categorization in 16.3% of individuals between footwear and barefoot conditions, again supporting use of a consistent footwear or barefoot protocol for a given individual when assessing injury risk or movement strategies over time.

The odds of errors significantly differed between footwear and barefoot conditions for four LESS items: knee valgus, ankle plantar flexion, narrow stance width, and toe-out foot position at initial contact (Items 4, 5, 8, and 10). The two first errors were 12.5% more prevalent in footwear, whereas the latter two were 8.7% and 17.5% less prevalent. Arguably, though, differences less than 10% are likely trivial in nature (i.e., narrow stance width). Hanzlíková and Hébert-Losier also found that these specific LESS errors differed between self-selected and 50% body height landing conditions, alongside knee valgus displacement (Item 15). Furthermore, review of the running literature indicate that footwear influences knee, ankle, and stride kinematics. A heel-to-toe drop of zero, for instance, is more commonly associated with a forefoot strike pattern in running studies compared to running in footwear with a drop of 8 mm or more. These findings combined suggest these specific LESS errors (Items 4, 5, 8, and 10) are more sensitive to change and alterations in protocol and footwear than the other errors. The differences in likelihood of specific LESS errors between footwear and barefoot conditions indicate differences in multi-joint strategies used to moderate impact forces during landing tasks, as shown elsewhere. Barefoot, participants were more likely to land with greater ankle plantar flexion and the front part of their foot. These observations are comparable to findings of a more plantar-flexed ankle and greater foot-ground angle at initial contact from a 30-cm DLJL task similar to the LESS when performed barefoot compared to with shoes. Landing in greater ankle plantar flexion during DLJL likely shifts loading between joints, with greater ankle but lesser knee joint loading. Indeed, participants with an ACL reconstruction landed from a 60-cm drop with greater ankle plantar flexion and absorbed a greater amount of force at the ankle compared to non-injured controls, presumably to protect their injured knee. Furthermore, research also indicates that single-leg landing with greater ankle plantar flexion from a drop jump increases total energy dissipation and reduces peak vertical loading rates. Since landing in greater plantar flexion may reduce the risk of knee and hip injuries, DLJL barefoot may be considered as a training tool in the early stages of ACL injury rehabilitation to reduce knee loads and peak vertical loading rates. In addition, our data indicate that maximal jump performance is not compromised barefoot, which is often of concern to coaches, clinicians, and athletes.

Although knee valgus at initial contact was one of the most frequent errors in both footwear and barefoot conditions, this error was 12.5% more prevalent in footwear. Previous research has identified knee valgus as a risk factor for ACL injury. These nine athletes all exhibited increased knee valgus when performing drop vertical jumps pre-injury. Therefore, this metric alone in the context of the LESS might suggest an increased ACL injury risk when wearing footwear compared to barefoot. However, knee valgus alone does not cause ACL injury. ACL injuries are moreover linked with multi-planar mechanisms, often with a hyperextended of slightly flexed knee undergoing a valgus motion with either internal or external rotation. Despite overt methodological limitations, more recent research continues to challenge that knee valgus during drop jumps is a valid predictor of ACL injury, with no association between 2D frontal plane knee and hip motion during drop jumps and noncontact ACL injuries.

In the current study, a threshold of five or more errors was used to categorize participants at high injury risk based on previous research. However, the predictive value of the LESS is debated in research given other studies indicating a lack of association between LESS scores and noncontact ACL injury. Noteworthy is that in these two studies, photographs of participants suggest performance of the LESS in shoes in one study and barefoot in the other, which might have influenced LESS scores at an individual level. The five-error threshold may be appropriate in footwear only. Furthermore, there is no population-specific LESS cut-off score established in the literature. For instance, there is a tendency in the literature for higher LESS scores in younger individuals. Hence, it remains to confirm whether the five-error threshold established from youth elite soccer players (age: 13.9 ± 1.8 y) apply to young active adults like those in the current study (age: 20.0 ± 2.3 y) in whom the LESS is often used. Nonetheless, over 75% of participants were categorized as high risk, which could reflect the inappropriate nature of the 5-error threshold in this cohort or the fact that most participants were not involved in jump-landing sports. Non-contact ACL injuries are multifactorial in nature, with the LESS examining gross movement patterns only. It is also worth noting that a series of studies suggest that the vertical drop jump and DLJL tasks are poor predictors of future ACL injury. Out of five biomechanical variables examined across these studies (knee valgus angle at initial contact, peak knee abduction moment, peak knee flexion angle, peak vertical ground reaction force, and medial knee displacement), only medial knee displacement during the drop vertical jump was linked to ACL injuries prospectively, but sensitivity (0.6) and specificity (0.6) were poor. In recent investigations, the ability to control the knees in the frontal plane during
landing from a DLJL was unable to distinguish between athletes who sustained an ACL injury to those who remained uninjured.\textsuperscript{40} Despite these findings, DLJL tasks can still be useful as part of neuromuscular training programs for reducing ACL injury incidence\textsuperscript{41} and guiding rehabilitation or return-to-sport decision making post ACL reconstruction.\textsuperscript{5, 42} The LESS can also be useful for monitoring the effectiveness of programs and changes in biomechanical patterns.\textsuperscript{43} Performing the DLJL in footwear and barefoot likely involves different multi-joint strategies, loads, and muscle recruitment and activation patterns, which might ultimately lead to different adaptations. As such, performing DLJL tasks in both footwear and barefoot within neuromuscular training programs could provide different stimuli to individuals. Given that participants wearing minimal footwear were excluded, the generalization of the current findings comparing DLJL measures between barefoot and different types of footwear needs confirmation.

**CONCLUSION**

Overall LESS scores were significantly greater and jump heights were significantly lower in footwear than barefoot, but differences were trivial and not clinically meaningful. At the group level, the proportion of participants categorized at high risk of injury was comparable between conditions; however, differences in specific landing errors, inconsistency in injury risk categorization, and clinically meaningful changes in LESS scores at an individual level were noted. In clinical settings or for screening purposes, performing the LESS with shoes is still recommended given that the predictive value of the LESS barefoot has not been established. If the DLJL is used in neuromuscular training programs, performing the task both with and without shoes can offer variety in landing strategies and potentially different stimuli and neuromuscular adaptations to individuals.

**CONFLICTS OF INTEREST STATEMENT**

The authors report no conflicts of interest.

**DATA AVAILABILITY**

The data that support the findings of this study are openly available in OSF at https://doi.org/10.17605/OSF.IO/KHS7V.\textsuperscript{44}

**ACKNOWLEDGMENTS**

The authors thank Dr. Shannon O’Donnell and Mr. Dalton Berry for their assistance during the data collection process. The authors also thank the participants for their voluntary participation.

Submitted: October 28, 2022 CDT, Accepted: May 06, 2023 CDT
REFERENCES


44. Hébert-Losier K. Landing Error Scoring System footwear dataset. OSF. Published online October 24, 2022. doi:10.17605/OSF.IO/KHS7V
Original Research

Concurrent Force Feedback on Load Symmetry in Total Knee Arthroplasty Patients

Steni Sackiriyas, Becky Heinert, Drew Rutherford, Gwen M Fritz, Thomas W Kernozek

1 Health Professions, University of Wisconsin–La Crosse, 2 Gundersen Health System, 3 La Crosse Institute for Movement Science, University of Wisconsin–La Crosse

Keywords: Retraining, Rehabilitation, Motor learning, Arthritis

International Journal of Sports Physical Therapy

Background and Purpose
Load asymmetry can be present before and after total knee arthroplasty (TKA), which may affect progress during knee rehabilitation in an outpatient sports medicine setting. Current rehabilitation primarily focuses on strength, pain, and range of motion deficits; however, recent evidence suggests the use of movement retraining strategies such as load feedback to address load asymmetry. Therefore, the purpose of this study was to examine how a single session of concurrent force feedback influences load symmetry during the leg-press and body-weight squat exercises in individuals following TKA. Additionally, a secondary purpose was to examine the retention of any changes over the course of a week.

Study design
Case-series study

Methods
This observational, repeated-measures study design examined the effect of concurrent force feedback training on the mean and standard deviation of load symmetry index during the leg press and squat exercises in 26 patients with TKA in an outpatient sports medicine clinic. The load asymmetry was measured with loadpad sensors placed underneath the each extremity during leg press and squat (baseline), after one training session consisting of concurrent force feedback during these exercises within a single physical therapy session (post feedback), and after seven to ten days of a washout period (post retention). Separate 2 x 3 repeated measures analysis of variance was used to compare the mean and standard deviation of load symmetry across exercise (leg press and squat) and across time (baseline, post feedback and post retention).

Results
There was a time effect for the mean load symmetry index (p=0.027) but not for the standard deviation (p=0.441) during these exercises. The leg press showed a greater mean symmetry index compared to the squat regardless of time (p=0.001).

Conclusions
A reduction in the mean load symmetry index following concurrent feedback training suggests improved use of the surgical limb during both leg press and squat exercises during the same therapy session but the more symmetric loading pattern was not retained one week later. Overall, the leg press showed greater mean asymmetry than the
INTRODUCTION

As knee OA progresses, individuals often progressively limit their participation in sport and recreational activities. In two separate reviews Barber-Westin & Noyes and Witjes et al. indicated that 34-100% of patients experiencing a TKA return to sport and recreational activities. It is unknown why such large differences exist in whether TKA patients return to sport and recreational activities. Also, despite the reported improvement in mobility and function of individuals following TKA in the literature, abnormal load symmetry (greater loading on the non-surgical limb compared to the surgical limb) has been reported to be present before and after surgical intervention. Numerous factors such as acquired knee pain on the surgical limb, quadriceps and hamstring weakness in both lower extremities, and acquired habitual movement patterns have been suggested as potential contributors to this abnormal loading symmetry following TKA. This abnormal load symmetry has been suggested to lead to contralateral knee and hip pain, and poor functional outcomes after TKA. For example, an increased load on the non-surgical limb following TKA could increase knee adduction moment and vertical ground reaction forces on the contralateral side that may lead to or accelerate knee osteoarthritis (OA) and the risk of surgical intervention on the contralateral limb.

Current rehabilitation protocols mostly focus on minimizing shear stimuli, avoiding excessive ligament strain in patients undergoing cartilage repair, minimizing pain, and improving strength, and range of motion in patients after TKA. However, these strategies may not address the habitual movement patterns and movement impairments such as abnormal load symmetry after TKA. Recent suggestions for comprehensive knee rehabilitation include the use of movement retraining strategies that utilize performance-based feedback to address these loading impairments. For example, concurrent feedback has been reported to improve performance during the acquisition phase of motor learning. Concurrent force feedback is a method where the total force from each extremity is displayed on a computer monitor during the performance, such as a leg press and squat. Forces shown may help the patient balance the forces traveling through the joints. Also, squat and leg press exercises, typically part of a current knee rehabilitation regimen, may address strength deficits but not address persistent movement impairments that involve reduced loading on the surgical limb during these exercises. Therefore, targeting and fostering greater load symmetry using concurrent force feedback during leg press and squat in the early stages of rehabilitation following TKA may be important. However, the use of force feedback to monitor and regulate load symmetry following TKA has not been extensively reported as part of standard rehabilitation programs.

Therefore, the purpose of this study was to examine how a single session of concurrent force feedback influences load symmetry during the leg-press and body-weight squat exercises in individuals following TKA. Additionally, a secondary purpose was to examine the retention of any changes over the course of a week. The hypothesis was that the load symmetry would improve in individuals with TKA immediately following concurrent force feedback training and persist after one week.

METHODS

STUDY DESIGN

In this repeated measures design, participants performed both body-weight squat and leg press, in a random order, as load symmetry were collected without providing feedback (baseline). Participants performed bodyweight squat in standing with self-chosen foot width. During the leg press, participants were seated with the seat adjusted to the proper length such that their knees could be fully extended and the leg press load was chosen based on their perceived exertion. Both exercises were performed at a set rate of movement and the body-weight squat was performed to 80° of knee flexion. The perceived exertion for both exercises were based on verbal indication by each participant not exceeding 5/10 on a Rated Perceived Exertion (RPE) scale (A scale that measures subjective reporting of effort that ranges from no exertion to maximum exertion experienced during an exercise or a physical activity) for all participants. No specific instructions were provided regarding body position or exercise form during these two exercises other than the rate and depth during baseline assessment.

Then, each participant received concurrent force feedback training (force feedback from each extremity) while performing both the leg press and body weight squat in a randomized order. They subsequently completed other weight-bearing exercises within a typical 20-minute physical therapy session without any feedback. Immediately after completion of this physical therapy session, participants were asked to perform the same leg-press and squat exercises while load symmetry data was collected with participants were blinded (load from each extremity was not shown or displayed) to concurrent force feedback (post feedback). Then, after seven to ten days of a washout period, the load symmetry data was collected similar to the baseline and post feedback testing conditions during the leg press and squat in a randomized order (post retention) prior to a standard physical therapy session.
PARTICIPANTS

Twenty-six patients (mean age 63.8 years, range 58–76) who desired to return to being recreationally active in sport and recreational activities after TKA were recruited as participants. Patients were included in this study if they were post TKA and were able to safely and independently perform an unsupported squat, ambulate without an assistive device, and complete seated leg press with a knee flexion angle of at least 80°. The mean time from surgery was 7.5 weeks (range from three weeks and two days to 14 weeks with no weight-bearing restrictions). All of the participants were actively participating in a supervised post-operative rehabilitation in an outpatient sports medicine setting at a regional medical center. Participants provided their written informed consent prior to participation in this study, and the study protocol was approved by the Ethics Committee on the use of human subjects in research at the University of Wisconsin – La Crosse. The IRB for this protocol was #17-TK-173.

INSTRUMENTATION

Two loadpad sensors (Novel GMBH, Munich, Germany), inserted in a protective cover, were placed on the ground or affixed to the base plate of a leg press machine (Eagle NX, Cybex, Rosemont, IL, USA) to measure and/or provide concurrent force feedback during the training session (Figure 1).

Force data from each extremity were collected using the sensors at 62.5 Hz that was streamed to an iPad (Apple Inc., Cupertino, CA, USA). Similar technology to the loadpad, the loadsol in-shoe sensors, developed by the same manufacturer, has been reported to have excellent reliability for determining peak ground reaction force with data collected simultaneously with a force platform for hopping (ICC=0.96), walking (ICC=0.89) and for running (ICC=0.92–0.94) on an instrumented treadmill as reported by Burns et al.22

PROCEDURE

Prior to baseline testing, subjects were provided with a verbal explanation and demonstration of the study protocol. For each testing scenario, the knee flexion angle was limited to 80° as measured by a hand-held goniometer. For the seated leg press, the machine was set to limit the knee flexion angle at 80° with a subject-specific weight based on their RPE. An adjustable stool was placed behind the subjects to maintain appropriate knee flexion angle while performing the body-weight squat. An auditory metronome (Seiko quartz metronome, Mahwah, NJ) was set at 36 beats per minute to constrain movement velocity to a set rate during both the leg press and squat for consistency between subjects. Each beat of the metronome occurred when the extremities were most flexed or extended. Participants were allowed several repetitions in each testing condition to familiarize themselves with the task. The subjects were not allowed to use any assistive device or their hands for support during the testing or training.

Baseline testing consisted of participants performing six repetitions of a body-weight squat and during self-selected,

Figure 1. A) Loadpad sensors used in the investigation were placed in a protective covering. B) During the squat performance and leg press performance, the loadpad sensors were placed under each foot. For the squat, the sensor was placed on the floor and for the leg press, it was affixed via double-sided tape to the foot plate of the leg press machine.
moderate-load, leg press exercises. The participants were blinded to their load symmetry during these two exercises and the order of exercises were randomized for each participant during this baseline assessment.

Then, all the participants performed one set of six repetitions of the leg press and body-weight squat exercises while receiving concurrent force feedback from the two loadpad sensors. The order of the exercises was randomly assigned. The force feedback data (from force measured each foot) were continuously displayed via a separate bar graph for each extremity measured in N (Figure 2) on the iPad display in front of the participant during each exercise so they could see the bar graph display.

The refresh rate on the iPad display from the loadpad sensors was 4.16 Hz. Before the training, the subjects were given verbal reminders to attempt to keep these two bars as equal as possible on the iPad which was placed on a stand in front of the participant during the squat and leg press exercise performance. After completing this feedback training, the participants completed 20 repetitions of each of the following supervised exercises during a standardized 20 minute physical therapy (PT) session: terminal knee extensions in standing, bipedal rocker board tilts forward-backward and side to side, and standing hip abduction bilaterally. Immediately after the 20-minute supervised PT session, the participants performed six repetitions of body-weight squat and leg press exercises in a randomized order without any force feedback while we collected load symmetry data (post feedback). After seven to ten days of a washout period, the load symmetry data were collected (post retention) similar to baseline and post feedback testing conditions with no force feedback prior to a standard physical therapy session.

STATISTICAL ANALYSIS

A sample size calculation performed using G-power (version 3.0.10, Germany) determined that at least a sample size of 16 was required for this study based on Zeni et al. 201325 based on the mean and standard deviations for vertical ground reaction force symmetry values from presurgical to six months postsurgical with alpha set to 0.05, beta to 0.2 and a calculated effect size of 1.07.

Force data were smoothed with a low-pass Butterworth filter (2nd order, dual-pass, 7 Hz cut-off frequency) and clipped to the beginning and end of the six squat or leg press repetitions from data recorded on the iPad. The load symmetry index (LSI) based on the bilateral force was determined based on the following formula as described elsewhere24:

\[ LSI = \frac{(\text{Force on the non-surgical limb} - \text{Force on the surgical limb})}{(\text{Force on the non-surgical limb} + \text{Force on the surgical limb})} \times 100\% \]

A load symmetry index of 0 indicates a perfect symmetry.9 All data processing were completed using the force values from each sensor within MATLAB® (Mathworks, Natick, MA). Mean load symmetry index is the mean value measured from the continuous force measurements from each force sensor at 62.5 Hz during the six repetitions of the squat and leg press exercises. Standard deviation of the load symmetry index was calculated and is used as an indicator of the variation in the load symmetry index during the continuous force measures from each sensor during the same timeframe for each exercise.

A 2 x 3 repeated-measures analysis of variance was performed on exercise (Squat, Leg press) and time (baseline, post feedback, post retention) to determine the mean and standard deviation of load symmetry index (α=0.05) using SPSS 28 (IBM, Inc., Armonk, NY). Follow-up tests using the Bonferroni procedure were performed as warranted. Partial eta2 was used to determine effect sizes within SPSS. A small effect was based on partial eta2 equal to 0.01, medium equal to 0.06, and large equal to 0.14.25

RESULTS

MEAN LOAD SYMMETRY INDEX

The mean and standard deviation of the load symmetry index was shown in Table 1 for each time point (baseline, post feedback, post retention) for leg press and squat exercises (Table 1). The mean load symmetry index for the squat showed a 42.7% decrease from baseline to post feedback but returned to baseline values post retention. The leg press exercise showed a 5.9% decrease in mean load symmetry index from baseline to post feedback but then showed an increase of 12.5% that depicted a regression in mean load symmetry index greater than the baseline. The mean load symmetry index was always greater due to greater loads being placed on the non-surgical limb (indicating an increased load compared to the non-surgical limb).

Figure 3 depicts the mean load symmetry index for all subjects across time points between the two exercises. There was an effect of time on the mean load symmetry in-

Figure 2. Concurrent feedback of force (N) displayed on the iPad in front of a representative subject during the squat and leg press exercises. Right indicates the right lower extremity and Left indicates the left lower extremity.
Table 1. Means ± standard deviation of load symmetry index (LSI) for leg-press and squat exercise across the time points. Values closer to zero indicates perfect symmetry.

<table>
<thead>
<tr>
<th>Movement Variable</th>
<th>Baseline</th>
<th>Post feedback</th>
<th>Post retention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Squat Mean LSI (%)</td>
<td>28 ± 16</td>
<td>16 ± 04</td>
<td>28 ± 22</td>
</tr>
<tr>
<td>Squat SD LSI (%)</td>
<td>22 ± 04</td>
<td>12 ± 02</td>
<td>12 ± 04</td>
</tr>
<tr>
<td>Leg Press Mean LSI (%)</td>
<td>34 ± 12</td>
<td>32 ± 12</td>
<td>36 ± 20</td>
</tr>
<tr>
<td>Leg Press SD LSI (%)</td>
<td>22 ± 06</td>
<td>20 ± 06</td>
<td>20 ± 06</td>
</tr>
</tbody>
</table>

Note: A load symmetry index of 0 indicates a perfect symmetry, as described in Christiansen et al."}
riods in this study. In other words, the participants relied more on the non-surgical limb during the leg-press than the squat exercise at all time periods. This finding may relate findings of Zheng et al., that reported a higher peak tibiofemoral compressive force (via analytical model) during a leg press compared to a squat exercise.26 Maybe to avoid increased tibiofemoral compressive force that could produce knee pain, the participants in this investigation relied more on the non-surgical limb for weight bearing during leg press than the squat. It is possible that one could manage their own body weight during squat but may not manage externally applied load during the leg press exercise. Another plausible reason for this difference in mean load symmetry may be imposed by the constraints necessary for maintaining postural control during each exercise. When one is seated during a leg press, a more similar contribution of force may not be required for success. However, when maintaining a standing position, postural control during a squat may require the performer’s center of gravity to be more centered between the legs compared to being seated during the leg press.

Regarding the absence of persistence of improved loading symmetry after the one week of washout period as evidenced by the return of load symmetry index to near baseline for both the exercises, this is not surprising and may be attributed to how concurrent feedback improves immediate practice performance but may not be sustained once the concurrent feedback has been removed.27,28 Chang et al.27 and Yamamoto & Ohashi28 suggested that considerable feedback must first be provided and then feedback frequency should be reduced to avoid dependency to promote long-term motor skill learning. A portion of this study results supported this training effect by showing an immediate increase in mean load symmetry immediately in a single session of training but then returned toward baseline levels after a short washout period.

Finally, the authors wanted to highlight that no observed difference in standard deviation of load symmetry index for both squat and leg press exercises over time periods indicated that minimal variation in the load symmetry performance pattern occurred in this study’s participants. This may be an indication of a repeatable pattern of loading symmetry in both exercises. Incorporating this study results and other literature related to the use of feedback suggest that using more practice may be needed with concurrent force feedback (increase the frequency of concurrent feedback training over more than one therapy session) to improve loading symmetry over time. Also, it may be then necessary to use different types of feedback (terminal: feedback provided as a summary after the completion of an action or performance and faded: reducing the feedback frequency over time) to translate the short-term improvements to a longer-term retention for improved function and return to sport and recreational activity.

Clinicians, such as physical therapists, may want to consider utilizing movement retraining strategies such as using concurrent force feedback training to promote improved load on the surgical limb following TKA when appropriate. With the dramatic projected increase of TKA from 85%29 in 2025 to 110%30 in 2050 to as high as 401% in 204031 and their association with future TKA in the uninvolved limb,6,13 the use of a movement retraining strategy with concurrent force feedback to address movement system impairments may be warranted.

STUDY LIMITATIONS

This study has several limitations, such as TKA subjects with varying Body Mass Index (BMI), variable time from surgery (three weeks to 14 weeks), pain levels, gender, and medication schedules that may have influenced findings and study participation. The status or level of participation in sport and recreational activities of each patient prior to TKA other than each had expressed a goal to return to such activities was unknown. Also, this study did not have a control or a comparison group that would have helped to delineate the influence of regular physical therapy exercises on the results. The loadpad device is likely not perfectly valid in measuring peak force but is reliable during the less dynamic tasks used in this study such as the squat and leg press. This preliminary investigation was solely to determine how force related feedback might be implemented in an outpatient sports medicine clinical setting. At present, there is no follow up data on each patient’s ability to return to sport and recreational activity.

CONCLUSION

Concurrent feedback training using load sensors may be used to improve the load symmetry during a single session of knee rehabilitation following TKA. Improvements were only shown in the short term and did not persist over the following week.

CONFLICT OF INTEREST

None declared.

FUNDING

None.

Submitted: February 09, 2023 CDT, Accepted: July 06, 2023 CDT

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-NC-4.0). View this license’s legal deed at https://creativecommons.org/licenses/by-nc/4.0 and legal code at https://creativecommons.org/licenses/by-nc/4.0/legalcode for more information.

International Journal of Sports Physical Therapy
REFERENCES


The Association of Joint Power Kinetic Variables with Running Injuries: A Case-Control Study
Matt Dewald, Jennifer Dalland, Josie Stockland

Keywords: injuries, joint power, kinetics, runners

Background
There is conflicting data on which kinetic variables are important to consider with running injuries. Furthermore, less is understood regarding differences in these variables when considering demographics such as age, sex, weight, and running speed. The primary question was what joint power kinetic variables were different between non-injured and injured runners.

Purpose
The purpose of this study was to identify if there were differences in joint power kinetic variables between non-injured runners and injured runners.

Study Design
Case-Control Study

Methods
Kinetic data were collected on 122 runners (26 non-injured and 96 injured) over three years with a Bertec force plated treadmill and Qualisys 3D motion capture. The subjects were considered eligible if they self-identified themselves as runners or had running as a key component of their activity. The subjects ran at a comfortable, self-selected pace while two 10-second trials of recordings were used to calculate the means of peak power generated at the hips, knees, and ankles of each gait cycle. Foot strike was categorized by kinematic data. Two sample T-tests were used to compare peak power variables at the hips, knees, and ankles between non-injured and injured runners. Logistic regression analyses examined how a combination of demographics and peak power variables were associated with injuries.

Results
No peak power variable at the hip, knee, or ankle was significantly different between injured and non-injured runners (p=0.07-0.87). However, higher hip power absorbed was found to be protective against injuries (odds ratio, .16; 95% CI .025-.88) when considering demographics using a logistic regression model including sex, foot strike, BMI, speed, age, and power variables from the hip, knee, and ankle. The area under the ROC curve was .74, which is acceptable discrimination.

Conclusion
When controlling for age, sex, BMI, foot strike, and speed; higher hip power absorbed was found to be protective against injury. This could be due to the hip muscles' unique role in absorbing force during early stance phase.
Level of Evidence

3b
©The Author(s)

INTRODUCTION

There is an obesity pandemic in the world, resulting in increased medical costs and decreased public health.1,2 There are multiple reasons for and solutions to this pandemic, with one solution being a healthy lifestyle that includes regular exercise.3,4 One of the most accessible exercise options to the public is running. It is well known that many runners suffer from running related injuries, with 19.4% to 79.3% injured every year.5 However, there are many myths and misunderstandings associated with running injuries. A better understanding of running injury risk is needed to minimize injuries, facilitate participation, and improve health conditions.

The causes of running injuries are multimodal and may include training errors, decreased sleep, poor nutrition, history of injuries, anatomical, physiological, and biomechanical factors. There is conflicting information about the importance and relevance of these variables as a whole.6-10 Nonetheless, a proportion of the risk is biomechanical, including running kinetics.6,11-15

While kinematic studies are much more prevalent in the literature, kinetic studies are beginning to fill in the gaps about running biomechanics. With improving technology, running kinetics are being investigated more often. However, presently many of these kinetic studies have investigated ground reaction forces and moments, not joint power.6,14 Joint power is a kinetic measure of the velocity of the joint moment, or the rate of work by the muscles at a particular joint.15

Currently, few studies have examined joint power kinetics and running, with limited studies investigating running injuries.15-17 Dicharry reported that running power kinetics are similar to walking at each joint, just with an increased amplitude.15 Xu et al. recognized that footfall strike pattern resulted in smaller knee power absorbed and higher ankle power absorbed compared to rearfoot strikers.16 Riley et al. compared overground running joint power to treadmill running and found a statistical difference in knee power generated and ankle power absorbed between the two modes of running.17 To the authors’ knowledge, there are no reports available that have assessed joint power of injured runners.

Therefore, the purpose of this study was to examine what joint power kinetic variables were different between non-injured and injured runners. The hypothesis was joint power kinetic variables would be different between non-injured and injured runners. This information will help identify what joint power kinetics are clinically relevant.

METHODS

STUDY DESIGN AND SUBJECTS

This was a case-control study of injured and non-injured runners. The subjects were considered eligible if they self-identified as runners or had running as a key component of their activity. The injured runners were current physical therapy patients, from provider referral or direct access, within the health system being treated for a running related injury. The injuries were determined by the referring provider or treating physical therapist. The non-injured runners declared no current running injury and completed analysis for performance goals. The sample size was based on existing data and supported by a sample size calculation using G*Power Version 3.1.9.7 based on .05 alpha, .8 power, and a medium-large effect size of .65 for comparing difference between two independent means with a ratio of four cases to one control resulting in 96 cases and 24 controls.

SETTING

All data were collected from the running lab at the health system’s primary sports performance and rehabilitation facility. Kinetic and kinematic data were collected from 122 runners (26 non-injured and 96 injured) over three years with a Bertec force plated treadmill and Qualisys 3D motion capture video analysis. The Qualisys Project Animation Framework running performance marker set, which includes 35 static and 35 dynamic markers, was the marker system used during the 3D motion analysis. Subjects completed an independent, self-selected warmup consisting of dynamic stretching or plyometric drills if desired, then were given three to five minutes on the treadmill to walk and run until they verbalized readiness to begin recordings. The data were de-identified and exported by an honest broker, an impartial biomechanical engineer not involved with any other portion of the study. Both the health system and the local university Human Subject Committees considered the study exempt, as this was secondary research with previously collected and de-identified data.

VARIABLES

The primary independent variables were joint power kinetics from the hip, knee, and ankle. These values were collected bilaterally and averaged. The peak power value, normalized to body weight, was recorded for both the power absorbed and the power generated at the joint when running.18 This resulted in six total kinetic variables to consider: hip power absorbed, hip power generated, knee power absorbed, knee power generated, ankle power absorbed, and ankle power generated.

Further variables included age, sex, height, BMI, foot strike, and speed. Height was self-reported and weight was collected on the force plate. Foot strike was categorized as
heel strike or non-heel strike based on the kinematic foot inclination angle with a positive value being heel strike and negative value being non-heel strike. To match what the subjects were doing functionally, they wore their own shoes and ran at a comfortable self-selected pace.

BIAS

To avoid selection bias between groups, subjects originated from the same general population. Controlling for demographics such as age, sex, weight, foot strike, and running speed in the adjusted logistic regression analyses further helped minimize selection bias. Performance bias was controlled with a standardized running assessment that was completed for both the injured and non-injured group preserving the fidelity to the protocol. The high validity and reliability of the measurement techniques used with the force plated treadmill and Qualisys 3D running analysis software ensures quality of the data.19-21

An honest broker was used to de-identify and export data resulting in a complete and thorough data set, this minimized information and attrition bias. There were no outliers or missing data in the de-identified data, however, one stride of the second trial for subject 58 ankle value was not calculated into the means due to abnormally low power. This was a reasonable way to manage that outlier as there was sufficient data on that subject to calculate an accurate mean for ankle power absorbed. Two 10-second trials of recordings were used to calculate the means of peak power generated at the hips, knees, and ankles of each gait cycle. In the end, all subjects had at least 15 steps (normally > 20) per trial.17

STATISTICAL METHODS

All statistical analysis was completed with SAS. All alphas were set at 0.05. Two sample T-tests were used to compare means of joint power kinetics between non-injured and injured runners. A log transformation was required with all variables in the T-tests except knee power absorbed.

A logistic regression was used to create a model with the binary dependent variable being injury status of the runner. The continuous independent variables considered included hip power absorbed, hip power generated, knee power absorbed, knee power generated, ankle power absorbed, ankle power generated, BMI, age, speed, height, and weight. Categorical variables considered included sex, foot strike, BMI categories, and age categories. The final model included age, sex, and BMI as categorical variables and hip power absorbed, hip power generated, knee power absorbed, knee power generated, ankle power absorbed, and speed as continuous variables.

Normality of the kinetic variables were assessed with Shapiro-Wilk, Kolmogorov-Smirnov, Histograms, and Q-Q Plots.22 Five of the six variables were right skewed, requiring a log transformation, as only knee power absorbed was normally distributed pre-transformation.

The predictive value of the logistic regression was assessed with the likelihood ratio of the global null hypothesis. Hosmer and Lemeshow Goodness-of-Fit test was referenced for fit statistics. Area under the ROC curve was given preference during model building, with R-Square adjusted used to compare models and R-Square values referenced for the final model to report total variation in injury status. Outlying and influential values were assessed by assessing the residual plots of Pearson, Deviance, Leverage, and DF-BETAS.22

Interactions were tested with custom model building. Multiple comparisons on the interaction term were compared between groups of interest to determine if there was an interaction. Multicollinearity was tested with the correlation coefficients, removing the highest values that resulted in significant model change.22

RESULTS

The average age of the subjects was 29.23 years-old (95% CI 27.03-31.45). The average height was 1.71 meters (95% CI 1.69-1.72). The average weight was 67.62 Kg (95% CI 65.57-69.88). The average BMI was 23.14 kg/m² (95% CI 22.56-23.73). All continuous descriptive statistics separated for injured and non-injured subjects can be found in Table 1.

Twenty-six subjects (21.31%) were classified as non-injured and 96 subjects (77.87%) as injured. Seventy-three subjects were female (59.84%) and 49 (40.16%) were male. The majority (111/122) of the subjects were classified as heel strikers (90.98%). When designating age categories, subjects under 40-years-old were classified as younger, with 27 (22.13%) subjects being older and 95 (77.87%) subjects being classified as younger. The 40-years-old threshold was used to match the categorization done by USA Track and Field symbolizing a "Masters" athlete.23 When using the CDC classifications of BMI, 85 (68.03%) were classified as healthy, four (3.28%) as obese, 26 (21.31%) as overweight, and nine (7.38%) as underweight.24 All injury locations are reported in Table 1, with some runners having more than one injury. Pertinent descriptive statistics of the injured and non-injured subjects can be found in Table 1.

When considering the normality of variables, five of the kinetic variables were right skewed and provided better insight with a log transformation, and only knee power absorbed was normally distributed. Following the transformation, there were no significant differences in the means of the peak kinetic data between the injured and non-injured runners when using a two-sample t-test (p=0.07-0.87). See Table 2 for the means (95% CI) and p-values of the two-sample t-test. Side by side box plots of the joint power variables can be seen in Figure 1.

Although no average peak power variable at the hip, knee, or ankle was significantly different between injured and non-injured runners (p=0.07-0.87); higher hip power absorbed was found to be protective against injuries (odds ratio, .16; 95% CI .025-.88) when considering demographics using a logistic regression model including categorical age, sex, BMI categories, speed, and power absorbed from the hip, knee, and ankle and power generated from the hip and knees. Ankle power generated was omitted secondary to multicollinearity. Analysis of the maximum likelihood esti-
mates of the final logistic regression found the only significant predictor of the variables included was hip power absorbed ($p = .041$).

There was a good area under the ROC curve at 74% (Figure 2) suggesting acceptable discrimination. However, the r-squared value was only 9%, suggesting the model is only responsible for 9% of the total variation in the injury status versus a model with no variables. The global null hypothesis did not find any predictive value in the model with a likelihood ratio of only .41. Lastly, throughout the model building, the Hosmer and Lemeshow Goodness-of-Fit test was significant with p-values above .05. However, with the last modification to the model, it’s p-value switched from indicating good fit to a p-value of 0.0499, resulting in a poor fit per the Hosmer and Lemeshow Goodness-of-Fit test.

Hip power absorbed was found to be protective against injury. Table 3 reports the Odds Ratios for injury risk. Figure 3 demonstrates the confidence limits with a forest plot, showing the hip power absorbed was completely under 1.00, indicating that higher hip power absorbed was associated with less injuries in runners. All other variables had confidence limits on either side of the 1.00 threshold, hence insignificant.

Pearson, Deviance, Leverage, and DFBETAS were used to assess residuals for outlying and influential values. Little change was seen in the model with potential outliers included or omitted, so all values were included.

Interactions, like outliers, did not significantly influence the model. Therefore, no effect modifiers or interactions were included. This was tested by including multiple comparisons on the interaction term. After checking the interaction term for significance, differences between the groups of interest were used to determine if the interaction was included.

Multicollinearity was tested by evaluating the correlation coefficients, starting with the highest value within the correlation matrix to lowest. In the end, ankle and knee joint power kinetics had values near 1 or -1. Each had a small effect on R-Square and ROC curve, but the best option was removing ankle power generated.

**DISCUSSION**

This is the first study that has reported an association between hip power absorbed and running injuries. When simply comparing the mean values of the injured group and the non-injured group of runners included in this study, there was no difference in the joint power. When both considering and controlling for age, sex, BMI, foot strike, and speed with logistic regression analysis; higher hip power absorbed was found to be associated with lesser odds of injury. This

**Table 1. Participant characteristics**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Injured (mean (95% CI))</th>
<th>Non-injured (mean (95% CI))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Participants, n (%)</td>
<td>96 (78.7)</td>
<td>26 (21.3)</td>
</tr>
<tr>
<td>Females, n (%)</td>
<td>57 (59.4)</td>
<td>16 (61.5)</td>
</tr>
<tr>
<td>Age (years), mean (95% CI)</td>
<td>28.1 (25.6-30.6)</td>
<td>33.4 (29.1-37.7)</td>
</tr>
<tr>
<td>Height (m), mean (95% CI)</td>
<td>1.71 (1.69-1.73)</td>
<td>1.70 (1.67-1.73)</td>
</tr>
<tr>
<td>Weight (kg), mean (95% CI)</td>
<td>67.5 (64.9-70.2)</td>
<td>67.9 (63.9-72.0)</td>
</tr>
<tr>
<td>BMI (kg/m$^2$), mean (95% CI)</td>
<td>23.3 (22.4-23.7)</td>
<td>23.5 (22.3-24.7)</td>
</tr>
<tr>
<td>Speed (m/sec), mean (95% CI)</td>
<td>3.1 (2.0-3.2)</td>
<td>3.1 (2.9-3.3)</td>
</tr>
<tr>
<td>Hip Power Absorbed</td>
<td>3.9 (3.6-4.1)</td>
<td>4.5 (3.9-5.0)</td>
</tr>
<tr>
<td>Knee Power Absorbed</td>
<td>5.0 (4.7-5.3)</td>
<td>4.8 (4.3-5.4)</td>
</tr>
<tr>
<td>Ankle Power Absorbed</td>
<td>11.8 (11.1-12.5)</td>
<td>12.2 (10.8-13.5)</td>
</tr>
<tr>
<td>Hip Power Generated</td>
<td>5.6 (5.2-5.9)</td>
<td>6.1 (5.5-6.8)</td>
</tr>
<tr>
<td>Knee Power Generated</td>
<td>7.4 (6.8-8.0)</td>
<td>7.1 (5.8-8.4)</td>
</tr>
<tr>
<td>Ankle Power Generated</td>
<td>10.6 (10.0-11.3)</td>
<td>10.4 (9.4-11.4)</td>
</tr>
<tr>
<td>Heel Strike, n (%)</td>
<td>88 (91.7)</td>
<td>23 (88.5)</td>
</tr>
<tr>
<td>Injury Location</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spine, Pelvis, and Abdomen, n %</td>
<td>7 (7.3)</td>
<td></td>
</tr>
<tr>
<td>Hip, n (%)</td>
<td>22 (22.9)</td>
<td></td>
</tr>
<tr>
<td>Knee, n (%)</td>
<td>44 (45.8)</td>
<td></td>
</tr>
<tr>
<td>Lower Leg, n (%)</td>
<td>45 (46.9)</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2. Two-sample T-tests of Joint Power Kinetic Variables**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Non-injured (95% CI)</th>
<th>Injured (95% CI)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log Hip Power Absorbed</td>
<td>1.43 (1.27-1.59)</td>
<td>1.28 (1.20-1.35)</td>
<td>0.07</td>
</tr>
<tr>
<td>Log Hip Power Generated</td>
<td>1.54 (1.41-1.66)</td>
<td>1.55 (1.48-1.62)</td>
<td>0.81</td>
</tr>
<tr>
<td>Knee Power Absorbed*</td>
<td>6.14 (5.48-6.80)</td>
<td>5.58 (5.24-5.93)</td>
<td>0.14</td>
</tr>
<tr>
<td>Log Knee Power Generated</td>
<td>1.78 (1.67-1.89)</td>
<td>1.68 (1.62-1.74)</td>
<td>0.12</td>
</tr>
<tr>
<td>Log Ankle Power Absorbed</td>
<td>1.88 (1.72-2.04)</td>
<td>1.94 (1.86-2.01)</td>
<td>0.47</td>
</tr>
<tr>
<td>Log Ankle Power Generated</td>
<td>2.31 (2.22-2.41)</td>
<td>2.32 (2.26-2.38)</td>
<td>0.87</td>
</tr>
</tbody>
</table>

*No log transformation was done on Knee Power Absorbed*
could be due to the hip muscles’ unique role in absorbing force during early stance phase and may warrant consideration in the context of running injuries.

Past kinetic studies have primarily been focused on moments and ground reaction forces. However, the concept of the hip being involved with and protective against running injuries is not new, with a significant amount of kinematic literature to support these findings. For example, a case-control study that stratified injuries to the patellofemoral joint, iliotibial band, medial tibial, and Achilles tendon found contralateral pelvic drop as the kinematic variable best at predicting injury, regardless of injury type. Further highlighting the importance of the hips, increased hip adduction is associated with patellofemoral pain, iliotibial band syndrome, and tibial stress fracture. Increased hip internal rotation is associated with iliotibial

Figure 1. Side by Side Box Plots
improved moments at the hip, along with improved pain, hip, and core strength. This could suggest that three weeks was not enough time to change hip biomechanics. Additionally, targeted gait training for eight sessions to control hip adduction and contralateral pelvic drop was found to improve those variables, along with pain again in subjects with patellofemoral pain. This all supports the conclusion that Willy and Davis had when suggesting that without specifically targeting the movement patterns during running, little improvements in biomechanics would be seen. Of clinical importance, successful gait retraining to improve hip biomechanics may be as simple as mirror biofeedback over eight sessions without the need for isolated strengthening interventions.

Overall, the authors’ expected to find more kinetic variables associated with running injuries. There may have been other causes for the non-associations and possible type-II errors that were not accounted for with the adjusted logistic regression. However, the R-Square value seems reasonable when considering the many variables to injuries, such as sleep, training errors, physiology, anatomy, society pressure, and psychology. Other considerations were that stratification could not be completed by injury or demographic categories because of low subgroup sample sizes, leading to large confidence intervals in the end. There appeared to be a few outliers, but they did not change the model when removed, so they were all included. Interactions did make some changes to the AUC when included, however they were not kept because the significance was not in comparison groups of interest. Subjects did wear their own shoes and ran at a comfortable self-selected pace, which matches what is done functionally; however, it does add possible confounding to the equation. The model controlled for speed, but not for shoe type. While many of the demographic variables were very similar between the injured and non-injured runners, the non-injured group’s mean age was five years older. The age was controlled for in the logistic regression, however this may have introduced bias. Lastly, the population of patients and clients may not be generalizable to runners of different demographics. The

Table 3. Odds Ratio for Injury Risk for Injured and Non-Injured Runners

<table>
<thead>
<tr>
<th>Effect</th>
<th>Odds Ratio</th>
<th>95% Confidence Intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female vs Male</td>
<td>.869</td>
<td>.281 - 2.559</td>
</tr>
<tr>
<td>Older vs Younger</td>
<td>.688</td>
<td>.226 - 2.160</td>
</tr>
<tr>
<td>Heavy vs Underweight BMI</td>
<td>.721</td>
<td>.035 - 5.130</td>
</tr>
<tr>
<td>Obese vs Underweight BMI</td>
<td>.358</td>
<td>.010 - 12.486</td>
</tr>
<tr>
<td>Overweight vs Underweight BMI</td>
<td>.263</td>
<td>.012 - 2.111</td>
</tr>
<tr>
<td>Speed</td>
<td>.640</td>
<td>.180 - 2.242</td>
</tr>
<tr>
<td>Hip Power Absorbed</td>
<td>.161*</td>
<td>.025 - .880</td>
</tr>
<tr>
<td>Ankle Power Absorbed</td>
<td>1.850</td>
<td>.384 - 9.118</td>
</tr>
<tr>
<td>Hip Power Generated</td>
<td>2.050</td>
<td>.304 - 14.697</td>
</tr>
<tr>
<td>Knee Power Generated</td>
<td>.431</td>
<td>.028 - 6.523</td>
</tr>
<tr>
<td>Knee Power Absorbed</td>
<td>.938</td>
<td>.755 - 1.164</td>
</tr>
</tbody>
</table>

*Hip Power Absorbed provided a statistically significant protective effect
subjects included in this study had the resources available to seek out and receive these services.

Joint power kinetics need to be considered in the big picture of running injuries. Recall that running injuries are very multimodal, with only 9% of the total variation in injury status being accounted for in this model. The results of this study by no means suggest that improving runners’ ability to absorb hip power will decrease their injury risk, it simply shows there may be an association. Prospective studies would be needed to confirm that. Future studies could look at strategies to 1) improve hip absorption power in injured runners and 2) assess whether there is a decrease in injury risk in doing so. For clinicians without access to kinetic measurements, it would also be interesting to assess how the hip kinematics were associated with the hip kinetics and injuries.

CONCLUSION

This study confirmed much of the previous understandings of running injuries as being multimodal, with a proportion of the risk associated with biomechanics. This study further identifies hip power absorbed as being associated with injuries, warranting further research such as how to change hip power absorbed in runners and if those changes result in less injuries with a prospective study design.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

Submitted: December 30, 2022 CDT, Accepted: May 16, 2023 CDT

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-NC-4.0). View this license's legal deed at https://creativecommons.org/licenses/by-nc/4.0 and legal code at https://creativecommons.org/licenses/by-nc/4.0/legalcode for more information.
REFERENCES


Original Research

No Effect of Return to Sport Test Batteries with and without Psychological PROs on the Risk of a Second ACL Injury: A Critical Assessment of Four Different Test Batteries

Ramana Piussi1,2,3,a, Rebecca Simonson1,2, Johan Högberg1,2, Roland Thomeé1,2, Kristian Samuelsson3,4,5, Eric Hamrin Senorski1,2,3

1 Sportrehab Sports Medicine Clinic, Stampgatan 14, SE-411 01 Gothenburg, Sweden, 2 Unit of Physiotherapy, Department of Health and Rehabilitation, Institute of Neuroscience and Physiology, Sahlgrenska Academy, University of Gothenburg, Box 455, SE-405 30 Gothenburg, Sweden, 3 Sahlgrenska Sport Medicine Center, Gothenburg, Sweden, 4 Department of Orthopaedics, Institute of Clinical Sciences, The Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden, 5 Department of Orthopaedics, Sahlgrenska University Hospital, Mölndal, Sweden, 6 Swedish Olympic Committee, Stockholm, Sweden

Keywords: ACL re-rupture, subsequent injury, psychological, patient-reported outcome

https://doi.org/10.26603/001c.81064

International Journal of Sports Physical Therapy

Background

Patients report psychological barriers as important when returning to sport, however, psychological outcome measures are seldom included in return to sport (RTS) assessment. There is a need for clinical trials to integrate psychological patient-reported outcomes (PROs) in return to sport batteries assessing patients treated with ACL reconstruction.

Objective

The aim of this study was to determine the association between passing clinical tests of muscle function and psychological PROs and sustaining a second ACL injury in patients who RTS after primary ACL reconstruction.

Design

Retrospective Cohort study

Methods

Patients’ sex, age, height and weight, and the results of strength and hop tests, as well as answers to PRO’s (including Tegner activity scale, the ACL Return to Sport after Injury scale (ACL-RSI) as well as the Quality of Life (QoL) subscale of the Knee injury and Osteoarthritis Outcome Score (KOOS)), were extracted from a rehabilitation-specific registry. Four different test batteries comprising muscle function tests and PROs were created to assess whether patients were ready to RTS. Passing each of the test batteries (yes/no) was used as an independent variable. A multivariable Cox proportional hazard model analysis was performed, with sustaining a second ACL injury (either ipsi- or contralateral; yes/no) within two years of RTS as the dependent variable.

Results

A total of 419 patients (male, n=214; 51%) were included, of which 51 (12.2%) suffered a second ACL injury within the first two years after RTS. There were no differences in passing rates in the different RTS test batteries comprising muscle function tests and PROs for patients who suffered a second ACL injury compared to patients who did not.

a Corresponding author:
Ramana Piussi
Address:
Sportrehab Sports Medicine Clinic, Stampgatan 14, SE-411 01 Gothenburg, Sweden
Ramana.piussi@gu.se +46761677520
Conclusion

No association between passing the RTS clinical tests batteries comprising muscle function and psychological PROs used, and the risk of a second ACL injury could be found.

Level of Evidence

3

©The Author(s)

INTRODUCTION

Anterior cruciate ligament (ACL) injuries occur most commonly in the active young population, with individuals between 14 and 25 years having the highest injury risk. In patients who undergo surgical treatment after an ACL injury, up to 22% will go on to sustain a second ACL injury when attempting to return to sport (RTS). Returning to sport after an ACL reconstruction entails an increased exposure to the risk of sustaining a second ACL injury. Therefore, authors have aimed to assess whether patients can be defined as "ready" to RTS; i.e. achieving physical and psychological status that would be associated with reduced, or minimal, risk of sustaining a second ACL injury. Patients have reported that psychological impairments such as fear of re-injury are greater and more difficult to overcome than physical impairments such as pain or muscular weakness, both during rehabilitation and when returning to sport after ACL reconstruction. Despite clinical practice guidelines (CPG's) highlighting the importance of assessing psychological factors during rehabilitation and RTS time, strength, and hop tests are the most commonly used criteria to determine readiness to RTS. Psychological factors are important to integrate in RTS testing. Furthermore, systematic reviews assessing association between passing physical performance based RTS tests and second ACL injury risk show no associations. Psychological factors are important for patients, and clinical trials should aim to integrate psychological patient-reported outcomes (PROs) in RTS batteries and assess whether passing RTS assessments including both physical and psychological outcome measures is able to reduce the risk of a second ACL injury.

The aim of this study was to determine the association between passing clinical tests comprising both muscle function as well as psychological PROs, and second ACL injury in patients who RTS after primary ACL reconstruction.

METHODS

To enhance the quality and transparency of this study, the REporting of studies Conducted using Observational Routinely-collected Data (RECORD)12 statement, extended from the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE)13 statement checklist for cohort studies, was followed. The present study was based on data extracted from a rehabilitation-specific outcome registry, Project ACL. The registry was established in 2014, with the aim of improving the care of patients after ACL injuries, regardless of treatment, surgical and rehabilitation or rehabilitation alone. The registry aims to improve care through utilization of standardized and continuous evaluation of rehabilitation specific outcomes in patients with ACL injury. The data in Project ACL consist of the results of unilateral isokinetic concentric strength tests, three unilateral hop tests (muscle function tests; MF) and responses to PROs (i.a. the ACL Return to Sport after Injury scale (ACL-RSI) and the Knee Injury and Osteoarthritis Outcome Score (KOOS) from patients with ACL injuries. The data are collected prospectively, according to a standardized schedule of follow-ups, starting from ACL injury or ACL reconstruction (baseline), followed by assessments at 10 weeks, 4, 8, 12, 18 and 24 months, and then every five years. The tests are administered and supervised by physical therapists specifically trained and regularly updated in the test protocol. The test procedures have previously been described in detail.14,15 Prior to participation in Project ACL, patients receive written information and informed consent is obtained.16 Ethical approval was obtained from the Swedish Ethical Review Authority (registration number 2020-02501).

STUDY EXECUTION

Patients registered in Project ACL were eligible. Following inclusion criteria were applied: age between 16-50 years; pre-injury activity level of > 6 on the Tegner Activity Scale (Tegner); suffered one ACL injury; treated with ACL reconstruction and participated in all the tests in the test battery. If no test data were available within six months of reaching Tegner > 6, or if patients had not reached Tegner > 6, i.e. returned to knee-strenuous sport, they were excluded. Patients who did not return to Tegner > 6 are patients who might 1) clinically struggle to return to sport, which might be due to physical limitations, 2) change of interest and/or 3) change of life goals; 4) concomitant injuries; 5) a combination of the above.

In this specific study, returning to Tegner > 6 was defined as RTS, since after Tegner level 6, only sport activities are possible choices, thereby excluding all patients who have knee demanding activities as an occupation. Patients who reported a return to knee-strenuous sport (Tegner > 6) before the four-month follow-up after ACL reconstruction were excluded, since a return to knee-strenuous sport within four months of ACL reconstruction was deemed unlikely as the phase of graft healing starts around the fourth month, and strength deficits are high between three and six months after ACL reconstruction.19,20 Once included, data available from the follow-up closest in time to when
patients reported at least Tegner 6 (RTS) were extracted for further analysis. A survival time of two years from reporting Tegner > 6 after ACL reconstruction was used. Patients were followed for two years or to second ACL injury, whichever occurred first.

Data from MF tests, PROs and demographics for included patients were extracted from Project ACL in April 2021.

MUSCLE FUNCTION TESTS

Before performing the MF tests, patients warmed up according to a standardized procedure consisting of 10 minutes on a stationary bike and sub-maximal trials on each test. The MF tests consisted of five tests: two isokinetic strength tests (knee extensors and knee flexors), and three hop tests. All tests were performed one limb at the time, starting with the reconstructed limb.

Torque (as a measure of strength) of the knee extensors and knee flexors was tested using an isokinetic dynamometer (Biodex System 4; Biodex Medical System, Shirley, NY, USA) at an angular speed of 90°/second. Strength testing with Biodex has been reported to be reliable (ICC = 0.95) when measuring muscle strength. Patients performed the strength test as previously described, sitting on the dynamometer chair, starting with knee extension. After the warm-up, patients performed three maximal trials, with 40 seconds’ rest between each trial. The rest period of 40 seconds was allowed to ensure patients were able to push as hard as possible, since the peak torque was the objective of the test. The highest values in Newton meters (Nm) were recorded.

The three unilateral hop tests were performed in the following order: the vertical hop, the hop for distance, and the 30-second side-hop test. For the vertical hop and the hop for distance, two trial repetitions were allowed before maximal testing. All the hop tests were performed with the patients holding their hands behind their back. For the vertical hop, the time from take-off to landing was converted into hop height in centimeters (cm) (Muscle lab, Ergostest Technology, Oslo, Norway). Three maximal trials were performed and the highest hop in cm was recorded. In the hop for distance, the distance in cm from toe at take-off to heel at landing was measured. Patients were required to perform a stable landing, without losing their balance or letting go of their hands. Three maximal trials were performed and the longest hop in cm was recorded. For the 30 seconds side-hop test, patients were required to perform as many hops as possible over two lines 40 cm apart. One 30-second trial was allowed, and the total number of hops (not touching the lines) was recorded. The hop tests were performed as described by Gustavsson et al. and were chosen as hop tests with the highest ability to discriminate hop performance in patients who had sustained an ACL injury and in patients who had undergone an ACL reconstruction.

The hop tests have been reported to have a high level of sensitivity and accuracy in patients with an ACL injury (87% and 84%, respectively) and in patients who had undergone ACL reconstruction (91% and 88%, respectively).25

PATIENT-REPORTED OUTCOMES

The PROs used for analysis in this study were the ACL-RSI as well as only the Quality of Life (KOOS QoL) subscale of the KOOS, i.e. no other subscales than the QoL were used from the KOOS. The ACL-RSI was chosen, as it has been reported as a PRO with high methodological quality for evaluating patients with an ACL injury.24 The ACL-RSI measures the patients’ emotion, confidence and risk appraisal of RTS. The ACL-RSI has 12 items. Each item is graded from 0 to 10, where 10 is the highest response, representing the strongest positive emotion, confidence and low risk appraisal regarding RTS. The results for each item are summarized in a total score, normalized on a 0–100 scale. The KOOS QoL was chosen, as knee-related quality of life is an important part of recovery and can be impaired up to 20 years after an ACL reconstruction. The KOOS has five sub-scales: pain, symptoms, function of daily living, function in sports and recreation and QoL, and in this study only the QoL subscale was used. Each item is rated from 0 to 4 on a 5-point Likert scale. A normalized score from 0 to 100 is calculated for each subscale, where 0 indicates the most severe symptoms and 100 indicates no symptoms.

Four different test batteries to assess whether patients are ready to RTS were created for analysis. The same test batteries were previously introduced in a publication from our study group (Table 1).27

For the two MF test, the vertical hop and the hop for distance were chosen as Abrams et al. reported that these tests were the most commonly used after ACL reconstruction. The five MF tests were chosen according to current consensus criteria.29 Psychological PROs were added to both the two MF and the five MF tests, creating two additional different test batteries.

To interpret the results of all of the MF tests, a cut-off value of $\geq$ 90% limb symmetry index (LSI) was chosen to define whether or not a patient passed a test. The Limb Symmetry Index is the ratio between the injured and uninjured limb expressed as a percentage. For the ACL-RSI, a score of 76.6 points was used as a cut-off or passing score, since this score has been reported to have maximal sensitivity for discriminating between patients who suffer a second ACL injury and patients who do not.22 For the KOOS QoL, a score of 62.5 points was chosen as a passing score, since Mueller et al. suggested that this score represents a threshold of "feeling well" after primary ACL reconstruction.

STATISTICS

Patients demographics were presented stratified by patient sex, but no statistical comparison between sexes was performed. To pass (yes/no), each of the test batteries was used as an independent variable for analysis. A multivariable Cox proportional hazard model analysis was performed, with suffering a second ACL injury (yes/no) as the dependent variable. Sub-analyses were performed depending on whether the second ACL injury was ipsilateral or contralateral. A hazard ratio with a value greater than 1 indicates a variable that is positively associated with a second ACL in-
jury, while a hazard ratio lower than 1 indicates a variable negatively associated with a second ACL injury.54 The results of the Cox proportional hazard model were reported with hazard ratios (HR), 95% confidence intervals (CI) and p-values. Statistical analyses were performed with the Statistical Analysis System (SAS) software version 9.4 (Copyright © 2013, SAS Institute Inc., Cary, NC, USA). Mean values with standard deviations (SD) or medians (min-max) were presented for demographic data for the entire cohort and stratified by sex. A significance level of 0.05 was set.

Several factors have been previously associated with second ACL injury risk: they include young age (15-19 years), female sex, high BMI (> 25), time to return to knee-demanding physical activity (> 9 months) and several surgically related factors, i.e. graft choice and fixation.16,35,36 To account for variables associated with a second ACL injury risk, demographic variables were compared with the outcome of interest, i.e. suffering a second ACL injury (yes/no), using Fisher’s exact test. Consequently, model analysis was adjusted for variables significantly associated to the risk for a second ACL injury, that is: time to RTS (months) and patient sex.

RESULTS

A total of 419 patients were included in the study (men, n=214; 51%). A flowchart of the inclusion process is presented in Figure 1 and demographics for the included patients are summarized in Table 2.

A total of 51 (12.2%) patients suffered a second ACL injury within the first two years after RTS, of which 51 (61% of second ACL injury) occurred ipsilaterally and 20 (39% of second ACL injury) occurred contralaterally (Table 2).

The number of second ACL injuries per month following primary reconstruction and the side of the second ACL injury are presented in Figure 2. The number of second ACL injuries per month following RTS (Tegner ≥ 6) and the side of the second ACL injury are presented in Figure 3.

The proportion of patients that passed the different test batteries, stratified by whether or not patients suffered a second ACL injury, is presented in Figure 4. There were no differences in passing rates for the respective test batteries between patients who suffered a second ACL injury and patients who did not. As such, the test batteries were not able to identify patients who were at increased risk of sustaining a second ACL injury. Results for the univariable and adjusted Cox proportional hazard model are presented in Table 3.

When stratifying patients depending on whether the second ACL injury occurred ipsilaterally or contralaterally, passing any of the test batteries was not able to identify patients who were at an increased risk of sustaining a second ACL injury (Table 4 and 5).

The Cox hazard ratios associated with a contralateral second ACL injury was only performed in a univariate manner (not adjusted) since no confounding variable showed a significant difference between groups and was therefore not deemed to be a confounding variable.

DISCUSSION

The main finding in this study was that passing various clinical test batteries consisting of muscle function and psychological PROs had no association with the risk of sustaining a second ACL injury in patients who had undergone primary ACL reconstruction.

The addition of psychological outcomes to the RTS test batteries in the present study was not able to identify patients running an increased risk of suffering a second ACL injury. However, the authors’ still assert that the inclusion of psychological outcome measures in the assessment of individual patients prior to RTS is meaningful, as patients express that psychological factors are of great importance,5 and second ACL injury risk is likely a complex puzzle incorporating physical and psychological dimensions.37,38 Results of interview studies suggest that some patients may not feel ready to resume sports participation and are afraid to RTS due to the risk of sustaining a new knee injury.39-41 RTS assessments seldom include the assessment of fear and readiness and, as a result, they may not measure factors that predispose patients to second knee injuries.42,43 Although psychological outcomes in terms of emotions, con-

Table 1. Various combinations of outcome measures grouped into test batteries for analysis.

<table>
<thead>
<tr>
<th>Type of test</th>
<th>Strength tests</th>
<th>Hop tests</th>
<th>PROs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two MF tests</td>
<td></td>
<td>Vertical hop</td>
<td>KOOS QoL</td>
</tr>
<tr>
<td>Two MF tests and two PROs</td>
<td></td>
<td>Hop for distance</td>
<td>ACL-RSI</td>
</tr>
<tr>
<td>Five MF tests</td>
<td>Knee extensors</td>
<td>Vertical hop</td>
<td>KOOS QoL</td>
</tr>
<tr>
<td></td>
<td>Knee flexors</td>
<td>Hop for distance</td>
<td>ACL-RSI</td>
</tr>
<tr>
<td>Five MF tests and two PROs</td>
<td>Knee extensors</td>
<td>Vertical hop</td>
<td>KOOS QoL</td>
</tr>
<tr>
<td></td>
<td>Knee flexors</td>
<td>Hop for distance</td>
<td>ACL-RSI</td>
</tr>
</tbody>
</table>

ACL-RSI = Anterior Cruciate Ligament Return to Sport; KOOS QoL = Knee injury and Osteoarthritis Outcome Score; Quality of Life subscale; MF = muscle function; PROs = patient-reported outcomes.
Figure 1. Flowchart with inclusion/exclusion criteria.

ACL = anterior cruciate ligament, n = number of patients, Tegner = Tegner activity scale; RTS = return to sport

The results of the clinical tests used in the current study were not associated with a second ACL injury, which confirms results summarized in two systematic reviews that investigated associations between passing RTS test batteries and occurrence of second knee injuries after an ACL reconstruction.\textsuperscript{10,11} The current results indicate that four different test batteries were unable to identify patients at increased risk of sustaining second knee injuries after ACL reconstruction.\textsuperscript{10,11} Future research should examine the test batteries and assess each component of the battery (e.g., hop test or PRO) to identify patients running the risk of new ACL injuries after ACL reconstruction. With regard to passing MF tests, a cut-off of an LSI of $\geq$ 90% was used to define "pass", based on consensus criteria published in 2015.\textsuperscript{30} Notably, the consensus criteria stating that 90% LSI is "successful outcome" after ACL reconstruction, does not clearly define what "successful" infers. Arguably, an LSI of $\geq$ 90% has been used as a proxy for muscular recovery,\textsuperscript{47-49} and consequently, as a logical step towards assessing patients for RTS. However, the sensitivity or specificity of the cut-off in relation to the risk of a second ACL injury has perhaps not been accounted for. Taken together, further studies determining cut-off values for all included tests of MF and their relationship to a second ACL injury after a primary ACL reconstruction are warranted, and clinicians working with patients who have undergone ACL reconstruction should look beyond results from test batteries and consider second ACL injury risk as a multifactorial con-
Table 2. Demographic data. Mean values, standard deviations (SD), count (n) and proportions (%).

<table>
<thead>
<tr>
<th></th>
<th>All subjects; n=419</th>
<th>Men; n=214</th>
<th>Women; n=205</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at reconstruction (years)</td>
<td>23.9 (7.5)</td>
<td>25.0 (7.3)</td>
<td>22.8 (7.5)</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>175.0 (9.8)</td>
<td>182.2 (6.2)</td>
<td>167.5 (6.7)</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>72.4 (12.3)</td>
<td>80.6 (9.7)</td>
<td>63.8 (8.4)</td>
</tr>
<tr>
<td>BMI</td>
<td>23.5 (2.5)</td>
<td>24.3 (2.5)</td>
<td>22.7 (2.2)</td>
</tr>
<tr>
<td>Time to return to knee-demanding physical activity (months)</td>
<td>11.0 (4.2)</td>
<td>10.2 (3.8)</td>
<td>11.8 (4.4)</td>
</tr>
<tr>
<td>Tegner pre-injury (level) n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>24 (5.7%)</td>
<td>13 (6.1%)</td>
<td>11 (5.4%)</td>
</tr>
<tr>
<td>7</td>
<td>64 (15.3%)</td>
<td>39 (18.2%)</td>
<td>25 (12.2%)</td>
</tr>
<tr>
<td>8</td>
<td>114 (27.2%)</td>
<td>41 (19.2%)</td>
<td>73 (35.6%)</td>
</tr>
<tr>
<td>9</td>
<td>141 (33.7%)</td>
<td>80 (37.4%)</td>
<td>61 (29.8%)</td>
</tr>
<tr>
<td>10</td>
<td>76 (18.1%)</td>
<td>41 (19.2%)</td>
<td>35 (17.1%)</td>
</tr>
<tr>
<td>Returned to pre-injury Tegner Yes n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n (%)</td>
<td>147 (35.1%)</td>
<td>75 (35.0%)</td>
<td>72 (35.1%)</td>
</tr>
<tr>
<td>Graft choice</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hamstring graft, n (%)</td>
<td>331 (85.1%)</td>
<td>167 (83.9%)</td>
<td>164 (86.2%)</td>
</tr>
<tr>
<td>Patellar graft, n (%)</td>
<td>53 (13.7%)</td>
<td>29 (14.6%)</td>
<td>24 (12.7%)</td>
</tr>
<tr>
<td>Other graft, n (%)</td>
<td>5 (1.3%)</td>
<td>3 (1.5%)</td>
<td>2 (1.1%)</td>
</tr>
<tr>
<td>Missing (graft type) n (%)</td>
<td>30</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Second ACL injury within 2 years of RTS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes n (%)</td>
<td>51 (12.2%)</td>
<td>23 (10.7%)</td>
<td>28 (13.7%)</td>
</tr>
<tr>
<td>Ipsilateral ACL injury n (%)</td>
<td>31 (61.0%)</td>
<td>15 (7.0%)</td>
<td>16 (7.8%)</td>
</tr>
<tr>
<td>Contralateral ACL injury n (%)</td>
<td>20 (39.2%)</td>
<td>8 (3.7%)</td>
<td>12 (5.9%)</td>
</tr>
</tbody>
</table>

ACL = anterior cruciate ligament; BMI = body mass index; cm = centimeters; kg = kilograms; n = number; RTS = return to sport; Tegner = Tegner activity scale, defined as Tegner ≥ 6

Figure 2. Incidence of second ACL injury starting from ACL reconstruction as baseline.

Green bar: ipsilateral second ACL injury; blue bar: contralateral ACL injury. ACL = anterior cruciate ligament; n = number

International Journal of Sports Physical Therapy
No Effect of Return to Sport Test Batteries with and without Psychological PROs on the Risk of a Second ACL Injury: A...

Figure 3. Incidence of second ACL injury after RTS (Tegner ≥6) as baseline.

Green bar: ipsilateral second ACL injury; blue bar: contralateral ACL injury. ACL = anterior cruciate ligament; n = number

LIMITATIONS

There are some limitations to this study. The exact rehabilitation that was administered and the extent of rehabilitation compliance are not known. Patients compliant with...
Table 3. Univariable and adjusted for time to return to sport (months) and patient sex Cox hazard ratios associated with a second ACL injury.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Event rate (fail/pass)</th>
<th>HR of a second ACL injury</th>
<th>p value</th>
<th>Adjusted HR of a second ACL injury</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 MF tests</td>
<td>28 of 236 / 23 of 183</td>
<td>0.93 (0.54 - 1.61)</td>
<td>0.79</td>
<td>0.82 (0.47 - 1.44)</td>
<td>0.50</td>
</tr>
<tr>
<td>2 MF tests and 2 PROs</td>
<td>33 of 258 / 5 of 34</td>
<td>0.83 (0.33 - 2.14)</td>
<td>0.71</td>
<td>0.76 (0.29 - 1.94)</td>
<td>0.56</td>
</tr>
<tr>
<td>5 MF tests</td>
<td>35 of 297 / 14 of 96</td>
<td>0.78 (0.42 - 1.45)</td>
<td>0.43</td>
<td>0.71 (0.38 - 1.32)</td>
<td>0.27</td>
</tr>
<tr>
<td>5 MF tests and 2 PROs</td>
<td>34 of 252 / 3 of 22</td>
<td>0.97 (0.30 - 3.17)</td>
<td>0.96</td>
<td>0.94 (0.29 - 3.10)</td>
<td>0.93</td>
</tr>
</tbody>
</table>

ACL = anterior cruciate ligament; HR = hazard ratio; BMI = body mass index; MF = muscle function; PROs = patient-reported outcomes. Adjusted for time to return to sport (Tegner ≥ 6; months) and patient sex

Table 4. Cox hazard ratios associated with an ipsilateral second ACL injury. Hazard ratio both univariable, as well as adjusted for time to return to sport (months) and patient sex.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Event rate (fail/pass)</th>
<th>HR of a second ACL injury</th>
<th>p value</th>
<th>Adjusted HR of a second ACL injury</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 MF tests</td>
<td>18 of 236 / 13 of 183</td>
<td>1.06 (0.52 - 2.16)</td>
<td>0.88</td>
<td>0.90 (0.44 - 1.84)</td>
<td>0.77</td>
</tr>
<tr>
<td>2 MF tests and 2 PROs</td>
<td>21 of 258 / 4 of 34</td>
<td>0.67 (0.23 - 1.94)</td>
<td>0.46</td>
<td>0.57 (0.20 - 1.67)</td>
<td>0.31</td>
</tr>
<tr>
<td>5 MF tests</td>
<td>20 of 297 / 10 of 96</td>
<td>0.63 (0.29 - 1.34)</td>
<td>0.23</td>
<td>0.54 (0.25 - 1.15)</td>
<td>0.11</td>
</tr>
<tr>
<td>5 MF tests and 2 PROs</td>
<td>22 of 252 / 3 of 22</td>
<td>0.63 (0.19 - 2.11)</td>
<td>0.46</td>
<td>0.53 (0.16 - 1.77)</td>
<td>0.30</td>
</tr>
</tbody>
</table>

ACL = anterior cruciate ligament; HR = hazard ratio; BMI = body mass index; MF = muscle function; PROs = patient-reported outcomes. Adjusted for time to return to sport (Tegner ≥ 6; months) and patient sex

Table 5. Univariable Cox hazard ratios associated with a contralateral second ACL injury.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Event rate (fail/pass)</th>
<th>HR of a second ACL injury</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 MF tests</td>
<td>10 of 236 / 10 of 183</td>
<td>0.76 (0.32 - 1.83)</td>
<td>0.55</td>
</tr>
<tr>
<td>2 MF tests and 2 PROs</td>
<td>12 of 258 / 1 of 34</td>
<td>1.51 (0.20 - 11.60)</td>
<td>0.69</td>
</tr>
<tr>
<td>5 MF tests</td>
<td>15 of 297 / 4 of 96</td>
<td>1.16 (0.39 - 3.50)</td>
<td>0.79</td>
</tr>
<tr>
<td>5 MF tests and 2 PROs</td>
<td>12 of 252 / 0 of 22</td>
<td>3725267 (0.00 - )</td>
<td>0.99</td>
</tr>
</tbody>
</table>

ACL = anterior cruciate ligament; HR = hazard ratio; BMI = body mass index; MF = muscle function; PROs = patient-reported outcomes. Adjusted for time to return to sport (Tegner ≥ 6; months) and patient sex

rehabilitation that consists of high-intensity training might be more prone to attempt RTS earlier and possibly have higher scores on PROs as well. Another limitation is that concomitant injuries such as cartilage, multi-ligamentous, or meniscal injuries, were not accounted for, which could potentially further increase the risk of a second ACL injury. Several statistical analyses were made in this study, raising the risk of type 1 error. However, the possibility of type 1 errors was accounted for by providing CIs for all hazard ratios as a measure of uncertainty, attempting to move away from the assumption that significant p-values mean "true" or clinically significant results. Another limitation is that data were analyzed for the available follow-up closest in time to when patients reported having returned to a Tegner ≥ 6. Project ACL's follow-ups are 10 weeks and 4, 8, 12, 18 and 24 months after baseline. The results for a patient who has returned to Tegner level ≥ 6 at 10 months after ACL reconstruction might therefore have been taken from the 8-month follow-up. However, muscle strength, knee function and psychological readiness can develop and change over a period of two months and the actual patient status at the time of RTS might therefore have been different compared with the time of RTS testing. This limitation was partially accounted for by excluding patients who did not participate in a follow-up within six months of reaching Tegner ≥ 6. The use of Tegner as a measure of functional status is a limitation in this study, as the Tegner assesses the level of knee-strenuous activity, but it does not account for the time of exposure or the intensity of exposure for a patient, which are factors that contribute to the risk of a second ACL injury. A further limitation is the exclusion of patients who did not return to Tegner ≥ 6 in the analysis. Patients who do not return to Tegner ≥ 6 might have presented with physical problems which obstructed their possibility of performing physical tasks and would have prevented their RTS. Therefore, it is important for clinicians to consider that results from the current study might not be applicable to patients who do not return to Tegner ≥ 6. A further limitation concerns the exclusion of patients who reported RTS before the four month follow-up. The use of LSI to assess muscle strength is a limitation itself as it does not account for movement quality, nor for eventual strength losses in the uninvolved leg, compared with pre-operative values. Additionally, a further limi-
tion might concern treating RTS as a dichotomous variable. The RTS involves a criteria-based progression starting from return to participation to RTS and, ultimately, return to performance, and can be described as a continuum that starts the day of injury.57 Taken together, these limitations affect the general external validity of results in the present study. However, it is important that clinicians not rely solely on the results of test batteries to assess the risk of second ACL injury in the process of sharing information (regarding test results and second ACL injury risk) with patients who are about to resume RTS after ACL reconstruction. Future prospective studies with large cohorts and different test batteries, with different cut-offs are warranted in order to develop greater knowledge of passing clinical test batteries including physical and psychological outcomes and their ability to reduce the risk of a second ACL injury.

CONCLUSION

No associations between passing any of four clinical tests batteries that included assessments of muscle function and psychological PROs and the risk of a second ACL injury were found in patients who had undergone primary ACL reconstruction.

GRANT SUPPORT

Local Research and Development Board for Gothenburg and Södra Bohuslän: grant number 937140. The funding source had no involvement in the study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

CONFLICT OF INTEREST

Author KS reports he is a Member of the Board of Directors of Getinge AB (publ).

ACKNOWLEDGEMENT

The authors thank biostatisticians Bengt Bengtsson and Nils-Gunnar Pehrsson from Statistiska Konsultgruppen for help with the statistical analyses and advice on the interpretation of data.

Submitted: November 10, 2022 CDT, Accepted: May 06, 2023 CDT
REFERENCES


Two-Dimensional and Three-Dimensional Biomechanical Factors During 90° Change of Direction are Associated to Non-Contact ACL injury in Female Soccer Players

Stefano Di Paolo¹, Alberto Grassi², Filippo Tosarelli³, Matteo Crepaldi³, Laura Bragonzoni³, Stefano Zaffagnini³, Francesco Della Villa³

¹ Life Quality Studies, University of Bologna, ² 2nd Orthopaedic and Traumatologic Clinic, Istituto Ortopedico Rizzoli, ³ Education and Research Department, Isokinetic Medical Group

Keywords: ACL, female football, biomechanics, 2D video-analysis, injury prevention, change of direction

Original Research

Background
The two-dimensional (2D) video-analysis of the change of direction (COD) technique has never been used to attempt to predict the risk of ACL injury in female football players.

Hypothesis/Purpose
The purpose of the present pilot study was to prospectively investigate the biomechanical predictors of ACL injury during a COD task in female football players using both gold standard 3D motion capture and a qualitative scoring system based on 2D video-analysis.

Study Design
Prospective cohort study

Methods
Sixteen competitive female football (soccer) players (age 21.4 ± 4.3) performed a series of pre-planned 90° COD tasks. 3D motion data was recorded through 10 stereophotogrammetric cameras and a force platform. 2D frontal and transverse plane joint kinematics were computed through video-analysis from three high-speed cameras. A scoring system based on five criteria was adopted: limb stability, pelvis stability, trunk stability, shock absorption, and movement strategy. The players were prospectively followed for the next two consecutive football seasons and the occurrence of severe knee injuries was registered.

Results
Four players (25%) experienced an ACL injury. In 3D analysis, ACL-injured players showed greater knee valgus, knee internal rotation, and lower knee flexion (p= 0.017 – 0.029). Lower hip flexion coupled with greater external rotation (p= 0.003 – 0.042), ankle eversion, and contralateral pelvic drop (p=0.001) were also noted. In 2D analysis, ACL-injured players showed greater internal foot rotation, contralateral pelvic drop, lower knee flexion, and contralateral trunk tilt (moderate-to-large effect size). Pelvis stability and trunk stability showed the highest predictive value towards ACL injury. Total score was significantly lower in ACL-injured players with a moderate effect size (d=0.45).
Conclusions

Both 3D and 2D methodologies depicted biomechanical risk factors and offered predictive insights towards the ACL injury risk. Awareness should rise in women’s football regarding the high risk of ACL injury and the strategies to assess and mitigate it.

Level of Evidence

3

©The Author(s)

INTRODUCTION

The rapid growth of women's football (soccer) participation is bringing about increasing rates and earlier occurrence of primary Anterior Cruciate Ligament (ACL) injury. Several authors have demonstrated that female footballers are at overall higher risk of ACL injury than their male counterparts and are less responsive to targeted preventative interventions e.g., the FIFA 11+ program. Therefore, a deeper comprehension of modifiable risk factors, such as ACL injury biomechanics, has been advocated to mitigate the injury risk in female football academies. Standardized methodologies are crucial to obtain objective measurements of players’ movement quality and identify ACL injury predictors in a user- and athlete-friendly fashion.

The 3D motion capture is the gold standard for biomechanical assessment of high-demanding tasks and has been used in prospective investigations of ACL injury predictors. Recently, 2D video-analysis has been proposed as a cost-effective alternative to 3D motion capture for high-demand tasks. Promising results have reported regarding the association between 2D angles (frontal plane knee, pelvis, and trunk angle; limited knee and hip flexion) and scoring systems (Movement Analysis Test, Cutting Movement Assessment Score) based on video-analysis and well-known ACL risk factors such as the knee abdication moment (KAM) and high vertical and posterior ground reaction forces (GRF).

Such 2D scoring systems allow for simpler movement analysis for the raters and quicker feedback to the athletes, making them an attractive solution to implement large screenings in young population. However, the efficacy of 2D video-analysis in determining the risk of the occurrence of ACL injury has never been assessed prospectively in male or female football players. The efficacy of 2D video-analysis and its agreement with 3D motion capture would support the use of dedicated methodologies to test female football players in primary ACL injury prevention.

The purpose of the present pilot study was to prospectively investigate the biomechanical predictors of ACL injury during a COD task in female football players using both gold standard 3D motion capture and a qualitative scoring system based on 2D video-analysis. The hypothesis was that both the 3D motion capture and the 2D scoring system would detect clinically-relevant biomechanical differences between non-injured and ACL-injured players such as shallower lower limb flexion and dynamic valgus.

MATERIALS AND METHODS

The study was approved by the Institutional Review Board (IRB approval ID number: 555/2018/Sper/IOR of 12/09/2018) of Area Vasta Emilia Romagna Centro (AVEC, Bologna, Italy) and registered on ClinicalTrials.gov (Identifier: NCT05840551). Signed informed consent were collected from all participants before enrollment.

PARTICIPANTS

This represents a secondary analysis of a larger cross-sectional study aimed to investigate the concurrent validity of different methodologies (wearable inertial sensors and 2D video-analysis) for examining high-demand motion tasks for ACL injury prevention and rehabilitation. The analysis was conducted in the Education and Research Department of the Isokinetic Medical Group during the regular football season (February 2019). Sixteen young competitive (first-to-third Italian Football division) female football (soccer) players (age 21.4 ± 4.3, range 18-31) were enrolled in the study. Exclusion criteria were: evidence of previous severe injury (> 28 days recovery time); previous surgery to the lower limbs; body mass index (BMI) > 35; and cardiopulmonary or cardiovascular disorders.

DATA COLLECTION

Each player performed a series of pre-planned 90° COD tasks at maximum speed possible in a specialized laboratory equipped with artificial turf. The players wore their own running shoes and clothes. The COD task consisted of a 6-meter frontal sprint followed by a 90° sidestep cut with foot strike on a force platform and a further 3-meter sprint. The complete acquisition setting was previously described. In brief, a 10-minute dynamic warm-up and few unrecorded repetitions of the movement were performed by each player to get confident with the setting. The players were instructed by a single sport & exercise medicine physician specialized in sports biomechanics (F.D.V.), who also verified the validity of each trial. Full foot contact on the force platform was required to consider a trial valid. Each player performed six valid trials with three right and three left strikes.

3D motion data and 2D video-analysis data were collected simultaneously. 3D Motion data were recorded through a set of 10 stereophotogrammetric cameras (VIC-CON Nexus, Vicon Motion Systems Ltd, Oxford, UK) and a force platform embedded in the floor (AMTI 400*600, Watertown, MA USA), and three high-speed cameras placed...
frontally and bilaterally towards movement direction to capture 2D data. The systems were synchronized for direct data comparison. The sampling rate was 120 Hz for cameras and force platform, and 100 Hz for the high-speed cameras.

The system calibration was performed at the beginning of the acquisition and repeated periodically during the session. Forty-two retroreflective markers were placed on each participant by a single expert user according to the full-body Plug-in Gait protocol. Anthropometric data were collected for each participant and model calibration was performed before data acquisition.

DATA PROCESSING – 3D MOTION CAPTURE

Marker trajectories were collected through the stereophotogrammetric cameras and interpolated through a spline-based algorithm. The lower limb and pelvis kinematics were computed through VICON Nexus software using the Cardan angles. The frontal, transverse and sagittal plane kinematics waveforms were extracted for the stance phase of the foot on the force platform. The positively defined rotations (+) were: knee varus and internal rotation, hip adduction and internal rotation, ankle inversion and internal rotation, and pelvis contralateral drop and rotation.

DATA PROCESSING – 2D VIDEO ANALYSIS

The 2D frontal and transverse plane joint kinematics was computed through video-analysis. The evaluation was performed in a specific VICON software environment through the recordings of the three high-speed cameras. Joint kinematics was evaluated at the frame of maximal knee flexion angle after the foot contact on the force platform. For each trial, the following 2D angles were calculated: Foot Projection Angle (FPA), Frontal Plane Knee Projection Angle (FPKA), Pelvis tilt Angle (PA), and Trunk tilt Angle (TA). Furthermore, a scoring system was adopted based on the 2D kinematics. Five scoring criteria were used to evaluate each COD trial: 
1. limb stability (LS),
2. pelvis stability (PS),
3. trunk stability (TS),
4. shock absorption (SA), and
5. movement strategy (MS).

The detailed description of each criteria was reported in a previous study. In brief, a score from 0/2 to 2/2 (non-adequate to adequate) was attributed to the trial for each criterion based on objective 2D measurements by a single sports physician. The total score was computed as the sum of each criterion sub-score (maximum total score: 10/10).

PROSPECTIVE FOLLOW-UP

The players were prospectively followed for the next two consecutive football seasons through regular phone calls. The occurrence of severe knee injuries (> 28 days recovery time) was registered. In case of an ACL injury, an MRI and a clinical examination were carried out to confirm the presence of the ACL tear. Additional information, e.g., the nature of the injury mechanism (non-contact, indirect contact, contact) and the context (training, match), were retrieved.

STATISTICAL ANALYSIS

Based on the occurrence of an ACL injury during the follow-up period, the players were divided into two groups: “No injury” and “ACL injury”. Kinematic data were retrieved for the injured and dominant leg for the latter and former group, respectively. The normal distribution of the data was verified through the Shapiro-Wilk test. The categorical variables were presented as a percentage over the total, while the continuous variables were presented as mean ± standard deviation (95% Confidence Interval – CI).

For the 3D kinematics waveforms, the two-tailed Student’s t-test with Statistical Parametric Mapping (SPM) was used to investigate the difference between the two groups.

For the 2D analysis, the two-tailed Student’s t-test was used to compare continuous 2D kinematics variables, and Cohen’s d effect size was reported. Effect size was considered small, medium, and large for Cohen’s d value of 0.2, 0.5, and 0.8, respectively. Furthermore, the Chi-square test for 3x2 (3 scores x 2 groups) contingency tables was computed to investigate the differences in the 2D scoring system percentage between the two groups for each of the five criteria. The Mann-Whitney U test was used to investigate the differences in the total score and in non-normally distributed variables, and Rank Biserial Correlation was reported as effect size measure. Differences were considered statistically significant for p<0.05. The statistical analyses were conducted in Matlab (The MathWorks, Natick, MA, US) and SPSS (IBM, Armonk, NY, US).

Given the pilot nature of the present study, no a-priori power-analysis was performed. However, the present cohort of 16 participants is consistent with previous pilot studies with similar purpose and methodological accuracy showing a minimum power of 0.80 for the analysis of frontal and sagittal plane knee and hip kinematics.

RESULTS

Overall, four players out of 16 (25% of the cohort) experienced an ACL injury during the two football seasons following the data collection. All injuries occurred during football matches with a non-contact mechanism. A complete ACL rupture in all injured athletes was confirmed through MRI examination. No other severe joint/muscle injuries were reported. Age at the time of data collection was significantly lower (p=0.009) in the ACL-injured group, while no differences (p>0.05) were found in demographics and approaching COD speed between ACL-injured and non-injured players.

3D MOTION CAPTURE

The ACL-injured players showed greater knee valgus (p=0.029) and internal rotation (p=0.017) during the early stance phase and lower knee flexion (p=0.023) during the late stance phase (Figure 1, first row). Less hip flexion (p=0.042) coupled with greater external rotation (p=0.003) were also noted in the ACL-injured group (Figure 1, second
row). Furthermore, ACL-injured players showed greater ankle evasion tendency and contralateral pelvic drop (p<0.001, Figure 1, third and fourth row).

2D VIDEO ANALYSIS

ACL-injured players showed greater internal foot rotation, pelvic contralateral drop, lower knee flexion (moderate effect size), and contralateral trunk tilt (large effect size, Table 2, Figure 2). Pelvis stability and trunk stability scores showed the highest predictive value towards ACL injury (significantly more 0/2 scores in ACL-injured players, Table 3, Figure 3). Total score was significantly lower in ACL-injured players (moderate effect size), that never scored greater than 7/10 (Table 5, Figure 4).

DISCUSSION

The main finding of the present study was that both the 3D motion capture and 2D analysis of the COD technique offered predictive insights towards the ACL injury risk in female football players. Significant differences between players experiencing an ACL injury in the two seasons after testing and non-injured players were consistent with the current literature on non-contact ACL injury biomechanics.21–24 Moreover, the 2D scoring system successfully identified the ACL-injured players, that showed a lower total score (Figure 4). Thus, such a scoring system has the potential to identify female footballers at high risk for ACL injury in a cost-effective fashion.

A video-analysis study by Lucarno et al. recently described the ACL situational patterns and injury mechanism of professional female footballers, highlighting the frequent occurrence of non-contact or indirect contact injuries with multiplanar biomechanics.23 Injury biomechanics included knee dynamic valgus, limited lower limb flexion, pelvis and trunk tilt and rotation, and ankle rotation, consistently with the one of male counterparts.25 In the present study, all the differences found between ACL-injured and non-injured players agreed with the typical biomechanical profile for non-contact ACL injury described by Lucarno et al. and Della Villa et al.23,25

The ACL-injured players exhibited greater knee valgus and internal rotation and hip external rotation (Figure 1). These patterns have been shown to increase the KAM and the load on knee ligaments.24,26,27 In LS score, none of the ACL-injured players’ tasks was rated as 2/2. However, no differences in FPKPA were found and the overall rating was shifted towards the 0/2 score (Figure 2). These findings suggest a good sensitivity but a lower specificity of the LS score and the FPKPA for the ACL injury risk, in line with previous literature on other functional tests.28,29

Lower hip and knee flexion angles were also observed in the ACL-injured players in both 3D and 2D joint angles. Though not happening in the early stance phase, these results suggest a stiffer kinematic strategy adopted by the ACL-injured players.23,30,31 Pelvic drop was detected during the whole stance phase in the players that sustained an ACL injury (Figure 1). Pelvic drop and trunk lean in the ACL-injured players were also detected through the PA and TA 2D angles with a moderate-to-large effect. The PS and TS criteria strongly were the best predictors of ACL injury in the 2D scoring system: the percentage of trials rated 0/2 in the ACL-injured players was double than the non-injured players in the PS, and triple in the TS. Moreover, ACL-injured players were never rated 2/2 in PS. The association between poor core stability and the risk of knee overloading has been proposed in literature.12,13 The video-analysis studies on professional athletes clearly identified that ACL injuries happen with pelvis and trunk in non-neutral position.23,25 Donelon et al. recently listed the maintenance of an upright pelvis and trunk position among the actions aimed to modify the COD technique and mitigate the ACL injury risk.22 The results of the present study suggest that the analysis of pelvis and trunk frontal plane kinematics in 2D could effectively detect core imbalance during the COD task and that the scoring system described could have a strong predictive power towards the risk of ACL injury.

Greater ankle eversion was noted in the players that sustained an ACL injury in the second half of the stance phase (Figure 1). The 2D analysis revealed a more internal foot progression angle at peak knee flexion. Both aspects of the foot position indicate that ACL-injured players landed on the force platform adopting a more anticipatory strategy.
than the non-injured players. An anticipatory strategy with a greater internal foot progression angle has been associated with faster cutting but in conflict with a safe mechanism. Foot eversion and rotation are also common in the professional football players that sustain an ACL injury. The analysis of foot progression angle could give valuable insights on the attitude of players COD technique and strongly influence joint loading propagation on the kinetic chain. The 2D total score proposed in the present study distinguished the players that sustained an ACL injury from the non-injured ones. Average total score was lower in the ACL-injured players (4.6 vs 2.4, moderate effect, Table 3) and none of the ACL-injured players’ trials were rated more than 7/10 (Figure 4). The total score is an intuitive measure that can be used by the professionals to track a player’s overall movement quality during the COD task and her progression over a prevention (or rehabilitation) program. From the results of the present pilot study, a total score of 4/10 or lower could be indicative or higher risk for ACL injury, while a score higher than 7/10 could be indicative of a protective COD technique.
Table 2. Joint kinematics (°) extracted from 2D video-analysis according to the occurrence of ACL injury during the following two football seasons.

<table>
<thead>
<tr>
<th>2D angle</th>
<th>No ACL injury (n=12)</th>
<th>ACL injury (n=4)</th>
<th>p-value</th>
<th>Cohen's d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foot Progression Angle</td>
<td>-15.2 ± 19.4 [-26.2 - -4.2]</td>
<td>-25.2 ± 19.9 [-44.7 - -5.7]</td>
<td>0.041</td>
<td>0.51</td>
</tr>
<tr>
<td>Frontal Plane Knee Projection Angle</td>
<td>36.8 ± 17.2 [27.1 - 46.5]</td>
<td>41.8 ± 8.1 [33.9 - 49.7]</td>
<td>n.s.</td>
<td>0.23</td>
</tr>
<tr>
<td>Pelvic Drop Angle</td>
<td>-9.9 ± 7.4 [-14.3 - -5.5]</td>
<td>-12.9 ± 9.1 [-21.8 - -3.9]</td>
<td>0.022</td>
<td>0.38</td>
</tr>
<tr>
<td>Trunk Tilt Angle</td>
<td>4.2 ± 7.4 [0.1 - 8.4]</td>
<td>-5.7 ± 11.7 [-17.2 - 5.8]</td>
<td>&lt;0.001</td>
<td>1.11</td>
</tr>
<tr>
<td>α (180°-Knee Flexion Angle)</td>
<td>118.7 ± 10.7 [112.6 - 124.8]</td>
<td>124.3 ± 10.5 [114.0 - 134.6]</td>
<td>0.026</td>
<td>0.53</td>
</tr>
</tbody>
</table>

Notes: Data were expressed as mean ± standard deviation (95% confidence intervals). n.s. = non-statistically significant difference (p>0.05). Positive rotations: foot external rotation, knee valgus, pelvic contralateral hike, trunk ipsilateral tilt.

Figure 2. Joint kinematics from 2D video-analysis. Data are presented as mean and standard deviation for the non-injured (blue) and ACL-injured (red) players.

Notes: Asterisks represent statistically significant differences between the two groups (p<0.05). FPA (Foot Progression Angle), FPKPA (Frontal Plane Knee Projection Angle), PA (Pelvic Drop Angle), TA (Trunk Tilt Angle), α (180°-Knee Flexion Angle). Positive rotations: foot external rotation, knee valgus, pelvic contralateral hike, trunk ipsilateral tilt.

The multiplanar risk factors detected in the players that sustained an ACL after testing were extremely congruent with the injury mechanism proposed in video-analysis studies on professional female footballers (Figure 5).25 Both the 3D motion capture and 2D video-analysis detected and agreed on most of the clinically-relevant differences. In the 2D video-analysis, frontal plane was indicative of a higher injury risk, but both frontal and sagittal view would be required to obtain a comprehensive measure of COD movement quality. Knee abduction angle, frequently used in multiple 2D assessments for this and other tasks (e.g., drop vertical jump, squats, and single-leg landings), might be not sufficient to explain the movement complexity and the relation with the ACL injury risk.28,29,33 Furthermore, approaching and landing strategies such as penultimate foot contact deceleration, anticipatory cut, flatfoot or heel strike, should be taken into account to finetune the assessment.11,12,22,34

In the present study, 25% of the cohort experienced an ACL injury. Previous larger prospective studies reported an injury rate of 7-11%.29,35,36 In the present study, the tests were performed in late February, thus close to one of the two peaks of injury occurrence identified in professional football players.25 All injuries occurred during matches and without a direct contact with an opponent. This is again in line with the predominance of non-contact/indirect contact mechanisms found in professional female footballers (88%.25). The fact that ACL-injured players were younger (p=0.009) than the rest of the cohort is consistent with the data reporting earlier occurrence (younger age) of the first ACL injury in female football.37 All these aspects deserve awareness from the football medicine community and highlight the urgent need for dedicated tools to counteract such a high injury risk since more and more female players turn professional earlier.

The present study has some limitations. First, the present pilot investigation involved a small cohort; thus, inferences should be interpreted with caution and are far from

Table 3. Frequencies (%) of single sub-scores for the qualitative scoring system from 2D video-analysis according to the occurrence of ACL injury during the following two football seasons.

<table>
<thead>
<tr>
<th>Score</th>
<th>No ACL injury (n=12)</th>
<th>ACL injury (n=4)</th>
<th>p-value</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limb Stability</td>
<td>87.0 7.4 5.6 95.8 4.2 0.0</td>
<td>n.s. 1.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pelvis Stability</td>
<td>48.1 27.8 24.1 87.5 12.5 0.0</td>
<td>0.003 11.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trunk Stability</td>
<td>18.5 46.3 35.2 58.3 20.8 20.8</td>
<td>0.002 12.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shock Absorption</td>
<td>16.7 59.3 24.1 25.0 66.7 8.3</td>
<td>n.s. 2.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motor Strategy</td>
<td>16.7 29.6 53.7 29.2 29.2 41.7</td>
<td>n.s. 1.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Score</td>
<td>4.6 ± 2.8 [3.0 - 6.2] 2.4 ± 1.8 [0.6 - 4.2]</td>
<td>0.001 0.45</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Total score data were as mean ± standard deviation (95% confidence intervals). n.s. = non-statistically significant difference (p>0.05).
being conclusive. Despite the sample size, the results of both the 2D and 3D analyses were consistent with current literature on ACL injury risk biomechanics.21-24 Prospective studies with larger cohorts are necessary to set a benchmark for objective measurements in ACL injury prevention in football. Second, the 90° COD task assessed in the present study was pre-planned. Recent studies suggest that unanticipated sidestepping tasks could offer important insights on ACL injury risk.38 Preventative testing should consider the use of both pre-planned and unplanned tasks to get a wider understanding of players’ injury risk. No muscular strength data were collected. Such data could have increased the level of detail of the players’ comparison. The data collection was performed in a single session and no post-operative investigation was performed on those who were injured. The longitudinal analysis of the ACL-injured players could be used to quantitatively inform the ACL rehabilitation programs.

The clinical relevance of the present work is that the adoption of a 2D video-analysis scoring system based on quantitative measurements has the potential to identify female football players at higher risk for ACL injury in a cost-effective fashion. The players could therefore benefit from additional preventative neuromuscular training. Given the rapid growth of female football and their greater predisposition to the ACL injury,4,59 dedicated methodologies assessing players’ movement quality and injury risk are fundamental to mitigate the injury rates in early football careers.

CONCLUSION

Significant kinematic differences during a planned COD task were identified between ACL injured and non-injured players with both the 3D and 2D methodologies. The biomechanical pattern of the ACL-injured players was strongly consistent with the video-analysis of ACL injured elite players.25 Both the 3D and 2D methodologies offered precious predictive insights to assess and mitigate the ACL injury risk in women’s football.
Figure 5. ACL injury frame (left) of one of the players involved in the prospective cohort. The frontal (middle, up) and sagittal (right, up) view of the 90° change of direction test performed with the involved limb is reported.

Notes: The associated 2D total score was 2/10 (bottom).

LIST OF ABBREVIATIONS

Anterior cruciate ligament (ACL)
Body Mass Index (BMI)
Body Weight (BW)
Change of Direction (COD)
Confidence intervals (CI)
Foot Projection Angle (FPA)
Frontal Plane Knee Projection Angle (FPKPA)
Ground Reaction Force (GRF)
Knee Abduction Moment (KAM)
Limb Stability (LS)
Movement strategy (MS)
Pelvis tilt angle (PA)
Pelvis stability (PS)
Shock absorption (SA)
Statistical Parametric Mapping (SPM)
Trunk tilt angle (TA)
Trunk stability (TS)

CONFLICT OF INTEREST

Each author certifies that he or she has no commercial associations (e.g., consultancies, stock ownership, equity interest, patent/licensing arrangements, etc) that might pose a conflict of interest in connection with the submitted article.

FUNDING

None.

ACKNOWLEDGMENTS

The authors want to thank GPEM s.r.l. for the support in data collection.

Submitted: April 20, 2023 CDT, Accepted: July 04, 2023 CDT
REFERENCES


Balance Error Scoring System Performance Differences in Figure Skaters Based on Discipline

L. Colby Mangum, Andrew Skibski, Luk Devorski, Lindsay Slater

1 Kinesiology and Rehabilitation Sciences, University of Central Florida, 2 Physical Therapy, University of Illinois at Chicago, 3 Athlete High Performance, United States Figure Skating

Keywords: sports performance, balance, screening, youth sports

https://doi.org/10.26603/001c.81598

International Journal of Sports Physical Therapy

Background
Balance and postural stability are required of figure skaters throughout on-ice performance. Spinning, jumping, and landing each rely on this skill set to maintain control while skaters manage changing demands for each skating discipline.

Hypothesis/Purpose
The aim of this study was to compare balance error scoring system (BESS) performance in figure skaters between disciplines and determine if age was related to BESS performance.

Study Design
Cross-sectional study.

Methods
Three hundred and fifty-eight figure skaters (age: 15.4±3.3 years, 213 females, 145 males) of multiple disciplines completed the BESS during the United States Figure Skating's Standardized Testing of Athleticism to Recognize Skaters (S.T.A.R.S.) combine. Errors during each condition of the BESS were recorded by trained evaluators. A 3x6 ANOVA was used to determine BESS differences based on skating discipline. Spearman’s rho (ρ) correlation coefficients were calculated for relationships between BESS errors and age.

Results
Ice dancers had more errors than singles and pairs for bipedal foam (p<0.001) but had fewer errors than single skaters for single leg foam (p=0.002). Tandem on a firm surface also showed an increase in errors for ice dancers and pairs skaters compared to singles (p<0.001). There were significant weak negative relationships noted between age and bipedal foam and single leg firm conditions (ρ=-0.14, -0.23, p<0.05).

Conclusion
Figure skaters of different disciplines have varying levels of static postural stability. Assessing postural stability in figure skaters can provide insight to improve performance and may identify skaters at risk of injury.

Level of Evidence
3

©The Author(s)
INTRODUCTION

Postural stability and balance are important to figure skaters of all ages as they influence flexibility and control. As figure skaters develop their on-ice skills from an early age, training may focus on specific jumps or spins, however the balance required to improve upon those initial skills can be measured independently from skating performance. Jumping and landing on a single leg requires power, strength, balance, and stability. Investigators have identified greater hamstring strength and hamstrings to quadriceps strength ratios in female figure skaters compared to female soccer players. Trunk flexion and extension forces while standing and were higher in an athlete sample, including figure skaters, that was evaluated against a non-athlete group. In junior elite figure skaters, knee extension strength was positively correlated to single axel and double axel height. Strength investigations in the figure skating population are more prevalent, revealing a gap in the balance and stability literature for this group of athletes. Clinicians can easily implement functional balance assessments to measure these characteristics.

The Balance Error Scoring System (BESS) is a clinical assessment originally developed for measuring postural control characteristics in healthy athletes without expensive equipment. Although the BESS is often utilized in concussion testing, this tool can be used to establish normative values for static postural stability that can be compared to values after musculoskeletal injury. The BESS involves testing balance in three different stance conditions: bipedal, single leg, and tandem. As an assessment tool for postural imbalances, both firm and foam surfaces are commonly used. An athletic setting may increase the amount of external attentional demands experienced by the test taker which may influence their BESS score. Ideally the test should be administered in a controlled environment.

Tests of postural stability such as the BESS can provide insight into athletic capabilities, but also can be influenced by factors such as age, ankle instability, or head injury. Although the BESS may not have a relationship with the occurrence of acute lower extremity injuries, the test is advantageous to understand how balance differs between levels of sport. In figure skaters, dynamic postural stability has been assessed with the Y-Balance test to compare between sexes, but static postural stability has only been tested in a limited manner with the stork pose, which tests only single leg balance. Static postural stability is influenced during figure skating as positions are held isometrically for extended periods of time, whether in bipedal, single leg or tandem positions. In adolescents and young adults, older age improves BESS performance, likely due to cognitive and physical development. Therefore, it is important to use broader assessments such as the BESS to more fully understand variations in static postural stability in figure skaters across disciplines and ages.

The purpose of this study was to compare balance error scoring system (BESS) performance in figure skaters between disciplines and determine if age was related to BESS performance. A secondary purpose was to compare BESS errors between males and females, regardless of discipline. It was expected that skaters would display fewer errors during the bipedal and firm surface conditions across all disciplines. It was hypothesized that singles skaters would perform better on the BESS. For the secondary aim, females were expected to exhibit fewer errors than males. Age was expected to have a negative relationship indicating that as skater age increased, their errors decreased in each BESS condition.

METHODS

STUDY DESIGN

A cross-sectional study evaluated the balance of participants in the United States Figure Skating’s Standardized Testing of Athletism to Recognize Skaters (S.T.A.R.S.) combine. Deidentification of all data was performed before analysis and therefore this study was deemed exempt by the University Institutional Review Board.

PARTICIPANTS

The 358 participants did not have any current injury that limited their ability to complete the BESS test during the S.T.A.R.S. combine. The BESS was administered as part of combine testing by United States Figure Skating trained evaluators, which has sufficient interrater reliability for stance condition scores.

TESTING PROCEDURES

Prior to balance testing, individuals self-reported age, sex, landing leg, and skating discipline. Landing leg length was measured from the greater trochanter to the lateral malleolus in centimeters using a tape measure and recorded. All skaters completed balance testing in the same order of conditions, starting with a firm then foam surface. The participants were barefoot when completing conditions as follows: bipedal firm, bipedal foam, single leg firm, single leg foam, tandem firm, and finally tandem foam. The single leg tasks were completed on their non-dominant or non-landing leg, and this leg was also placed behind the landing leg in the tandem stance conditions. Skaters were asked to balance with their eyes closed for 20-seconds for each trial. Errors were counted real-time by evaluators according to standardized BESS scoring with a maximum of 10 errors for each condition.

STATISTICAL ANALYSES

Participant characteristics (age, sex, leg length, skating discipline) were used to calculate descriptive statistics. Errors observed were averaged for each BESS condition grouped by skating discipline and sex. A 3x6 ANOVA with Bonferroni post-hoc tests was performed to examine differences between discipline and condition. Mean differences (MD), standard error (SE), and 95% confidence intervals (CI) were calculated for significant effects identified from ANOVA. Independent t-tests were used to evaluate sex differences. For non-normally distributed data, assessed by Shapiro-
Table 1. Errors on Balance Error Scoring System by Skating Discipline

<table>
<thead>
<tr>
<th>BESS Condition</th>
<th>Singles</th>
<th>Pairs</th>
<th>Ice Dance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bipedal Firm</td>
<td>0.01±0.07</td>
<td>0.00±0.00</td>
<td>0.02±0.18</td>
</tr>
<tr>
<td>Bipedal Foam**.***</td>
<td>0.35±1.00</td>
<td>0.00±0.00</td>
<td>1.24±2.89</td>
</tr>
<tr>
<td>Single Leg Firm</td>
<td>2.26±2.48</td>
<td>1.75±1.68</td>
<td>1.89±2.11</td>
</tr>
<tr>
<td>Single Leg Foam**</td>
<td>5.13±3.28</td>
<td>5.20±1.73</td>
<td>3.80±3.82</td>
</tr>
<tr>
<td>Tandem Firm*</td>
<td>1.68±2.23</td>
<td>3.53±2.24</td>
<td>3.81±3.68</td>
</tr>
<tr>
<td>Tandem Foam</td>
<td>3.66±2.44</td>
<td>3.13±2.19</td>
<td>3.84±2.80</td>
</tr>
</tbody>
</table>

*p<0.05, difference between singles and pairs
**p<0.05, difference between singles and ice dancers
***p<0.05, difference between pairs and ice dancers
BESS, balance error scoring system.

Wilk tests, Spearman’s rho correlation coefficients were utilized to calculate relationships between age and BESS performance in each condition. Alpha was set a priori to p<.05. All SPSS v.28.0.1.0 (SPSS, Chicago, IL) was used for all statistical analyses.

RESULTS

There were 358 participants included in this study including singles, pairs, and dancers (age: 15.4±3.3 years, leg length: 84.7±7.3 cm, 213 females, 145 males). Mean and standard deviation for errors completed in each BESS condition by discipline are summarized in Table 1. The ANOVA revealed between group differences based on discipline for the bipedal foam (F: 16.42, p<0.001), single leg foam (F: 6.56, p=0.002), and tandem firm (F: 23.90, p<0.001). There were no significant interactions between discipline and BESS condition. In the bipedal foam stance, ice dancers had more errors than both pairs (MD: 1.24, SE: 0.28, 95% CI: 0.57, 1.91) and singles skaters (MD: 0.89, SE: 0.18, 95% CI: 0.47, 1.32). Singles skaters had more errors than ice dancers in the single leg foam stance (MD: -1.34, SE: 0.39, 95% CI: -2.26, -0.41). In the tandem firm position, ice dancers had more errors than singles skaters (MD: 2.13, SE: 0.32, 95% CI: 1.35, 2.91), and pairs had more errors than singles skaters (MD: 1.85, SE: 0.49, 95% CI: 0.67, 3.03). There were no differences between males and females in any BESS conditions (Table 2), although males were significantly older (Males: 16.7±3.5 years, Females: 14.5±2.9years, p<0.001) and had a longer leg length (Males: 88.8±7.4 cm, Females: 81.9±5.8 cm, p<0.001). Age had a weak, negative relationship with bipedal foam (p=-.14, p=.001) and with single leg firm (p=-.25, p<0.001).

DISCUSSION

The results of this study support that BESS performance varies based on skating discipline. There were differences between disciplines for the bipedal foam, single leg foam, and tandem firm stances. The BESS effectively identifies static postural instability, especially in stances performed on foam surfaces.6 Singles skaters perform higher difficulty jumps and more technical elements during their routines, so they require a greater degree of postural stability.17 Ice dancers do not perform jumps, so they may not require the same level of postural stability needed for take-off and landing by singles skaters, as demonstrated by the poorer BESS performance in the bipedal foam and tandem firm stances (Table 1).

Past research has found varying performance on other measures of athleticism across skating discipline, specifically testing agility, strength, and flexibility.1 For example, the Y-Balance test is a dynamic test that has been used to measure the posterolateral reach in figure skaters, which corresponds with the landing position.15 A shorter reach distance during the posterolateral reach within the Y-Balance test has been correlated with skating speed. The Star Excursion Balance Test, from which the Y-Balance test was derived, has been incorporated into a previous study evaluating a proprioceptive training protocol in 10–18-year-old figure skaters.18 All participants in their sample showed improvement in the Star Excursion Balance Test following the strength and proprioception training.18 It has been reported that females have a greater Y-Balance composite score than males,15 but skating discipline has not previously been assessed. Comparing Y-Balance and BESS performance can provide insight into both dynamic and static stability.

Noting differences in postural stability allows for a discipline-specific understanding of the sport’s requirements. It has also been reported that dancers have decreased lumbar and hamstring flexibility compared to singles and pairs skaters, which may impact stability.1 The same decrease in postural stability and balance manifested in our present study with a greater number of errors in the bipedal foam and tandem firm stances (Table 1). This discipline-specific approach to balance can impact training and performance, injury prevention and rehabilitation strategies, and establishment of reference values for the BESS.

There were no discipline-specific differences in BESS performance for the bipedal firm, single leg firm, or tandem foam stances. These findings suggest that the easier, initial position, the bipedal firm stance, may not provide an adequate challenge to postural stability. Single leg firm can be a more challenging position, which is supported by the current findings when noting that single leg firm had more errors than the other firm stances in this sample, regardless of lack of difference between disciplines. However, single
Leg balance on a firm surface has been assessed in a figure skating population using the stork pose, which found greater hold times in ice dancers than singles and pairs skaters. Therefore, measuring stability with an endurance component may provide an additional challenge to compare between these athletes. Very few errors were recorded across groups for the bipedal firm and single leg firm positions (Table 1). The small number of errors agrees with previous findings which reported few BESS errors and minimal sway measured with a force platform in these positions. Conversely, the tandem foam stance may have been similarly challenging due to sway for all disciplines, as that condition averaged approximately three errors for all disciplines. Three errors during tandem foam is comparable to other studies investigating the BESS in healthy collegiate baseball players (2.9 errors), and a mixed sample of youth (3.8 errors), high school (3.5 errors), and college athletes (2.9 errors). These similarities in errors for tandem foam reinforce the notion that the highest degree of postural instability was in this stance as quantified by a force platform measuring sway. Together, these non-significant findings indicate that the least and most difficult conditions were not able to distinguish between skating disciplines in this sample. The overall body of evidence surrounding the BESS promotes its use to assess postural stability for baseline testing, musculoskeletal injury, and head injury purposes. Comparative findings have been reported when examining head injuries, where the bipedal firm and single leg firm stances do not differ between concussed and non-concussed athletes.

There were weak negative correlations between age and bipedal foam and single leg firm conditions, implying that as age increased, errors decreased. Previous findings suggest a stronger relationship between age and BESS performance in adolescents and young adults, where college athletes scored significantly better than both high school and youth athletes. However, in this sample of elite young athletes, these relationships were less extreme. Youth athletes can increase balance and postural stability through sports and resistance training, and these participants’ athletic development and postural stability likely is greater than comparable non-athletes of the same age. Similarly, there has been an increase in agility and strength reported in more advanced skaters (junior and senior level) compared to novices. These findings suggest early sport specialization and a focus on strength and balance training throughout the stages of development in young athletes can improve athletic capabilities and performance.

Although not the primary aim of the study, sex differences in BESS performance were assessed due to the frequency of this comparison in the current figure skating testing literature. There was no statistically significant difference between males and females on BESS performance in any condition. This conflicts with a study of normative scores on the BESS which reported fewer errors by females than males in a similar age range of young athletes. However, they found no differences in BESS errors between collegiate-level males and females. Comparing to college athletes may be more representative of our sample, as the figure skaters compete at a national level, and for the elite, an international level. They may be more athletically mature and specialized in their discipline despite the average age of the current participants (15.4±3.5 years) matching more closely to the high school cohort.

There were some limitations to the present study. Participants’ injury history, and previous musculoskeletal or head injuries were not available and may have been confounding variables which altered BESS performance. The performance of the BESS barefoot or shod, does not consider how skaters balance while in their skating boot and on the ice, and limits the authors’ ability to understand their balance on-ice. Also, access to the skaters’ level (e.g., novice, junior, senior) was not provided, which could have provided further understanding of the role of experience or skill level in postural stability. Age was relied upon for this comparison, but age does not necessarily relate to skill level. Future research should compare static postural stability such as the BESS across levels in figure skating. This could provide athletes, clinicians, and coaches information to use for training programs, injury rehabilitation, and talent identification.

**CONCLUSION**

The results of this study support the hypothesis that BESS performance in some stances varies based on skating discipline in elite figure skaters. Singles skaters tended to display the least errors, specifically on the bipedal foam and tandem firm stances, but ice dancers had fewer errors than singles in a more challenging stance, the single leg foam. These findings suggest the idea that static postural stability can vary across skating discipline, and therefore using

**Table 2. Errors on Balance Error Scoring System by Sex**

<table>
<thead>
<tr>
<th>BESS Condition</th>
<th>Females (n=213)</th>
<th>Males (n=145)</th>
<th>All (n=358)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bipedal Firm</td>
<td>0.00±0.00</td>
<td>0.02±0.18</td>
<td>0.01±0.12</td>
</tr>
<tr>
<td>Bipedal Foam</td>
<td>0.62±1.61</td>
<td>0.60±1.60</td>
<td>0.62±1.60</td>
</tr>
<tr>
<td>Single Leg Firm</td>
<td>2.14±2.40</td>
<td>1.98±2.11</td>
<td>2.08±2.28</td>
</tr>
<tr>
<td>Single Leg Foam</td>
<td>4.56±3.46</td>
<td>4.86±3.32</td>
<td>4.68±3.40</td>
</tr>
<tr>
<td>Tandem Firm</td>
<td>2.51±2.92</td>
<td>2.78±3.11</td>
<td>2.62±2.99</td>
</tr>
<tr>
<td>Tandem Foam</td>
<td>3.68±2.62</td>
<td>3.64±2.45</td>
<td>3.66±2.55</td>
</tr>
</tbody>
</table>

BESS, balance error scoring system.
the BESS to test postural stability off-ice in figure skaters may provide insight on performance improvement and injury risk. Additional factors such as age and sex do not tend to display a strong association with BESS performance in this population.

DECLARATION OF CONFLICTING INTERESTS

The authors have no conflicts of interest to report.

FUNDING

The authors received no financial support for the research, authorship, and/or publication of this article.

Submitted: February 03, 2023 CDT, Accepted: May 06, 2023 CDT

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-NC-4.0). View this license's legal deed at https://creativecommons.org/licenses/by-nc/4.0 and legal code at https://creativecommons.org/licenses/by-nc/4.0/legalcode for more information.
REFERENCES


Original Research

Maximal and Explosive Muscle Strength During Hip Adduction Squeeze and Hip Abduction Press Test Using A Handheld Dynamometer: An Intra- and Inter-tester Reliability Study

Lasse Ishøi1,2, Kristian Thorborg1, Laura Krohn1, Lars Louis Andersen3, Asger Møller Nielsen1, Mikkel Bek Clausen1,2

1 Sports Orthopedic Research Center – Copenhagen (SORC-C), Department of Orthopedic Surgery, Copenhagen University Hospital, 2 National Research Centre for the Working Environment, 3 University College Copenhagen

Keywords: Groin pain, Hip strength, Reliability, Squeeze test, Rate of force development

https://doi.org/10.26603/001c.83259

International Journal of Sports Physical Therapy

Background

Hip adduction and abduction muscle function plays an important role for risk of groin pain in athletes. Maximal isometric strength can be obtained clinically using a handheld dynamometer. However, in very strong athletes this is challenging, as external fixation of the dynamometer is needed for reliable measures. An alternative to unilateral testing, is the long-lever hip adduction squeeze test and a novel bilateral hip abduction press test. While promising intra-tester reliability has been found for maximal strength during the long-lever hip adduction squeeze test, inter-tester reliability may be more challenging during both maximal and explosive strength measurements.

Hypothesis/purpose

The aim of the present study was to assess intra- and inter-tester reliability of maximal, and explosive strength during the long lever hip abduction squeeze test and the long lever hip abduction press test in healthy adults using a hand-held dynamometer.

Study design

Intra- and interrater reliability study.

Methods

Forty-nine healthy subjects were included for intra- (n=20) and inter-tester reliability (n=29). Subjects performed the hip adduction long lever squeeze test and the bilateral hip abduction press test in a randomized order. Maximal isometric strength and early (0-100 ms) and late (0-200 ms) phase rate of force development (explosive muscle strength) was obtained using a hand-held dynamometer. Relative reliability for all tests was assessed using ICC2,1 two-way mixed model with absolute agreement, thereby taking bias between testers into account.

Results

Maximal isometric strength showed good intra- and inter-tester reliability for adduction (ICC: 0.93-0.97) and abduction (ICC: 0.88-0.92). For 0-200 ms rate of force development, both the squeeze and press test showed good intra-tester reliability (ICC: 0.85-0.87), whereas inter-tester reliability was good for hip adduction squeeze (ICC: 0.75) and moderate for hip abduction press (ICC: 0.71). For 0-100 ms rate of force development, the hip abduction press test showed good intra-tester reliability (ICC: 0.78). Remaining tests for intra- and inter-tester reliability showed moderate reliability (ICC: 0.50-0.71).

* Corresponding Author:
Lasse Ishøi
Copenhagen (SORC-C), Bispebjerg Bakke 23, 2400 Copenhagen, Denmark
Lasse.ishoei@regionh.dk
Conclusion

Assessment of maximal isometric strength in hip adduction squeeze and abduction press test showed good intra- and inter-tester reliability, whereas only 0-200 ms rate of force development demonstrated good intra-tester reliability of both tests. Therefore, rate of force development should preferably be conducted by the same tester, while the long lever squeeze and press test can reliably be used within- and between testers to measure maximal isometric strength.

Level of Evidence

3

©The Author(s)

INTRODUCTION

Groin pain is common in sport affecting up to 50% of football and ice hockey players during a season.1-3 Low hip adduction strength increases the risk of adductor-related injuries in male football players,4,5 whereas low eccentric hip adduction strength relative to abduction strength is associated with adductor muscle strains in ice hockey players.6 In addition, a decrease in hip adduction strength precedes groin pain,7 and is also observed in football players with prior season longstanding groin pain >6 weeks.8 Therefore, researchers have proposed evaluation of hip adduction and abduction strength, including the adduction/abduction ratio, as a clinical screening tool to detect early groin pain, aid secondary prevention strategies,9 and guide readiness to play following groin injuries.10-12

Simple and reliable methods to measure unilateral hip adduction and abduction strength using hand-held dynamometry was described in 2010 by Thorborg et al.13 Since then, several additional studies have confirmed these findings highlighting the clinical usefulness of handheld dynamometry.14-17 One of the advantages of handheld dynamometry is the ease of use without the need for a comprehensive setup, however, for valid and reliable measures the tester must possess enough strength compared to the force output of the patient or athlete to be able to fixate the dynamometer.18 If this is not the case, the dynamometer will move indicating that concentric rather than maximal isometric strength is obtained, leading to a lower strength output.19 Thorborg et al.18 showed how inadequate upper body strength of the tester led to lower tested maximal isometric strength compared to a tester with high upper body strength. One way to overcome this is to externally fixate the handheld dynamometer against a wall14 or using a rigid belt.15,20 which has been applied successfully in several studies. Such a setup may not always be feasible, which calls for alternative ways of estimating hip adduction and abduction strength in strong athletes without the need for external fixation. Healthcare professionals working with athletes have to a wide extent adopted the bilateral hip adduction squeeze test in various forms21 which provides a gross measure of hip adduction strength and function.11,22,23 The long lever hip adduction squeeze test elicits high hip adduction torque,24 and is thus considered appropriate to measure maximal strength17 and to stress the groin tissue.25 The long lever hip adduction has previously been applied to obtain strength profiles of both elite football26 and ice hockey players,11 and to aid secondary prevention of groin injuries.9 Previous studies have only established intra-tester reliability (ICC: 0.92; SEM %: 4.5).17 Similarly, bilateral hip abduction strength can be measured using the long lever press test; a novel test which has been utilized in elite ice hockey players.11 However, determination of intra- and inter-tester reliability is lacking in the literature.

Assessment of rate of force development (RFD), the ability to rapidly produce force over a short period of time,27 may provide value in the management of patients with hip and groin pain beyond measures of maximal strength. In patients who have undergone hip arthroscopy for femoroacetabular impingement syndrome, hip flexion rate of torque development was lower in the operated compared to the non-operated hip despite maximal hip flexion strength being normalized,23 a tendency which is also present in other types of musculoskeletal pain conditions.28 Therefore, assessment of hip adductor and abductor RFD may aid return-to-play decisions, as sports activities such as kicking, skating, and change of direction rely on rapid force development.29-33 Good test-retest reliability of bilateral hip adduction and abduction RFD using the Groin-Bar testing system (Vald Performance, Albion, Australia) and a user-independent portable dynamometer have been observed.34,35 Both studies applied a short lever arm by placing the force pads between the knees rather than ankles, thereby limiting the torque production across the hip joint17 and compromising detection of groin pain.25

The aim of the present study was to assess intra- and inter-tester reliability of maximal, and explosive strength during the long lever hip adduction squeeze test and the long lever hip abduction press test in healthy adults using a hand-held dynamometer.

METHODS

STUDY DESIGN AND SUBJECTS

Intra- and inter-tester reliability of maximal isometric strength and rate of force development of 0-100 ms (RFD100) and 0-200 ms (RFD200) was assessed with a handheld dynamometer during the long lever hip adduction squeeze test17 and a novel long lever bilateral hip abduction press test using a rigid belt.11 All subjects were included by convenience sampling from two different setting. Subjects for the intra-tester part were included from the Physiotherapy Department at Metropolitan University College, Den-
mark, whereas subjects for the inter-tester part were included from sub-elite sports clubs in the Capital Region of Denmark (Figure 1). All subjects were between 18-40 years old, and subjects included from sport clubs for the inter-tester reliability part had to be injury-free at the time of testing. Exclusion criteria for all subjects were any current pain in the hip and groin region, knee, or low back considered to influence their ability to exert a maximal and rapid muscle contraction.

The reporting adheres to the Proposed Guidelines for Reporting Reliability and Agreement Studies (GRRAS) and approval by the Ethics Committee of the Capital Region, Denmark (16041360) was obtained prior to commencement. All subjects gave their written informed consent in accordance with the Declaration of Helsinki.

TESTERS
All three testers (one for the intra-tester part and two for the inter-tester part) were final year physiotherapist students. The tester involved in the intra-tester part had a previous half year experience with handheld dynamometers but not specifically related to the tests applied in this study, whereas the testers for the inter-tester part had no previous experience with handheld dynamometry before being involved in this study. All testers received 1-2 hours of practice supervised by LI and KT, followed by self-practice for maximum of one week. LI and KT approved all testers prior to data collection.

DATA COLLECTION
The force signal for bilateral hip adduction and abduction using a hand-held dynamometer (HHH) were recorded with a sampling frequency of 100 Hz (Hoggan microFET2, Hoggan Scientific L.L.C., Salt Lake City UT, USA). From this maximal isometric strength and rate of force development for 0-100 ms and 0-200 ms (detailed below) was determined. The test procedure for the long lever squeeze test followed a standardized reliable set-up (ICC2,1; 0.90-0.97). With the subject in supine position and hips and knees straight (0° flexion), the dynamometer was placed 5 cm proximal to the most prominent point on the medial malleolus of the dominant leg. The participant’s legs were slightly abducted corresponding to the length of the testers forearm and the dynamometer. The forearm and the HHD was placed between the ankles, and the subject was subsequently instructed to perform a bilateral hip adduction squeeze. The bilateral hip abduction press test was conducted with the subject in supine position. The HHD was placed 5 cm proximal to the most prominent point of the lateral malleolus and fixated with a rigid belt around the legs. The participant’s legs were slightly abducted corresponding to the position of the hip adduction squeeze test. The subject was instructed to press against a rigid belt placed around the ankles and the HHD by performing a bilateral hip abduction. During both tests, subjects were instructed to push or press as “fast and hard” as possible, and to keep pushing or pressing until instructed to relax (approximately 3-4 s).

Both tests consisted of two submaximal trials at 50% and 100% of self-perceived maximum effort, followed by three valid trials, separated by one-minute rest. After each trial subjects were asked to score pain in the hip and groin on an 11-point Numerical Rating Scale (NRS 0-10). The test was terminated in case of pain >3. For the inter-tester part, the two testers were blinded to strength measures obtained by the other tester.

The force data were transmitted from the HHD to a commercial software program (TBS, Hoggan, Scientific L.L.C., Salt Lake City, USA) and extracted to a custom-made spreadsheet (Microsoft Excel, USA) for analyses. Force was recorded in Newton (N). RFD100 (100 ms) and RFD200 (200 ms) was calculated as the mean change in N per second in each time interval (100 ms and 200 ms) with the onset threshold of force (t=0 ms) set at 6.7 N above baseline force. Maximal isometric strength was determined as the peak value.

The sequence of measurements and testers were randomized, and an identical sequence was used during the retest session. Subjects rested 15 minutes between the test-retest sessions.

**Figure 1. The flow of participants.**
STATISTICAL METHODS
Systematic bias between-sessions (intra-tester) and between-testers (inter-tester) was assessed as differences in mean values using paired t-tests with a significance level set at $p<0.05$. Relative reliability was assessed as Intraclass Correlation Coefficient (ICC) with a two-way random effects model and absolute agreement definition using the “irr” package in R. Absolute agreement was chosen as this examines the relative reliability without incorporating a systemic error term (this means that the ICC value reflects any systematic variation between testers/sessions, and was thus used because of systematic bias was detected). The relative reliability was interpreted as poor (ICC<0.50), moderate (0.50≤ICC≤0.75), and good (ICC>0.75). Absolute reliability was expressed as 1) the standard error of measurement (SEM) calculated as $SD_{pool} \times \sqrt{1-ICC}$, and 2) SEM% calculated as: $\left(\frac{SEM}{mean_{pool}}\right) \times 100$. Minimal detectable change % (MDC%) was calculated using SEM%, both at the individual level (MDCind% = SEM% × $1.96 \times \sqrt{2}$) and group level (MDCgroup% = SEM% $\times \frac{1.96 \times \sqrt{2}}{\sqrt{n}}$, where $n$ is the sample). Bland-Altman plots were constructed using the “BlandAltmanLeh” package in R. All statistical analyses were calculated in R Studio (v. 3.6.1).

RESULTS
PARTICIPANTS
Fifty-four subjects were recruited. Four subjects were excluded due to technical errors in the software program used for data collection and one subject was excluded due to pain during testing that affected performance. Therefore, a total of 49 subjects were included. Twenty subjects (males: 10; females: 10, mean age ± SD: 25.5 ± 4.2, mean body mass ± SD: 75 ± 12.5, mean height ± SD: 177.7 ± 12.8) were included for intra-tester reliability. Due to loss of data, twenty subjects were included for abdution testing, whereas 19 subjects were included for abduction testing. Twenty-nine male subjects (mean age ± SD: 23.5 ± 5.5, mean body mass ± SD: 89 ± 11.0, mean height ± SD: 187.5 ± 12.5) were included for inter-tester reliability (Figure 1).

INTRA-TESTER RELIABILITY
No systematic bias between sessions was observed ($p>0.62$). Good intra-tester reliability was observed for peak force and RFD200 for both hip adduction squeeze and hip abduction press (ICC 0.83-0.97), whereas only hip abduction press showed good reliability for RFD100 (ICC 0.76) (Table 1). The absolute intra-tester reliability (SEM %) for peak force, RFD100 and RFD200 ranged from 5.9-7.7 %, 16.9-25.2 %, and 10.5-11.8 %, respectively (Table 1). MDCind % for peak force, RFD100 and RFD200 ranged from 10.9-21.2 %, 47.0-69.9 %, and 29.0-52.7 %, respectively (Table 1). Bland-Altman plots are depicted in Figure 2.

INTER-TESTER RELIABILITY
Systematic bias was observed between testers for all tests except adduction squeeze MVC test (Table 2). Good inter-tester reliability was observed for peak force for both hip adduction squeeze and hip abduction press (ICC 0.91-0.95). For rate of force development, all measures showed moderate reliability (ICC 0.5-0.75) (Table 2). The absolute inter-tester reliability (SEM %) for peak force and rate RFD measures ranged from 5.7-6.2 % and 10.8-18.3 %, respectively, while MDCind % ranged from 15.8-17.1 and 30.0-50.8 for peak force and RFD measures, respectively (Table 2). Bland-Altman plots are depicted in Figure 3.

DISCUSSION
This study introduces a new and reliable way to measure maximal and explosive bilateral hip abduction strength using a hand-held dynamometer using a test setup unaffected by the tester’s strength. The current findings show that maximal isometric strength can be reliably measured within and between testers with low measurement error in both tests, while late-phase RFD (0-200ms) for both tests also showed good intra-tester reliability. All remaining RFD measures showed moderate reliability but with high imprecision based on the confidence intervals crossing threshold of <0.50 signifying poor reliability, and wide Limits of Agreement. This suggest that measures of maximal isometric hip abduction and abduction strength can be obtained by different testers, whereas late-phase RFD only should be obtained by the same tester as the confidence interval for inter-tester reliability is too large to ensure reliable measurements. The early-phase RFD shows large measurement error regardless of whether this was obtained by the same or different testers, and thus provide little to no utility.

MAXIMAL ISOMETRIC STRENGTH
In elite sport settings there is often a need to efficiently measure athletes’ lower body strength in a short time frame, such as during periodic testing. This can be achieved by using handheld dynamometers which easily accommodate various testing setups, such as unilateral hip adduction (ICC: 0.93) and abduction (ICC: 0.97). In strong athletes, external fixation may be needed dependent on the tester’s upper body strength to obtain valid measures, but this is not always feasible due to time constraints when large cohorts need testing. An alternative to unilateral testing of the hip is the long lever hip abduction squeeze test and the bilateral hip abduction press test. Both tests are quick and can easily be applied in even strong athletes without the need for an external fixation setup, thus providing a feasible way of measuring hip abduction and abduction strength in athletes. The current findings for intra-tester reliability of the long lever squeeze test are consistent with previous findings reported by Light et al. (ICC: 0.92; SEM %: 4.5). Other hip squeeze test variations have shown similar reliability using either a sphygmome-
<table>
<thead>
<tr>
<th>Isometric hip actions</th>
<th>Session 1 mean (SD)</th>
<th>Session 2 mean (SD)</th>
<th>Difference session 1-session 2 peak [CI 95%]</th>
<th>Paired t-test p-value</th>
<th>ICC (2,1) * [CI 95%]</th>
<th>SEM</th>
<th>SEM %</th>
<th>MDC\textsubscript{ind} (%)</th>
<th>MDC\textsubscript{group} (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MVC - N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HADD</td>
<td>175.3 (48.3)</td>
<td>173.5 (47.3)</td>
<td>1.8 [-7.3; 11.0]</td>
<td>0.68</td>
<td>0.92 [0.81; 0.97]</td>
<td>13.3</td>
<td>7.7</td>
<td>21.2</td>
<td>4.9</td>
</tr>
<tr>
<td>HABD</td>
<td>156.2 (33.7)</td>
<td>156.4 (37.8)</td>
<td>-0.2 [-4.6; 4.2]</td>
<td>0.91</td>
<td>0.97 [0.92; 0.99]</td>
<td>6.1</td>
<td>3.9</td>
<td>10.9</td>
<td>2.5</td>
</tr>
<tr>
<td>RFD 0-100 – N/s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HADD</td>
<td>700.0 (300.7)</td>
<td>708.0 (282.6)</td>
<td>-7.9 [-132.6; 116.7]</td>
<td>0.89</td>
<td>0.62 [0.23; 0.84]</td>
<td>177.5</td>
<td>25.2</td>
<td>69.9</td>
<td>16.0</td>
</tr>
<tr>
<td>HABD</td>
<td>844.1 (327.4)</td>
<td>838.7 (257.9)</td>
<td>5.3 [-91.3; 102.0]</td>
<td>0.91</td>
<td>0.76 [0.49; 0.90]</td>
<td>142.5</td>
<td>16.9</td>
<td>47.0</td>
<td>10.8</td>
</tr>
<tr>
<td>RFD 0-200 – N/s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HADD</td>
<td>535.2 (179.3)</td>
<td>526.8 (159.1)</td>
<td>8.5 [-36.1; 53.1]</td>
<td>0.69</td>
<td>0.86 [0.67; 0.94]</td>
<td>62.6</td>
<td>11.8</td>
<td>32.7</td>
<td>7.5</td>
</tr>
<tr>
<td>HABD</td>
<td>574.0 (161.3)</td>
<td>581.6 (134.8)</td>
<td>-7.6 [-48.8; 33.6]</td>
<td>0.70</td>
<td>0.83 [0.62; 0.93]</td>
<td>60.5</td>
<td>10.5</td>
<td>29.0</td>
<td>6.7</td>
</tr>
</tbody>
</table>

MVC (Maximal voluntary contraction); Nm (Newton meter); Nm/s (Newton meter/second); ICC (Intraclass Correlation Coefficient); SEM (Standard Error of Measurement); MDC\textsubscript{ind} (Minimal Detectable Change on an individual level); MDC\textsubscript{group} (Minimal Detectable Change on a group level); SD (Standard Deviation); HABD (hip abduction); HADD (hip adduction); RFD (Rate of force development); *ICC used for absolute assessment between session.
Table 2. Inter-tester reliability of peak force, and rate of force development for hip adduction and hip abduction (n=29)

<table>
<thead>
<tr>
<th>Isometric hip actions</th>
<th>Tester 1 mean (SD)</th>
<th>Tester 2 mean (SD)</th>
<th>Difference Tester 1-Tester 2 Mean [CI 95%]</th>
<th>Paired t-test p-value</th>
<th>ICC (2,1) * [CI 95%]</th>
<th>SEM</th>
<th>SEM %</th>
<th>MDCind (%)</th>
<th>MDCgroup (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MVC - N</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HADD</td>
<td>278.3 (57.6)</td>
<td>274.2 (62.3)</td>
<td>4.1 [-4.7; 12.9]</td>
<td>0.35</td>
<td>0.93 [0.85-0.96]</td>
<td>15.7</td>
<td>5.7</td>
<td>15.8</td>
<td>2.9</td>
</tr>
<tr>
<td>HABD</td>
<td>190.1 (38.1)</td>
<td>178.9 (36.8)</td>
<td>11.2 [6.4; 16.1]</td>
<td>&lt;0.01</td>
<td>0.91 [0.56-0.97]</td>
<td>11.4</td>
<td>6.2</td>
<td>17.1</td>
<td>3.2</td>
</tr>
<tr>
<td>RFD 0-100 - N/s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HADD</td>
<td>1511.2 (360.7)</td>
<td>1245.3 (305.7)</td>
<td>265.8 [159.1; 372.6]</td>
<td>&lt;0.01</td>
<td>0.50 [-0.01-0.77]</td>
<td>252.8</td>
<td>18.3</td>
<td>50.8</td>
<td>9.4</td>
</tr>
<tr>
<td>HABD</td>
<td>1187.7 (308.6)</td>
<td>1055.6 (283.0)</td>
<td>132.1 [44.5; 219.6]</td>
<td>&lt;0.01</td>
<td>0.64 [0.30-0.82]</td>
<td>180.6</td>
<td>16.1</td>
<td>44.6</td>
<td>8.3</td>
</tr>
<tr>
<td>RFD 0-200 - N/s</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HADD</td>
<td>1005.3 (206.2)</td>
<td>900.8 (195.7)</td>
<td>104.5 [62.4; 146.6]</td>
<td>&lt;0.01</td>
<td>0.75 [0.19-0.91]</td>
<td>103.0</td>
<td>10.8</td>
<td>30.0</td>
<td>5.6</td>
</tr>
<tr>
<td>HABD</td>
<td>759.9 (167.1)</td>
<td>704.5 (162.2)</td>
<td>50.4 [7.1; 93.6]</td>
<td>0.02</td>
<td>0.73 [0.48-0.87]</td>
<td>85.8</td>
<td>11.8</td>
<td>32.6</td>
<td>6.1</td>
</tr>
</tbody>
</table>

MVC (Maximal voluntary contraction); N (Newton); N/s (Newton/second); ICC (Intraclass Correlation Coefficient); SEM (Standard Error of Measurement); MDCind (Minimal Detectable Change on an individual level); MDCgroup (Minimal Detectable Change on a group level); SD (Standard Deviation); HABD (hip abduction); HADD (hip adduction); RFD (Rate of force development); *ICC with two-way random effects and absolute agreement definition.
ter (ICC: 0.81-0.94; SEM %: 1.60-3.27)\textsuperscript{42,43} or the Groinbar (ICC: 0.85-0.94; SEM %: 8.2).\textsuperscript{34,44} The inter-tester reliability of the long lever squeeze test has not yet been examined, but the current findings are consistent with those reported when using other variations of the squeeze test (short lever squeeze test, ICC: 0.91-0.92)\textsuperscript{45,46} and peak force measured unilaterally (ICC:0.92-0.94).\textsuperscript{15,47} This study is the first to examine intra- and inter-tester reliability of the long lever hip abduction press test. The current data are comparable to previous literature reporting on reliability for hip abduction strength testing, using a bilateral short-lever test in the Groinbar (ICC: 0.82)\textsuperscript{34} or performed in a user-independent portable device (ICC: 0.91).\textsuperscript{35}

RATE OF FORCE DEVELOPMENT

The present study is the first to establish reliability data on RFD measures in the long-lever squeeze test and the novel hip abduction press test. These data indicate that late-phase RFD (0-200ms) showed good intra-tester reliability for both tests, while remaining RFD measures showed poor reliability. Two previous studies have reported reliability for RFD measured during bilateral hip adduction squeeze and abduction press tests; both studies applied a short lever position, precluding direct comparison with the present study. Desmyttere et al. found moderate to good intra-tester reliability of peak RFD using a 200 ms moving average during bilateral short lever adduction and abduction testing using the Groinbar (ICC 0.81 [95 % CI: 0.65-0.90] and 0.68 [95 % CI: 0.42-0.83]).\textsuperscript{34} In contrast to the present study, good reliability has been reported for early-phase RFD (0-100 ms) using a user-independent device to measure bilateral hip adduction and abduction strength in the short lever position, while early-phase RFD has also showed good reliability in both abduction and adduction when performed unilaterally.\textsuperscript{14} These discrepancies in testing properties may be explained by different set-ups, but suggest that, if explosive bilateral hip adduction and abduction strength in the long-lever position is of interest across athletes or over time, this should be obtained by the same tester and only late-phase RFD should be considered. The lower reliability for early-phase RFD in our study could be explained by the timeframe of 0-100 ms being too short to coordinate a simultaneous fast contraction with both legs.

APPLICATION TO CLINICAL PRACTICE

The procedure examined in this study is feasible to be included in clinical practice, as a quick and simple method for testing maximal isometric strength and late-phase explosive strength in strong athletes without potential bias related to the tester size or strength.\textsuperscript{15,18} In research settings, evaluation of changes is often of interest at a group level. The MDC\textsubscript{group} % reported in the present study is calculated based on 20 participants for the intra-tester and 29 participants for the inter-tester part. Using an equivalent sample size, changes at group level would be detected with 5 % for peak force and 7.5 % for late-phase RFD (0-200 ms) can be detected with 95 % certainty,\textsuperscript{40} when using both a single- and multiple-tester setup. When applied on individual patients or athletes in the clinical setting, changes would have to exceed ~10-20 % for peak force and ~30 % for late-phase RFD (0-200 ms) to be detected with 95 % certainty. This further indicated by the width of the Limits of Agreement of the Bland Altman plot, suggesting that high uncertainty between two measures is expected. Although these numbers are high they should be considered in relation to findings in injured athletes. As an example, prior-season groin pain lasting more than six weeks, is related to in average 19 % decrease in peak strength during the long lever squeeze test in currently uninjured players.\textsuperscript{8} For explosive muscle strength, even larger deficits may be expected; Nunes et al. observed a decrease in explosive strength of 33 % in the hip abductors in females with patellofemoral pain.\textsuperscript{48} However, explosive hip adduction and abduction strength and its relation to groin pain have not yet been established.

METHODOLOGICAL CONSIDERATIONS AND LIMITATIONS

A limitation in the present study is the systematic bias observed in the inter-tester data, which could have been affected by dynamometer placement, instruction, or encouragement during testing. However, it might be possible to minimize these differences by further standardization of testing procedures and more focus on tester calibration. Since the adductor muscle force angle depends on the testers forearm length, this may also contribute to the systematic bias, thus influencing inter-tester reliability. However, the authors did not collect data related to forearm length, thus it cannot be concluded if this affected the reliability in the present study. A further limitation is that only healthy subjects were included. Thus, future studies should be performed including athletes with groin pain, to understand if groin pain may affect reliability.

CONCLUSION

Assessment of maximal isometric strength in hip adduction squeeze and abduction press test showed good intra- and inter-tester reliability, whereas only 0-200 ms rate of force development demonstrated good intra-tester reliability of both tests. Therefore, rate of force development should preferably be conducted by the same tester, while this is less important for isometric peak torque.

CONFLICTS OF INTEREST

The authors report no conflicts of interest.

Submitted: December 16, 2022 CDT, Accepted: May 16, 2023 CDT
Figure 2. Bland-Altman plots for inter-tester reliability.
Figure 3. Bland-Altman plots for inter-tester reliability.
REFERENCES


The Validity and Reliability of a Smartphone Application for Break-Point Angle Measurement during Nordic Hamstring Exercise

Toshiaki Soga1, Shota Yamaguchi1, Takayuki Inami1, Hiromi Saito1, Nadaka Hakariya1, Naoto Nakaichi1, Sho Shinohara1, Kei Akiyama1, Norikazu Hirose1a

1 Graduate School of Sport Sciences, Waseda University. 2 Institute of Physical Education, Keio University. 3 Faculty of Sport Sciences, Waseda University

Keywords: injury prevention, injury risk factor, eccentric force

Background

A recently developed smartphone application (Nordic Angle) allows the automatic calculation of the break-point angle (BPA) during Nordic hamstring exercise (NHE) without transferring the collected data to a computer. The BPA is the point at which the hamstrings are unable to withstand force. However, the validity of the BPA values obtained by this method has not been examined.

Hypothesis/Purpose

This study aimed to evaluate the validity and reliability of the Nordic Angle by comparing the BPA values of the Nordic Angle with those of two-dimensional motion analysis software that can calculate the angles and angular velocities of various joints.

Study Design

Cohort assessing Validity and Reliability

Methods

The validity of the Nordic Angle BPA data was verified by Spearman’s correlation test for consistency with the movement analysis data, and the magnitude of the correlation was indicated by rs. The agreement between these measurements was examined using the Bland-Altman analysis. The reliability of the Nordic Angle and motion analysis was examined using the intraclass correlation coefficient (ICC) (1,k) based on data from repeated trials within a day.

Results

Although the spearman correlation between the Nordic angle and the angle determined using motion analysis did not reach statistical significance (p = 0.052), a very large correlation was present (rs = 0.75). The difference between the mean values of the Nordic Angle and motion analysis was 0.4 ± 2.1°, and the limits of agreement ranged from -3.9° to 4.6°. In two BPA measurements, the Nordic Angle showed perfect reliability (ICC = 1.00, p < 0.001), while motion analysis showed nearly perfect reliability (ICC = 0.97, p < 0.001).

a Corresponding author:
Norikazu Hirose
Faculty of Sport Sciences, Waseda University, Japan
totsu_hirose@waseda.jp

Nordic Angle is a smartphone application developed in collaboration with S-CADE. Co., Ltd. and Athletic Training Laboratory at Waseda University (instructed by Professor Norikazu Hirose). The first author of the article collected and analyzed the data. However, the authors of the current study assert no conflicts as they do not receive any fees from S-CADE. Co., Ltd. for developing the Nordic Angle, and the installation of the Nordic Angle app is free.
Conclusion
The Nordic Angle, which has both validity and reliability, may be appropriate for field measurement because it allows immediate feedback of BPA and the measurement of many athletes.

Level of evidence
3b

BACKGROUND
Hamstring strain injury (HSI) occurs in many sports activities.\(^1\) For example, HSI reportedly accounted for 12% of all injuries in football,\(^2\) with recurrence rates exceeding 50%.\(^3\) It is important to prevent the initial occurrence and recurrence of HSI to maintain and improve the performance of athletes.

The risk factors for HSI include eccentric knee flexion strength,\(^4,5\) fascicle length of the biceps femoris long head,\(^2\) and a history of HSI.\(^6\) Isokinetic eccentric knee flexion strength assessment is a popular screening tool for identifying athletes who are at risk of future HSI.\(^4\) However, the use of isokinetic equipment is limited by various factors, including time, cost, and poor portability.\(^7\) Although eccentric knee flexion strength is a kinetic variable, the Nordic break-point test of the Chinese University of Hong Kong uses a kinematic variable and has recently been reported as a unique method for predicting eccentric knee flexion strength.\(^8\) As a feature of the Nordic hamstring exercise (NHE), further forward leaning of the upper trunk is associated with longer moment arm from the knee joint to the center of mass, which increases the knee flexion moment.\(^9\) The point at which the hamstring force can no longer tolerate this increased eccentric knee flexion angle begins to accelerate toward the ground is known as the break point. A very large correlation has been observed between eccentric knee flexion strength and knee flexion angle at the break point measured by smartphone application (SensorLog) (\(r = 0.88, p < 0.001\)).\(^5\) Although the break-point angle (BPA) during NHE does not directly measure eccentric knee flexion strength, the regular measurement of the BPA might indicate whether there is increased HSI risk. However, there is a disadvantage in measuring BPA by using smartphone applications, namely, the need to transfer the data to a personal computer for analysis.\(^8\) Therefore, providing immediate feedback and measuring a large number of athletes are expectedly difficult.\(^8\)

A very recent development of a smartphone application (Nordic Angle) allowed automatic calculation of BPA without transferring the collected data to a personal computer. However, the validity of the BPA values obtained by this method has not been examined. Therefore, this study aimed to examine the validity and reliability of the Nordic Angle by comparing the BPA values of the Nordic Angle with those of two-dimensional motion analysis software.

DESCRIPTION

STUDY DESIGN
After the warm-up, the participants performed 12 NHE repetitions while being recorded with an iPhone 12 high-speed camera. Based on the same movie data recorded with an iPhone camera, BPA during NHE was calculated using two-dimensional motion analysis software and the Nordic Angle (Figure 1). The validity of the Nordic Angle data was examined using the Spearman’s correlation test for consistency with motion analysis data, and the agreement between these measurements was examined using the Bland-Altman analysis. The reliability of Nordic Angle and motion analysis was examined using the intraclass correlation coefficient (ICC) (1,k) based on data from repeated trials within a day.

PARTICIPANTS
The sample size was calculated a priori based on Pearson correlation coefficient analysis (G*Power version 3.1, Heinrich Heine Universität Düsseldorf, Germany). Its input parameters (correlation \(r = 0.9,\) alpha = 0.05, power = 0.8) were set with reference to the effect size of a previous study.\(^10\) This setting resulted in a sample size of five. Therefore, seven male volunteers (age, 23.9 ± 2.0 years; height, 174.3 ± 5.4 cm; and body mass, 67.3 ± 7.9 kg; all measured in mean ± standard deviation (SD)) participated in this study. The inclusion criterion was healthy males between the ages of 18 and 35 who participate in sports activities and can perform exercises without musculoskeletal pain. The exclusion criterion was participants who could not perform NHE because of a current injury to the lower extremities. The experimental protocol was approved by the institutional review board of Waseda University’s ethical committee (approval number: 2022-345), and all procedures in this study were performed in accordance with the Declaration of Helsinki. All participants were informed of the purpose and procedure of the study, and informed consent was obtained from all participants.

PROCEDURES
Before the experiment, participants initially performed a warm-up consisting of light aerobic activity (2 minutes of alternating stepping on a 15-cm high box), a 15-second static hamstring stretch (standing hamstring stretch on one leg), and one set of 10 repetitions of dynamic hamstring stretching (forward leg swing). Hamstring stretching was performed on both legs. Thereafter, participants were given
two repetitions to gain familiarity with the movement being assessed, and then performed 12 repetitions of NHE for data collection. There was at least one minute of rest between each repetition. The experiment was conducted under the supervision of an examiner, who was qualified as a National Strength and Conditioning Association Certified Strength and Conditioning Specialist.

NORDIC HAMSTRING EXERCISE

Participants were instructed to start in a kneeling position on a bench of application of approximately 50 cm with elbows bent and hands open in front of them (Figure 1). The examiner instructed participants to keep their posture straight from the knee to the head and to lean forward "as slowly as possible". A certified examiner confirmed that the participant performed the NHE correctly.

TWO-DIMENSIONAL MOTION ANALYSIS

The speed of the camera was set to 240 frames per second (fps), and the height was approximately 0.8 m, while the camera was positioned application approximately 3 m from the right side of the participants. After the recorded movie was transported to a personal computer, two-dimensional motion analysis was performed on the obtained movie data using the Frame-DIAS V software (DKH Inc., Tokyo, Japan). Reflective markers were attached to the greater trochanter and the lateral malleolus. The knee flexion angle was calculated by digitizing the greater trochanter, center of knee joint, and lateral malleolus. The digitization of the center of the knee joint was performed visually by the examiner. The line connecting the greater trochanter with the center of the knee joint and the line connecting the center of the knee joint with the lateral malleolus created an angle that was defined as the knee flexion angle. The knee flexion angle was calculated using the following formula:

\[
\theta = \frac{a_1 b_1 + a_2 b_2}{\sqrt{a_1^2 + a_2^2 b_1^2 + b_2^2}},
\]

where \(\theta\) is the knee flexion angle, \(a_1\) and \(a_2\) are the points of the line connecting the greater trochanter with the center of the knee joint, and \(b_1\) and \(b_2\) are the points of the line connecting the center of the knee joint with the lateral malleolus.

The knee extension angular velocity was calculated using the following formula:

\[
\text{deg/s} = \frac{\theta_{t+1} - \theta_{t-1}}{2\Delta t},
\]

where \(\theta\) is the knee flexion angle, and \(\Delta\) is the time.

BPA was defined as the angle at which the knee extension angular velocity exceeded 50°/s. The kinematic data were smoothed with a Butterworth filter for two-dimensional motion analysis. The order of the Butterworth filter was as follows: (1) the characteristic was low-pass, (2) the cutoff frequency was 6 Hz, and (3) the number of extended data for the correction of both ends was 20.

STATISTICAL ANALYSIS

Values were expressed as mean ± SD. The average BPA values from motion analysis and the Nordic Angle were calculated from the BPA values of the 12 NHE repetitions performed by each participant. The Shapiro–Wilk test was performed for the analysis of the normal distribution of the BPA data of motion analysis and the Nordic Angle. As a result, there was no confirmed normal distribution in the BPA of the Nordic Angle. Because BPA data were

Figure 1. Nordic Angle screen information during the Nordic hamstring exercise.

1: Knee flexion angle (0° indicates the fully extended position of the knee); 2: Break-point angle determined by Nordic Angle; 3: Current time from movie data start.
The validity and reliability of a smartphone application for break-point angle measurement during Nordic hamstring exercise

![Image](image_url)

**Figure 2. The correlation between break-point angle of the Nordic Angle and that of motion analysis.**

not normally distributed, the validity of the Nordic Angle data was examined using the Spearman’s correlation test for consistency with the motion analysis data. Although normality was not confirmed in the BPA data, the agreement between these measurements was examined using the Bland–Altman analysis. The Wilcoxon signed-rank test was performed to confirm the difference in BPA between the Nordic Angle and motion analysis data. The limits of agreement in the Bland–Altman analysis were calculated by multiplying SD by ±1.96. The reliability of the Nordic Angle and motion analysis was examined using ICC (1,k). Correlation is indicated as rs for Spearman’s correlation test, and its magnitude of the correlation was established based on the following criteria: rs = 1, perfect correlation; 1 ≥ rs ≥ 0.9, nearly perfect; 0.9 > rs ≥ 0.7, very large; 0.7 > rs ≥ 0.5, large; 0.5 ≥ rs ≥ 0.3, moderate; 0.3 > rs ≥ 0.1, small; and 0.1 ≤ rs, trivial. Statistical analyses were performed using SPSS version 27 (IBM SPSS, Armonk, NY, USA). The significance level was set at p < 0.05.

**OUTCOMES**

**THE VALIDITY OF THE NORDIC ANGLE**

**Figure 2** showed the correlation between the BPA of Nordic Angle and that of motion analysis. Although the spearman correlation between the Nordic angle and the angle determined using Motion analysis did not reach statistical significance (p = 0.052), it represents a very large correlation (rs = 0.75).

**Figure 3** shows the agreement between the BPA of Nordic Angle and that of motion analysis. The BPAs of Nordic Angle and motion analysis were 60.2 ± 10.4° and 59.9 ± 9.9°, respectively, and there was no significant difference between the Nordic Angle and motion analysis data (p = 0.50). The difference between the mean values of the Nordic Angle and motion analysis was 0.4 ± 2.1°, and the limits of agreement ranged from -3.9° to 4.6°.

**Figure 3. The agreement between break-point angle of Nordic Angle and that of motion analysis, and the limits of agreement.**

The solid line showed the difference between the mean values of Nordic Angle and motion analysis and the dotted line showed the mean values in motion analysis.

**THE RELIABILITY OF NORDIC ANGLE AND MOTION ANALYSIS**

In two BPA measurements, the Nordic Angle showed perfect reliability (ICC = 1.00, p < 0.001), while motion analysis showed nearly perfect reliability (ICC = 0.97, p < 0.001).

**DISCUSSION**

This study aimed to examine the validity and reliability of the Nordic Angle by comparing the BPA values of the Nordic Angle with those of two-dimensional motion analysis software. The main outcomes of this study are as follows: (a) Although the spearman correlation between the Nordic angle and the angle determined using Motion analysis did not reach statistical significance (p = 0.052), it represents a very large correlation (rs = 0.75); (b) the mean difference between BPA data from the Nordic Angle and motion analysis was 0.4 ± 2.1°, with no significant difference; (c) the limits of agreement ranged from -3.9° to 4.6°; and (d) the Nordic Angle showed perfect reliability (ICC = 1.00, p < 0.001).

The results of this study indicated a non-significant, yet very large correlation between Nordic Angle and motion analysis (Figure 1). In addition, no significant difference was observed between the BPA value of the Nordic Angle and that of the motion analysis (Figure 2), suggesting that BPA calculation by the Nordic Angle was valid. As previously reported, HSI risk might be predicted by eccentric knee flexion strength during isokinetic dynamometry. A very large correlation was observed between eccentric knee flexion strength during isokinetic dynamometry and BPA during NHE. If there is also a strong association between BPA measured by the Nordic Angle and eccentric knee flexion strength, regular BPA measurements using the Nordic Angle may provide insights into the potential increase or decrease in HSI risk.
This study also assessed the limits of agreement between measures (Figure 3). Delahunt et al. investigated the changes in eccentric knee flexion strength during isokinetic dynamometry and BPA during NHE after a 6-week NHE training in recreationally active men. Their results showed that eccentric knee flexion strength increased by 15%, and BPA changed by only -5.6°. In the current study, the limits of agreement of BPA ranged from -3.9° to 4.6°, which may either overestimate or underestimate eccentric knee flexion strength. Therefore, measuring BPA by the Nordic Angle with three to five trials and using average of these trials for BPA calculation might be recommended.

In this study, the Nordic Angle demonstrated perfect reliability (ICC = 1.00, p < 0.001). On the other hand, while the motion analysis also showed nearly perfect reliability (ICC = 0.97, p < 0.001), there was a slight difference between the first and second analysis. Thus, the Nordic Angle, which demonstrates acceptable validity and reliability, may be appropriate for field measurement because it allows immediate feedback of BPA and measurement in many athletes.

It is important to note that this study did not examine the relationship between BPA using the Nordic Angle and eccentric knee flexion strength during isokinetic dynamometry. Therefore, it remains unclear whether BPA calculated using the Nordic Angle reflects eccentric knee flexion strength. Since eccentric knee flexion strength during isokinetic dynamometry at 60°/s is an HSI risk factor, the relationship between BPA calculated by the Nordic Angle and eccentric knee flexion strength at 60°/s should be examined.

LIMITATIONS

This study has several limitations. The first is the small sample size. This created two clusters and risked overestimating the validity of BPA by the Nordic Angle. In addition, the sample size was set based on the effect size of a previous study when determining the sample size, but the effect size in this study was not as large as that of that previous study. Future studies are needed to test the validity of the BPA for the Nordic Angle when the sample size is increased. The second was the use of Bland-Altman analysis on non-normally distributed data. The limits of agreement in the Bland-Altman analysis might not be exact values. It is expected that the limits of agreement in the Bland-Altman analysis might be tested in the future when the sample size is increased. The third is that the measurement method was only applied to two-dimensional motion analysis. In the future, the validity of three-dimensional motion analysis needs to be examined. The final limitation is that data from repeated trials within a day make it impossible to draw conclusions about inter-session testing.

CONCLUSION

In conclusion, this study examined the validity and reliability of the Nordic Angle by comparing the BPA values of the Nordic Angle with those of two-dimensional motion analysis software. The results indicate a very large (non-significant) correlation between the Nordic Angle measurement and the angle determined using Motion analysis, however, the mean difference between the Nordic Angle measurement and motion analysis data was not statistically significantly different (0.4 ± 2.1°). The limits of agreement ranged from -5.9° to 4.6° and the Nordic Angle showed perfect reliability (ICC = 1.00, p < 0.001). The Nordic Angle might be appropriate for field measurement because it provides immediate feedback regarding the BPA in many athletes. However, the small sample size suggests that the validity of the BPA may be overestimated.

DISCLOSURES

Nordic Angle is a smartphone application developed in collaboration with S-CADE. Co., Ltd. and Athletic Training Laboratory at Waseda University (instructed by Professor Norikazu Hirose). The first author of the article collected and analyzed the data. However, the authors of the current study assert no conflicts as they do not receive any fees from S-CADE. Co., Ltd. for developing the Nordic Angle, and the installation of the Nordic Angle app is free.

ACKNOWLEDGEMENTS

The authors acknowledge the facilities and assistance of the Graduate School of Sport Sciences, Waseda University. The experiments complied with the current laws of the country in which they were conducted. The datasets generated and/or analyzed during the current study are not publicly available, but are available from the corresponding author, who was an organizer of the study. We would like to thank Professor Susumu S. Sawada for his advice on data analysis and statistical analysis. This work was supported by the JST SPRING (grant number JPMJSP2128).

Submitted: December 29, 2022 CDT, Accepted: May 24, 2023 CDT

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-NC-4.0). View this license’s legal deed at https://creativecommons.org/licenses/by-nc/4.0 and legal code at https://creativecommons.org/licenses/by-nc/4.0/legalcode for more information.

International Journal of Sports Physical Therapy
REFERENCES


Original Research

Evaluating Psychometric Properties of the International Knee Documentation Committee Subjective Knee Form in a Heterogeneous Sample of Post-Operative Patients

Ryan D Richardson1, Madeline P Casanova1, Ashley J Reeves1, Seungho Ryu1, Adam C Cady2, Russell T Baker1•a

1 WWAMI Medical Education, University of Idaho, 2 Woodland Hills Orthopedics, Kaiser Permanente

Keywords: structural validity, psychometric analysis, knee pathology, patient reported outcomes

https://doi.org/10.26603/001c.83940

International Journal of Sports Physical Therapy

Background
The International Knee Document Committee Subjective Knee Form (IKDC-SKF) is a patient-reported outcome measure used in orthopedics and sports medicine. Further psychometric assessment is necessary to confirm measurement properties in a large, heterogenous sample.

Purpose
The purpose of the study was to assess the psychometric properties of the IKDC-SKF in a large, heterogenous sample.

Study Design
Cross-Sectional Study

Methods
An exploratory factor analysis (EFA) was conducted to identify a sound latent structure and to assess internal consistency in a large sample of patients who underwent knee arthroscopy. A confirmatory factor analysis (CFA) was conducted to confirm structural validity. Multi-group invariance was conducted to assess factorial stability across sex and age groups, while longitudinal invariance procedures were performed to assess stability over time.

Results
A 3-factor, 9-item IKDC-SKF short form was identified with EFA procedures. The model was confirmed with CFA (CFI = 0.983; TLI = 0.975; IFI = 0.983; RMSEA = 0.057), while a sound 2-factor, 6-item model was also identified (CFI = 1.0; TLI = 0.999; IFI = 1.0; RMSEA = 0.11). The 9-item IKDC-SKF short form was invariant across groups but not time; removal of a single item (i.e., 8-item IKDC-SKF short form) resulted in longitudinal invariance. The 6-item IKDC-SKF short form was invariant across groups and time.

Conclusion
The 6-item, 8-item, and 9-item short form versions of the IKDC-SKF exceed contemporary fit recommendations and present as plausible alternatives to the IKDC-SKF with improved measurement properties, reduced scale response burden, and evidence of multi-group and longitudinal invariance. Further, the 6- and 8-item IKDC-SKF short forms may be used to assess group differences or change across time.
Level of evidence here

Level 3
©The Author(s)

INTRODUCTION

The International Knee Documentation Committee (IKDC) Subjective Knee Form (SKF) is a 19-item joint-specific patient-reported outcome measure (PROM) used in orthopedics and sports medicine, with applications in clinical practice and research. The IKDC-SKF is intended to be used across a variety of knee pathologies (e.g., ACL injury, meniscal injury, cartilage damage, patellofemoral pain syndrome) and has been translated into many languages (e.g., Chinese, Arabic, German). Initial assessment of scale properties indicated the English version of the IKDC-SKF had adequate test-retest reliability (ICC range = 0.85-0.99) but large ceiling effects, while translated versions have been reported to have good responsiveness (i.e., change can be detected) without floor or ceiling effects. However, other analysis procedures necessary to establish the measurement properties of the IKDC-SKF for use in clinical practice and research are either lacking (e.g., multi-group invariance testing) or have identified potential concerns with the scale.

For example, internal consistency, a measure of how similar the items are in a unidimensional scale or unique subscale, should be calculated for each construct included in a scale. High alpha levels (i.e., ≥ 0.90) have been interpreted as evidence of strong internal consistency; however, these high values may be more likely to indicate item redundancy, inclusion of too many items or parallel items, construct underrepresentation, or reduced construct precision. Low values (i.e., ≤ 0.70), in contrast, indicate poor internal consistency within a scale or construct. While alpha values ≥ 0.70 and < 0.90 are often considered acceptable, a range of ≥ 0.80 and < 0.90 has been recommended for scale development. Reported Cronbach’s alpha values for the IKDC-SKF have ranged from 0.77 to 0.97 across different versions of the scale. The reported Cronbach’s alpha values outside the recommended range, particularly those well above 0.90, along with those calculated for the entire scale, raise concerns regarding the item design, internal consistency, and dimensionality of the IKDC-SKF. The results suggest further scale modification is needed to reduce redundancy, construct underrepresentation, and response burden, while also improving model fit.

Psychometrically sound reflective scales should also have consistent structural validity, which is often established through exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) procedures or Rasch analysis. When factor analysis is used, initial procedures should follow recommended extraction techniques and factor identification should follow recommended procedures, as under factoring or over factoring issues occur in scale development. While principal component analysis (PCA) can be used initially to reduce the number of items and factors, a common factor approach is preferred, and PCA should not be used as a substitute for EFA and CFA to determine the underlying latent factors. Further, once EFA procedures have been used to identify latent factors, CFA procedures in new samples to confirm the factor structure are recommended. An important step in this process is the identification of latent factors, which is often performed using multiple criteria, such as eigenvalues (e.g., Kaiser-Guttman criterion of values greater than 1.0), scree plots, percent variance explained, or minimum-average partial correlation. Parallel analysis is another approach that has been recommended because it performs well across PCA or EFA procedures for correctly identifying factors. As it relates to the IKDC-SKF, researchers have used PCA, EFA, and Bayesian CFA procedures to establish scale structure; however, multiple factor solutions have been found and best practice recommendations have not always been followed. For example, PCA has resulted in three-component solutions, with researchers supporting a one-factor solution despite recommendations to not use PCA. Others have used EFA and Bayesian CFA methods which resulted in two-factor solutions. Thus, the factor structures (e.g., number of factors) have not been consistent across studies and the solutions have varied in the number of items to include in the final scale (e.g., 15 items across two-factors; all 18 items across the PCA solution).

Short form versions have also been identified from EFA, CFA, and Rasch analysis procedures. The first short form (i.e., 15 items) solution, however, was not identified with EFA procedures using the most contemporary methods for item retention and factor identification (e.g., parallel analysis) and CFA results have indicated further modification of the scale is warranted to identify a sound short form version for use in clinical practice and research. The need for further item removal and the identification of a parsimonious short form was also supported with Rasch analysis; however, final model solutions differed with one retaining 5 items and the other retaining 8 items. Other concerns with these studies are the use of small samples (i.e., 774 and 160, respectively) and respondent pools (e.g., healthy respondents) who are not representative of the patient population with which the scale is used in clinical practice and research. Thus, further research is needed to make clearer recommendations on a parsimonious IKDC-SKF short form that can be used in practice and research.

Finally, multi-group and longitudinal measurement invariance and hypothesis testing assessment results, which helps ensure scale suitability for use in research and clinical practice, have not been reported for the IKDC-SKF. Multi-group invariance testing should be conducted to ensure factorial stability exists across different popula-
tions, which establishes measurement properties are equivalent across various subgroups (e.g., sex, age, injury type). Establishing multi-group invariance of PROMs allows clinicians and researchers to answer substantive questions regarding group differences.\textsuperscript{20,26,36} Longitudinal invariance testing is valuable for PROMs because it helps establish if the underlying constructs are adequately measured across repeated testing to allow clinicians or researchers to interpret score changes as true change.\textsuperscript{20,25,36} Establishing multi-group and longitudinal invariance then allows for hypothesis testing by determining if the scale can be used to measure differences between relevant groups or across time.\textsuperscript{20,25,26}

Thus, further psychometric assessment of the IKDC-SKF is warranted given the reported inconsistencies, concerns with scale measurement properties, lack of invariance analysis results, and inconsistent findings on a short form version. Performing EFA and CFA procedures in large, diverse, and separate samples is valuable for determining and then confirming or refuting the structural validity of the IKDC-SKF or an identified short form version. These procedures will allow for identification of a parsimonious and psychometrically sound scale when following contemporary factor analysis procedure recommendations. Additionally, assessing the internal consistency of the identified factors (i.e., one, two, or three factors) is warranted to further confirm internal consistency and measurement precision without item redundancy. Finally, conducting multi-group and longitudinal invariance testing will provide insight into whether the scale can be used to measure group differences and change over time. Establishing these scale properties provides clinicians and researchers with a psychometrically sound scale to track patient progress or compare groups. Therefore, the purpose of the study was to assess the psychometric properties of the IKDC-SKF in a large, heterogeneous sample. This included four separate mechanisms: 1) to conduct EFA following best practice recommendations to identify a sound latent structure, which may include alternate forms (i.e., short forms), of the IKDC-SKF in a large, heterogeneous sample; 2) to assess the internal consistency of any identified constructs; 3) to use CFA procedures to confirm the structural validity of the identified scale structure in a separate sample; and 4) to perform relevant multi-group and longitudinal invariance procedures on the identified scale to inform practitioners and researchers on scale use for assessing group differences and change over time.

METHODS

A sample of patient data obtained from the Surgical Outcome System (SOS, Arthrex, Naples, Florida) was used for the study. Patients provided informed consent prior to using the SOS and were emailed PROMs at predetermined intervals. Institutional Review Board (IRB) approval for the project was granted by the Cedar-Sinai Office of Research Compliance and Quality Improvement as part of a larger research project using SOS data. University IRB was not required because the deidentified data set was not considered human subject research.

For this study, patients who were classified in an arthroscopic knee surgery group and who had completed the IKDC-SKF at baseline (i.e., pre-arthroscopic knee surgery) were included in the study. For longitudinal invariance, only patients who completed the IKDC-SKF at four time points (i.e., baseline [pre-arthroscopic knee surgery], three months post-surgery, six months post-surgery, and 12-months post-surgery) were included in the analysis.

INSTRUMENTATION

\textbf{INTERNATIONAL KNEE DOCUMENTATION COMMITTEE – SUBJECTIVE KNEE FORM}

The IKDC-SKF is a 19-item knee joint specific PROM.\textsuperscript{37} The IKDC-SKF includes one dichotomous item, four 11-point Likert scale items, and fourteen 5-point Likert scale items. Eighteen of the items are summed into one score which ranges from 0 to 100 (item #19 is not included in the score).\textsuperscript{37} A higher score represents less dysfunction, less pain, and greater knee function.\textsuperscript{37}

DATA ANALYSIS

A total of 1,959 individuals completed the IKDC-SKF prior to knee arthroscopy and were exported from the SOS database into the Statistical Package for Social Sciences (SPSS v. 25.0, Chicago, IL) and Analysis of Moment Structures (AMOS v. 25.0, Chicago, IL) for analysis. Cases with a z-score equal to or greater than ± 3.3 were classified as univariate outliers and were subsequently removed. The dataset was also assessed for multivariate outliers using Mahalanobis distance; cases with a $p < 0.001$ according to the Chi-square test were removed from the data set. Respondent data were not excluded if demographic information was missing because the primary study purpose was to assess the IKDC-SKF. Finally, histograms and descriptive statistics (i.e., skewness and kurtosis values) were used to assess the normality of the data. Following data cleaning, the data set was randomly split into two equal samples ($n_1$ and $n_2$).

\textbf{EXPLORATORY FACTOR ANALYSIS}

An EFA with maximum likelihood extraction and direct oblimin rotation was conducted on sample $n_1$ to identify a parsimonious scale. Bartlett’s test of sphericity ($< 0.001$) and Kaiser-Meyer-Olkin values ($> 0.80$) were assessed, with values outside of the specified ranges constituting a violation of the test.\textsuperscript{38} Items were assessed individually and removed one at a time until a parsimonious solution was identified.\textsuperscript{22,25} Item removal was guided by theoretical (e.g., item content), design-related (e.g., item structure)\textsuperscript{19} and statistical (e.g., low factor loadings $< 0.40$, high cross-loadings $> 0.50$, high bivariate correlations with another item, poor contribution to internal consistency) criteria.\textsuperscript{15,23,25,38} Factor retention was guided by eigenvalues $\geq 1.0$, scree plot examination, and factors that explained $\geq 5.0\%$ of the variance.\textsuperscript{25,28,58,39} Parallel analysis was used to confirm or refute factor retention; eigenvalues of the original
data set were compared to a randomly ordered data set to inform final factor retention.\textsuperscript{40} Cronbach’s alpha was also calculated for each factor retained. Items were considered for removal if the alpha value was > 0.90; item removal was guided by statistical guidelines (i.e., which item was most redundant), theory (e.g., item content), and item design. The final EFA solution resulted in a parsimonious IKDC short form to be confirmed with CFA.

\section*{CONFIRMATORY FACTOR ANALYSIS OF PROPOSED IKDC-SKF SHORT-FORM\textsuperscript{*}}

Sample \(n_2\) was used to conduct a CFA of the proposed IKDC-SKF short form to confirm model structure using maximum likelihood estimation in AMOS. Model fit indices used for evaluation included the Comparative Fit Index (CFI) > 0.95, Tucker-Lewis Index (TLI) > 0.95, and root mean square error of approximation (RMSEA) < 0.06. Models with fit indices values outside of the specified ranges indicated poor model fit and were interpreted as not supporting the proposed factor structure of the IKDC-SKF.\textsuperscript{20,41} CFA procedures also included assessing localized areas of strain and the interpretability, size, and statistical significance of the model’s parameter estimates (i.e., factor variances, covariances, and indicator errors).\textsuperscript{25} If indicated, additional items were removed, and the CFA procedures were repeated with the new model.

\section*{MULTI-GROUP INVARiance TESTING}

Multi-group invariance testing between participant sex and age groups was conducted on the full sample (i.e., samples \(n_1\) and \(n_2\) combined) in three stages: 1) structural invariance to assess equivalent factor structure between subgroups; 2) metric invariance to assess equal factor loadings between subgroups; and 3) scalar invariance to confirm equal loadings and intercepts between subgroups. Each model was more restricted than the previous model\textsuperscript{20} and each step was used to assess whether the items were being interpreted equally across selected subgroups (i.e., sex, age group). These steps ensure the meanings of the common factors are consistent across groups and that mean scores are not contaminated by outside factors (e.g., group specific attributes), which then allows for substantive questions to be answered to support hypothesis testing (e.g., comparison of subgroup means).\textsuperscript{20} If the metric model held, subgroups could be tested for equal variances on the latent constructs, and if the scalar model held, subgroups could be tested for equal latent means. For the purposes of multi-group analysis by age, participants were split into groups defined as youth (<18 years old), emerging adult (18-25 years old), early adulthood (26-40 years old), middle age (41-65 years old), and older adult (>65 years old);\textsuperscript{42} however, the older adult group was not analyzed because of its small sample size (\(n = 50\)). The \(\chi^2\text{diff}\) and CFI\textsuperscript{diff} tests were both used to assess invariance, and the scale was considered invariant at each stage if the CFI\textsuperscript{diff} was ≤ 0.01 as compared to the configural model and the fit indices previously described were met. If the model was not found to be invariant at a given step, item loadings (i.e., metric model) or item intercepts (i.e., scalar model) were released one by one, and the model was retested. Once a problematic item was identified (i.e., the one that improved CFI to be closest to the CFI of the configural model), it was removed, and the model was re-run. For the substantive questions, if the CFI\textsuperscript{diff} was > 0.01 compared to the configural model, it was deemed that the subgroups were not equal on the tested statistic (e.g., latent means). In these cases, another model was run in which one group served as the comparison group to determine relative latent variances or means for the other subgroups (i.e., greater than, less than, or equal to the comparison group). The \(\chi^2\text{diff}\) test was not weighted as heavily in the invariance process because of the effect sample size has on this statistic.\textsuperscript{20,21}

\section*{LONGITUDINAL INVARiance TESTING}

Longitudinal invariance testing was evaluated using the same procedures outlined in the multi-group invariance section to confirm similar interpretation of items and common factors across time points. If all models held (i.e., all fit indices cut-off values were met), it indicated that substantive properties (e.g., change over time) could be evaluated, allowing for clinician assessment of patient scores over time (e.g., did scores change from baseline to 12-months post-arthroscopy). The same procedures were used as described in multi-group invariance testing to identify any problematic items and create a more parsimonious scale, when indicated.

\section*{CORRELATION ANALYSES}

Bivariate correlation analysis was conducted using scores from the 18-item IKDC-SKF and scores from any generated IKDC-SKF short forms. The preferred percentage of variance explained was set at \(r > 0.90\) (\(R^2 = 0.81\)).\textsuperscript{43,44}

\section*{RESULTS}

A total of 55 cases were removed during the data cleaning process (i.e., identified outliers) leaving 1,904 cases for analysis; the 1,904 total cases were then randomly split into two even data sets (i.e., 952 cases in \(n_1\) and \(n_2\)). For the full sample, participants were 32.06 ± 14.16 years of age (range: 11-80 years) and included 874 males and 802 females. For sample \(n_1\), participants were 32.42 ± 14.38 years old (range: 11-74 years) and included 441 males and 388 females. For sample \(n_2\), participants were an average of 31.69 ± 13.93 years old (range: 12-80 years) and included 453 males and 414 females.

\section*{EXPLORATORY FACTOR ANALYSIS}

The initial exploratory factor analysis (EFA) using all 18 items resulted in a four-factor solution with items that had low loadings and high cross-loadings. Parallel analysis indicated a three-factor solution was sufficient when all 18 items were used. Items were removed during the EFA procedures one at a time and the solution was respecified until an acceptable solution was identified; a total of nine
Table 1. Exploratory factor analysis of the IKDC-SKF

<table>
<thead>
<tr>
<th>Item</th>
<th>Factor 1 Athletic Activities</th>
<th>Factor 2 Activity Level</th>
<th>Factor 3 ADLs</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>0.925</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>0.850</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.713</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>0.857</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>0.797</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>0.731</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td>0.786</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td>0.657</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td>0.647</td>
</tr>
<tr>
<td>Eigenvalue (% variance)</td>
<td>4.67 (51.87)</td>
<td>1.21 (13.44)</td>
<td>0.94 (10.43)</td>
</tr>
<tr>
<td>Cronbach's Alpha</td>
<td>0.88</td>
<td>0.84</td>
<td>0.76</td>
</tr>
</tbody>
</table>

ADLs = Activities of Daily Living.

Items were removed, resulting in a 9-item, 3-factor solution, with three items in each factor (Athletic Activities, Activity Level, Activities of Daily Living [ADLs]; Table 1). The solution accounted for 75.74% of the variance and Cronbach’s alpha values fell within the suggested range for each subscale (ADLs: 0.76; Activity Level: 0.84; Athletic Activities: 0.88; Table 1) with item loadings ranging from 0.65 to 0.95. While certain criteria (e.g., scree plot, percent variance explained) supported the 3-factor structure solution, parallel analysis with the nine items supported a two-factor structure. The 3-factor, 9-item scale was retained for CFA as further modification could be conducted during those analysis procedures to support or refute factor structure.

CONFIRMATORY FACTOR ANALYSIS 9-ITEM IKDC-SKF SHORT FORM

The CFA of the 3-factor, 9-item IKDC-SKF short form met all model fit criteria (CFI = 0.983; TLI = 0.975; IFI = 0.985; RMSEA = 0.057; chi-square = 97.667; p < .001; Figure 1) and had factor loadings ranging from 0.55 to 0.87. Construct correlations ranged from 0.64 to 0.75, with the highest correlation between Athletic Activities and ADLs (56.25% shared variance). Modification indices indicated significant cross-loadings and potential model misspecification were present.

Although model fit indices were exceeded, inspection of the model (e.g., item design, latent variable correlations, modification indices) and consideration of the parallel analysis findings led to further refinement and the identification of a 2-factor, 6-item modified IKDC-SKF short form. The 2-factor (Activity Level and ADLs), 6-item model supported by parallel analysis, demonstrated excellent model fit (CFI = 1.0; TLI = 0.999; IFI = 1.0; RMSEA = 0.11; chi-square = 8.943; p = 0.347; Figure 2), and addressed concerns (e.g., high latent variable correlations, cross-loadings) identified in the 9-item IKDC-SKF short form. Invariance testing (multigroup and longitudinal) was conducted on both the 9-item and 6-item IKDC-SKF short forms to provide further insight on both proposed factor structures.

MULTIGROUP INVARIANCE TESTING

Multigroup invariance testing across sex and age groups was conducted using participant responses to the IKDC-SKF at baseline (i.e., pre-arthroscopy).

SEX

IKDC-SKF 9-ITEM SHORT FORM

A total of 1,676 individuals (males = 874; females = 802) reported sex and were used for analysis. Both individual models (i.e., males, females) met all fit indices criteria (Table 2). The configural model fit indices also met all recommended values (CFI = 0.990; RMSEA = 0.031; Table 2). The metric and scalar models passed the CFI_diff test, warranting examination of an equal variances and equal means model. The equal variance model passed the CFI_diff test, indicating variances were equal across groups. The equal means model also passed the CFI_diff test, indicating the means were equal for all latent variables across males and females.

IKDC-SKF 6-ITEM SHORT FORM

A total of 1,676 individuals (males = 874; females = 802) reported sex and were used for analysis. Both individual models (i.e., males, females) met all fit indices criteria (Table 3). The configural model also met all recommended model fit values (CFI = 0.990; RMSEA = 0.031; Table 3). The metric and scalar models passed the CFI_diff test, warranting examination of an equal variances and equal means model. The equal variance model passed the CFI_diff test, indicating variances were equal across groups. The equal means model also passed the CFI_diff test, indicating the means were equal for all latent variables across males and females.
Table 2. Invariance across sex for the 9-item IKDC-SKF Short Form

<table>
<thead>
<tr>
<th>Measure 1 (Male)</th>
<th>Chi-Square</th>
<th>df</th>
<th>Chi-Square Diff</th>
<th>CFI</th>
<th>CFI Diff</th>
<th>TLI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>56.039</td>
<td>24</td>
<td>N/A</td>
<td>0.992</td>
<td>N/A</td>
<td>0.989</td>
<td>0.039</td>
</tr>
<tr>
<td>Measure 2 (Female)</td>
<td>68.907</td>
<td>24</td>
<td>N/A</td>
<td>0.987</td>
<td>N/A</td>
<td>0.980</td>
<td>0.048</td>
</tr>
<tr>
<td>Model A (Configural)</td>
<td>124.947</td>
<td>48</td>
<td>N/A</td>
<td>0.990</td>
<td>N/A</td>
<td>0.985</td>
<td>0.031</td>
</tr>
<tr>
<td>Model B (Metric)</td>
<td>127.682</td>
<td>54</td>
<td>2.735 (6)</td>
<td>0.990</td>
<td>NC</td>
<td>0.987</td>
<td>0.029</td>
</tr>
<tr>
<td>Model C (Equal Latent Variances)</td>
<td>136.188</td>
<td>57</td>
<td>11.241 (9)</td>
<td>0.990</td>
<td>NC</td>
<td>0.987</td>
<td>0.029</td>
</tr>
<tr>
<td>Model D (Scalar)</td>
<td>137.173</td>
<td>60</td>
<td>12.226 (12)</td>
<td>0.990</td>
<td>NC</td>
<td>0.988</td>
<td>0.028</td>
</tr>
<tr>
<td>Model E (Equal Latent Means)</td>
<td>151.039</td>
<td>63</td>
<td>26.092 (15)</td>
<td>0.989</td>
<td>0.001</td>
<td>0.987</td>
<td>0.029</td>
</tr>
</tbody>
</table>

**Bold italic font**: CFI value exceeded; df = degrees of freedom; Diff = difference; CFI = Comparative- Fit Index; TLI = Tucker-Lewis Index; RMSEA = root mean square error of approximation; N/A = not applicable; NC = no change

Table 3. Invariance across sex for the 6-item IKDC-SKF Short Form

<table>
<thead>
<tr>
<th>Measure 1 (Male)</th>
<th>Chi-Square</th>
<th>df</th>
<th>Chi-Square Diff</th>
<th>CFI</th>
<th>CFI Diff</th>
<th>TLI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>14.813</td>
<td>8</td>
<td>N/A</td>
<td>0.997</td>
<td>N/A</td>
<td>0.994</td>
<td>0.031</td>
</tr>
<tr>
<td>Measure 2 (Female)</td>
<td>10.948</td>
<td>8</td>
<td>N/A</td>
<td>0.998</td>
<td>N/A</td>
<td>0.997</td>
<td>0.021</td>
</tr>
<tr>
<td>Model A (Configural)</td>
<td>25.761</td>
<td>16</td>
<td>N/A</td>
<td>0.997</td>
<td>N/A</td>
<td>0.995</td>
<td>0.019</td>
</tr>
<tr>
<td>Model B (Metric)</td>
<td>28.506</td>
<td>20</td>
<td>2.745 (4)</td>
<td>0.998</td>
<td>+0.001</td>
<td>0.997</td>
<td>0.016</td>
</tr>
<tr>
<td>Model C (Equal Latent Variances)</td>
<td>37.002</td>
<td>22</td>
<td>11.241 (6)</td>
<td>0.996</td>
<td>0.001</td>
<td>0.995</td>
<td>0.020</td>
</tr>
<tr>
<td>Model D (Scalar)</td>
<td>32.529</td>
<td>24</td>
<td>6.768 (8)</td>
<td>0.998</td>
<td>+0.001</td>
<td>0.987</td>
<td>0.015</td>
</tr>
<tr>
<td>Model E (Equal Latent Means)</td>
<td>44.991</td>
<td>26</td>
<td>19.23 (10)</td>
<td>0.995</td>
<td>0.002</td>
<td>0.994</td>
<td>0.021</td>
</tr>
</tbody>
</table>

**Bold italic font**: CFI value exceeded; df = degrees of freedom; Diff = difference; CFI = Comparative- Fit Index; TLI = Tucker-Lewis Index; RMSEA = root mean square error of approximation; N/A = not applicable; NC = no change

AGE GROUP

**IKDC-SKF 9-ITEM SHORT FORM**

A total of 1,762 individuals (youth = 321; emerging adults = 416; early adulthood = 558; middle age = 467) who reported an age (range: 11-65 years) were used for analysis. Baseline models (i.e., youth, emerging adults, early adults, middle age) met all fit indices (Table 4). The configural model fit indices met all recommended values (CFI = 0.995; RMSEA = 0.019; Table 4). The metric and scalar models passed the CFI_diff test, warranting examination of an equal variances and equal means model. The equal variance model passed the CFI_diff test, indicating variances were equal across groups. The equal means model did not pass the CFI_diff test, indicating the means were not equal for all latent variables between age groups. When means were not constrained, the middle age group had significantly lower means than all groups (i.e., more dysfunction, more pain, and less knee ability) across all three latent variables (i.e., ADLs, Activity Level, and Athletic Activities). Additionally, the early adulthood group had a significantly lower mean (i.e., more dysfunction, more pain, and less knee ability) than the youth and emerging adult groups for the ADL latent variable. Statistically significant mean differences were not found for any latent constructs between the youth and emerging adult groups.

**IKDC-SKF 6-ITEM SHORT FORM**

A total of 1,762 individuals (youth = 321; emerging adults = 416; early adulthood = 558; middle age = 467) who reported an age (range: 11-65 years) were used for analysis. Baseline models (i.e., youth, emerging adults, early adults, middle age) met all fit indices (Table 5). The configural model fit indices met all recommended values (CFI = 0.995; RMSEA = 0.019; Table 5). The metric and scalar models passed the CFI_diff test, warranting examination of an equal variances and equal means model. The equal variance model passed the CFI_diff test, indicating variances were equal across groups. The equal means model did not pass the CFI_diff test, indicating the means were not equal for all latent variables between age groups. When means were not constrained, the middle age group had significantly lower means (i.e., more dysfunction, more pain, and less knee ability) than all groups across both latent variables (i.e., ADLs and Activity Level). Additionally, the early adulthood group had a significantly lower mean (i.e., more dysfunction, more pain, and less knee ability) than the youth and emerging adult groups for the ADL latent variable. There
Table 4. Invariance across age group for the 9-item IKDC-SKF Short Form

<table>
<thead>
<tr>
<th></th>
<th>Chi-Square</th>
<th>df</th>
<th>Chi-Square Diff</th>
<th>CFI</th>
<th>CFI Diff</th>
<th>TLI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure 1 (Youth)</td>
<td>21.997</td>
<td>24</td>
<td>N/A</td>
<td>1.00</td>
<td>N/A</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Measure 2 (Emerging Adult)</td>
<td>38.464</td>
<td>24</td>
<td>N/A</td>
<td>0.992</td>
<td>N/A</td>
<td>0.988</td>
<td>0.038</td>
</tr>
<tr>
<td>Measure 3 (Early Adult)</td>
<td>56.629</td>
<td>24</td>
<td>N/A</td>
<td>0.988</td>
<td>N/A</td>
<td>0.982</td>
<td>0.049</td>
</tr>
<tr>
<td>Measure 4 (Middle Age)</td>
<td>36.968</td>
<td>24</td>
<td>N/A</td>
<td>0.993</td>
<td>N/A</td>
<td>0.989</td>
<td>0.034</td>
</tr>
<tr>
<td>Model A (Configural)</td>
<td>154.051</td>
<td>96</td>
<td>N/A</td>
<td>0.993</td>
<td>N/A</td>
<td>0.989</td>
<td>0.019</td>
</tr>
<tr>
<td>Model B (Metric)</td>
<td>188.910</td>
<td>114</td>
<td>34.859(18)</td>
<td>0.990</td>
<td>0.003</td>
<td>0.988</td>
<td>0.019</td>
</tr>
<tr>
<td>Model C (Equal Latent Variances)</td>
<td>225.757</td>
<td>123</td>
<td>71.706(27)</td>
<td>0.987</td>
<td>0.006</td>
<td>0.987</td>
<td>0.022</td>
</tr>
<tr>
<td>Model D (Scalar)</td>
<td>236.479</td>
<td>132</td>
<td>82.428(36)</td>
<td>0.987</td>
<td>0.006</td>
<td>0.985</td>
<td>0.021</td>
</tr>
<tr>
<td>Model E (Equal Latent Means)</td>
<td>332.783</td>
<td>141</td>
<td>178.732(45)</td>
<td>0.975</td>
<td>0.022</td>
<td>0.975</td>
<td>0.028</td>
</tr>
</tbody>
</table>

**Bold italic font:** CFI value exceeded; df = degrees of freedom; Diff = difference; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = root mean square error of approximation; N/A = not applicable; NC = no change

were no significant mean differences for any latent constructs between the youth and emerging adult groups.

LONGITUDINAL INvariance testing

A total of 792 individuals completed the IKDC-SKF at all four time points and were retained for longitudinal invariance. The average age of participants in this subsample was 33.02 ± 15.00 years (range: 11-78 years; 354 females; 353 males).

IKDC-SKF 9-ITEM SHORT FORM

All baseline models (i.e., baseline, 3-months post-surgery, 6-months post-surgery, 12-months post-surgery) met all fit indices (Table 6). The configural model fit indices met all recommended values (CFI = 0.993; RMSEA = 0.019; Table 6). The metric model passed the CFI<sub>diff</sub> test, warranting examination of an equal variances model. The equal variance model did not pass the CFI<sub>diff</sub> test, indicating variances were not equal across time points for latent variables. The scalar model, however, did not pass the CFI<sub>diff</sub> test, indicating potential item-level bias which did not support testing of the equal latent means model. Follow-up analysis indicated slight item bias for item #15 (i.e., "How does your knee affect your ability to run straight ahead?").

Due to the item bias findings, invariance testing was conducted on an 8-item IKDC-SKF short form (i.e., the remaining items from the 9-item scale after item #15 was removed). All baseline models (i.e., baseline, 3-months post-surgery, 6-months post-surgery, 12-months post-surgery) met model fit indices (Table 7). The configural model fit indices met all recommended values (CFI = 0.997; RMSEA = 0.014; Table 7). The metric and scalar models passed the CFI<sub>diff</sub> test, warranting examination of an equal variances and equal means model. The equal variance model did not pass the CFI<sub>diff</sub> test, indicating variances were not equal across time points for latent variables. The equal means model also did not pass the CFI<sub>diff</sub> test, indicating means were not equal across time. When not constrained to be equal, Activity Level, ADLs, and Athletic Activities latent means at 3-, 6-, and 12-months post-surgery were sigif-
significantly higher than baseline (i.e., pre-arthroscopy) scores (i.e., less dysfunction, less pain, and higher knee ability), except for Activity Level latent means at three months. Scores increased/improved across time, except for Activity Level at three months, indicating patients reported scores with improved function, pain, and knee ability after surgery.

**IKDC-SKF 6-ITEM SHORT FORM**

All baseline models (i.e., baseline, three months post-surgery, six months post-surgery, 12-months post-surgery) met all fit indices (Table 8). The configural model fit indices met all recommended values (CFI = 0.998; RMSEA = 0.012). The metric model and scalar model passed the CFI_{diff} test, warranting examination of an equal variances and equal means model. The equal variance model did not pass the CFI_{diff} test, indicating variances were not equal across time points. The equal means model also did not pass the CFI_{diff} test, indicating means were significantly different across time points. When not constrained, Activity Level and ADL latent means at three-, six-, and 12-months post-surgery were significantly higher than baseline (i.e., pre-arthroscopy) scores (i.e., less dysfunction, less pain, and higher knee ability), except for Activity Level at six months post-surgery. Scores increased/improved across time, except for Activity Level at three months, indicating patients reported improved scores with function, pain, and knee ability after surgery.

**CORRELATION ANALYSIS**

Individual scores for the IKDC-SKF 9-item short form were highly correlated ($r = 0.924$, $R^2 = 0.854$) with the scores for the original 18-item IKDC-SKF. Individual scores for the IKDC-SKF 6-item short form were highly correlated ($r = 0.889$, $R^2 = 0.790$) with the scores for the original 18-item IKDC-SKF. Scores for the IKDC-SKF 9-item short form were also highly correlated ($r = 0.940$, $R^2 = 0.884$) with the scores for the 6-item IKDC-SKF short form. Finally, scores on the modified 8-item (3-dimension) IKDC-SKF short form were highly correlated with scores on the original 18-item IKDC-SKF ($r = 0.919$, $R^2 = 0.845$), the 9-item IKDC-SKF short form ($r = 0.992$, $R^2 = 0.984$), and the 6-item IKDC-SKF short form ($r = 0.962$, $R^2 = 0.925$).

**Table 6. Longitudinal Invariance for the 9-item IKDC-SKF Short Form**

<table>
<thead>
<tr>
<th></th>
<th>Chi-Square</th>
<th>df</th>
<th>Chi-Square Diff</th>
<th>CFI</th>
<th>CFI Diff</th>
<th>TLI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure 1 (Pre)</td>
<td>53.313</td>
<td>24</td>
<td>N/A</td>
<td>0.990</td>
<td>N/A</td>
<td>0.985</td>
<td>0.043</td>
</tr>
<tr>
<td>Measure 2 (3 months)</td>
<td>52.882</td>
<td>24</td>
<td>N/A</td>
<td>0.992</td>
<td>N/A</td>
<td>0.987</td>
<td>0.039</td>
</tr>
<tr>
<td>Measure 3 (6 months)</td>
<td>37.748</td>
<td>24</td>
<td>N/A</td>
<td>0.997</td>
<td>N/A</td>
<td>0.995</td>
<td>0.027</td>
</tr>
<tr>
<td>Measure 4 (12 months)</td>
<td>65.321</td>
<td>24</td>
<td>N/A</td>
<td>0.992</td>
<td>N/A</td>
<td>0.989</td>
<td>0.047</td>
</tr>
<tr>
<td>Model A (Configural)</td>
<td>614.743</td>
<td>474</td>
<td>N/A</td>
<td>0.993</td>
<td>N/A</td>
<td>0.990</td>
<td>0.019</td>
</tr>
<tr>
<td>Model B (Metric)</td>
<td>740.369</td>
<td>492</td>
<td>125.626 (18)</td>
<td>0.987</td>
<td>0.006</td>
<td>0.984</td>
<td>0.025</td>
</tr>
<tr>
<td>Model C (Equal Latent Variances)</td>
<td>1108.942</td>
<td>501</td>
<td>494.199 (27)</td>
<td>0.969</td>
<td>0.024</td>
<td>0.961</td>
<td>0.039</td>
</tr>
<tr>
<td>Model D (Scalar)</td>
<td>908.477</td>
<td>510</td>
<td>293.734 (36)</td>
<td>0.980</td>
<td>0.013</td>
<td>0.975</td>
<td>0.031</td>
</tr>
</tbody>
</table>

**Table 7. Longitudinal Invariance for the 8-item IKDC-SKF Short Form**

<table>
<thead>
<tr>
<th></th>
<th>Chi-Square</th>
<th>df</th>
<th>Chi-Square Diff</th>
<th>CFI</th>
<th>CFI Diff</th>
<th>TLI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure 1 (Pre)</td>
<td>24.382</td>
<td>17</td>
<td>N/A</td>
<td>0.997</td>
<td>N/A</td>
<td>0.995</td>
<td>0.023</td>
</tr>
<tr>
<td>Measure 2 (3 months)</td>
<td>27.127</td>
<td>17</td>
<td>N/A</td>
<td>0.996</td>
<td>N/A</td>
<td>0.994</td>
<td>0.027</td>
</tr>
<tr>
<td>Measure 3 (6 months)</td>
<td>25.566</td>
<td>17</td>
<td>N/A</td>
<td>0.997</td>
<td>N/A</td>
<td>0.996</td>
<td>0.025</td>
</tr>
<tr>
<td>Measure 4 (12 months)</td>
<td>28.512</td>
<td>17</td>
<td>N/A</td>
<td>0.997</td>
<td>N/A</td>
<td>0.996</td>
<td>0.029</td>
</tr>
<tr>
<td>Model A (Configural)</td>
<td>402.353</td>
<td>350</td>
<td>N/A</td>
<td>0.997</td>
<td>N/A</td>
<td>0.995</td>
<td>0.014</td>
</tr>
<tr>
<td>Model B (Metric)</td>
<td>473.321</td>
<td>365</td>
<td>70.968 (15)</td>
<td>0.993</td>
<td>0.004</td>
<td>0.991</td>
<td>0.019</td>
</tr>
<tr>
<td>Model C (Equal Latent Variances)</td>
<td>811.086</td>
<td>374</td>
<td>408.733 (24)</td>
<td>0.973</td>
<td>0.024</td>
<td>0.964</td>
<td>0.038</td>
</tr>
<tr>
<td>Model D (Scalar)</td>
<td>564.617</td>
<td>380</td>
<td>162.264 (30)</td>
<td>0.989</td>
<td>0.008</td>
<td>0.975</td>
<td>0.031</td>
</tr>
<tr>
<td>Model E (Equal Latent Means)</td>
<td>1505.912</td>
<td>386</td>
<td>1103.559 (36)</td>
<td>0.930</td>
<td>0.067</td>
<td>0.910</td>
<td>0.061</td>
</tr>
</tbody>
</table>

**Bold italic font:** CFI value exceeded; df = degrees of freedom; Diff = difference; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = root mean square error of approximation; N/A = not applicable; NC = no change
Table 8. Longitudinal Invariance for the 6-item IKDC-SKF Short Form

<table>
<thead>
<tr>
<th>Measure</th>
<th>Chi-Square</th>
<th>df</th>
<th>Chi-Square Diff</th>
<th>CFI</th>
<th>CFI Diff</th>
<th>TLI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure 1 (Pre)</td>
<td>15.254</td>
<td>8</td>
<td>N/A</td>
<td>0.995</td>
<td>N/A</td>
<td>0.992</td>
<td>0.034</td>
</tr>
<tr>
<td>Measure 2 (3 months)</td>
<td>15.961</td>
<td>8</td>
<td>N/A</td>
<td>0.996</td>
<td>N/A</td>
<td>0.992</td>
<td>0.035</td>
</tr>
<tr>
<td>Measure 3 (6 months)</td>
<td>6.284</td>
<td>8</td>
<td>N/A</td>
<td>1.000</td>
<td>N/A</td>
<td>1.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Measure 4 (12 months)</td>
<td>12.202</td>
<td>8</td>
<td>N/A</td>
<td>0.999</td>
<td>N/A</td>
<td>0.997</td>
<td>0.026</td>
</tr>
<tr>
<td>Model A (Configural)</td>
<td>210.782</td>
<td>188</td>
<td>N/A</td>
<td>0.998</td>
<td>N/A</td>
<td>0.997</td>
<td>0.012</td>
</tr>
<tr>
<td>Model B (Metric)</td>
<td>264.528</td>
<td>200</td>
<td>53.746(12)</td>
<td>0.994</td>
<td>0.004</td>
<td>0.992</td>
<td>0.020</td>
</tr>
<tr>
<td>Model C (Equal Variances)</td>
<td>554.915</td>
<td>206</td>
<td>344.133(18)</td>
<td>0.969</td>
<td>0.029</td>
<td>0.958</td>
<td>0.046</td>
</tr>
<tr>
<td>Model D (Scalar)</td>
<td>344.718</td>
<td>212</td>
<td>133.936(24)</td>
<td>0.988</td>
<td>0.01</td>
<td>0.984</td>
<td>0.028</td>
</tr>
<tr>
<td>Model E (Equal Latent Means)</td>
<td>1214.108</td>
<td>218</td>
<td>1003.326(30)</td>
<td>0.911</td>
<td>0.087</td>
<td>0.887</td>
<td>0.076</td>
</tr>
</tbody>
</table>

*Bold italic font*: CFI value exceeded; df = degrees of freedom; Diff = difference; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = root mean square error of approximation; N/A = not applicable; NC = no change

DISCUSSION

Best practice recommendations for assessing the structural validity have not always been followed or reported in measurement studies of the IKDC-SKF,\(^{20,22-31}\) which may explain the inconsistent structural findings reported.\(^{1,32,33,37}\) Further, multiple short form versions of the IKDC-SKF have been suggested in the literature, but initial efforts have primarily used small samples that do not well represent the patient population who completes the IKDC-SKF.\(^{12,34}\) Therefore, assessment of the IKDC-SKF using recommended classical test theory procedures was warranted and the purpose of our study was to conduct EFA, CFA, and invariance testing procedures on the IKDC-SKF in a large, heterogeneous sample of patients to assess the measurement properties of the scale or an alternate, psychometrically sound short form version of the scale. EFA resulted in a 9-item, 3-factor IKDC-SKF short form (IKDC-SKF-9; Appendix 1) supported by CFA and multi-group invariance testing; however, the proposed model did not meet all recommended measurement criteria and did not pass longitudinal invariance requirements. Due to potential concerns with the identified 9-item version, subsequent 8-item (3-factor; IKDC-SKF-8; Appendix 2) and 6-item (2-factor; IKDC-SKF-6; Appendix 3) IKDC-SKF short forms (Appendices 1 and 2) were also tested with CFA and invariance procedures.

FACTOR STRUCTURE

The IKDC-SKF has been reported to have a unidimensional\(^{32,37}\) and a multidimensional\(^{1,8,33}\) factor structure with different items in the final models. Study methodology differences may contribute to the inconsistent findings as differences in samples (e.g., size, respondent population [e.g., healthy,\(^{12}\) ACL injury,\(^{1,34}\)) and analysis methodologies (e.g., EFA/CFAs,\(^{33}\) PCA,\(^{32}\) Bayesian SEM,\(^{1}\) Rasch,\(^{12,34}\) factor and item retention criteria) exist between studies. For example, our study included a large, heterogenous sample of patients who had undergone arthroscopic knee surgeries, while others have included smaller samples, healthy respondents,\(^{12}\) or focused on different patient populations (e.g., ACL reconstruction,\(^{1,34}\) meniscal lesions,\(^{45}\) various patient pathologies\(^{32,33,35}\)). Research\(^{1,8,33}\) using more contemporary and recommended factor analysis measurement techniques has generally supported a multidimensional structure; however, evidence exists to suggest structural validity and model fit could be improved with further item reduction.\(^{8,12,34}\)

Three-dimensional and two-dimensional factor structures that exceeded most recommended contemporary fit criteria were identified.\(^{20,41}\) The retained solutions included fewer items than those found in prior research on the IKDC-SKF.\(^{1,8,33}\) Removing items with poor fit (e.g., cross-loadings, item redundancy) or poor design (e.g., item structure, item reading level, etc.) can improve internal consistency and scale structure, while also reducing response burden with a more concise instrument.\(^{20,38,43}\) The proposed short form versions improved model precision and scale structure without losing much of the information captured with all 18 items. Despite using nine or fewer items, scores on the short form versions accounted for 79% (r = .889), 84% (r = .919), and 85% (r = .924) of the variance in scores on the 18-item original IKDC-SKF with the 6-, 8-, and 9-items, respectively. Our correlational findings are in line with prior research using similar procedures to produce short-form versions of previously established PROMs\(^{43,44}\) and indicate the newly proposed models capture similar enough information to warrant use in comparison to the original scale. One concern, however, was the identification of an internal consistency value (0.76) for the ADLs construct that is outside of the preferred 0.80 to 0.90 range for scale development.\(^{15,17}\) Thus, the ADLs construct may not have the preferred precision for assessing the construct; however, scale design recommendations caution against using constructs with fewer than three items.\(^{20,21,46}\) Future work may be recommended to alter or add items to measure the ADLs construct more precisely. Until that time, researchers and clinicians should be aware that this construct
does not meet the strictest contemporary recommendations for internal consistency.

**MULTI-GROUP INVARIANCE TESTING**

This study is the first to perform multi-group invariance testing with the IKDC-SKF. Multi-group invariance testing helps to ensure the association between the items and dimensions are stable between groups, which supports scale validity and allows for an instrument to be used to assess group differences (e.g., group mean differences in older individuals compared to younger individuals would be outside scale measurement error). Both the 9-item and 6-item IKDC-SKF short form versions in this study were found to be invariant across configural, metric, and scalar models for sex and age groups, indicating the short form models have sound measurement properties across the tested groups. Thus, researchers and clinicians could use these versions of the scale to assess differences among these groups. The findings also allow for substantive testing of whether variances or means are equal between groups, which can also support scale validity.

The current multi-group findings could provide theoretical support for the validity of the two short form versions if the findings align with expectations based on the literature. For example, widespread pain, which often includes longstanding knee pain, is more common in patients over 50 years of age, and self-reported knee pain has been found to be higher in people over the age of 40 compared to those under the age of 40. Additionally, OA, a leading diagnosis and cause of disability in older populations, has a higher prevalence and more radiographic signs with increases in age and population longevity, with those over the age of 45 accounting for over 98% of total knee arthroplasties. Further, the presence of all types of knee abnormalities (e.g., osteophytes, cartilage damage, ligamentous damage, OA) has been found to increase with age, and it has been reported that 85% or more of patients 50 years of age or older demonstrate articular cartilage changes to at least one knee compartment compared to 52% of patients between the ages of 20 to 29 years of age and 15% of patients 20 years or younger. Researchers have also indicated knee functional difficulties increase with age.

Gradual functional deterioration was found across the lifespan on the Knee Injury and Osteoarthritis Outcome Score (KOOS) and reported functional impairment was more apparent with functional tasks of greater difficulty (e.g., sport and recreational functional activities) across adults aged 18-84. Similarly, Baldwin et al. reported group mean score differences across age groups (e.g., 18 to 29 years, 30 to 39 years, 40 to 49 years, 50 to 59 years), with consistent findings of less knee impairment on the pain, ADL, sport/recreation, and quality of life constructs of the KOOS for those under 40 years of age. Thus, if the proposed IKDC-SKF short forms are measuring the intended constructs, it would be expected to find similar patterns in our multi-group invariance results (e.g., higher levels of impairment in the older age groups in our sample).

The proposed models were able to identify age group differences at initial examination: statistically significant group differences were found across age groups for the 6-item and 9-item solutions. Specifically, the middle age (41-65 years of age) group had lower means (i.e., more impaired knee health) than the other younger age groups (i.e., early adult, emerging adult, youth) at baseline (i.e., prearthroscopic surgery). Additionally, the early adulthood (26-40 years of age) group had a lower mean score for the ADL construct than the youth and emerging adult groups. The middle age group reporting greater impairment across all three factors (i.e., ADLs, Activity Level, Athletic Activity) than the three younger age groups align with expectations based on KOOS findings and expectations for functional impairment across the life span. Similarly, the early adult age group reporting greater impairment in the Activity Level factor but similar mean scores for ADLs with younger age groups also aligns with expectations based on the literature. The lack of a statistically significant difference for the Athletic Activity construct might be explained by the final items included in that construct; however, prior KOOS Sport/Recreation construct findings indicate meaningful age group differences were not found until after 40 years of age for items assessing this type of construct. Finally, statistically significant differences were not found between the emerging adult and youth groups, which aligns with prior KOOS findings and expectations for the presence of knee pathological changes being less likely to have occurred this early in the lifespan.

A weakness of our results, however, is the low number of responses (n = 30) in the older adult (66 years of age or older) group, which prevented us from including this group in the multi-group analyses of the 6- and 9-item short form versions. It would be valuable to confirm that similar group differences are found in older or elderly populations. Similarly, it would be valuable to conduct this analysis across different pathology groups (e.g., total knee arthroplasty patients vs. arthroscopy patients) to ensure the scale has the necessary measurement properties to assess groups differences based on pathology and if greater levels of diagnosed pathology results in greater reported knee health impairment on the proposed short forms.

**LONGITUDINAL INVARIANCE TESTING**

To the authors knowledge, this study is also the first to assess longitudinal invariance of the IKDC-SKF or proposed short forms. Longitudinal invariance testing is valuable because it allows for the determination of whether the items and dimensions are stable across time, which supports scale validity and allows for an instrument to be used to assess change over time. We found that the 6-item IKDC-SKF short form was invariant across time based on the configural, metric, and scalar model findings. The 9-item IKDC-SKF short form was not invariant across time, and further analysis revealed item #15 exhibited bias. Follow-up analysis indicated that the remaining items (i.e., 9-item IKDC-SKF short form except for item #15; 8-item IKDC-SKF short form) and factor structure were invariant across time. The findings allow for the assessment of score change over time to determine when and where patient reported improvement occurred following arthroscopic surgery.
surgery on the 6- and 8-item short form versions. Finding expected improvement over time would support scale validity, while also indicating whether patients perceived improvements in their condition across time.20,21,25

In this study, individuals reported the lowest scores (i.e., greatest impairment in knee health) at baseline and the highest scores (i.e., lowest knee impairment) at 12-months post-surgery. The score improvements were statistically significant for all latent means (e.g., Activity Level, ADLs) at 3-, 6-, and 12-months post-surgery except Activity Level at three months post-surgery across the 6- and 8-item versions of the scale. The current findings indicate patients reported statistically significant improvements across all dimensions six months post-surgery and the improvements were maintained at 12-months post-surgery. The longitudinal findings were consistent across both the 6- and 8-item short form versions of the IKDC-SKF. Thus, the two versions of the scale identified patient-reported improvement across the measured latent constructs across time similarly.

The current findings support scale validity as the results are consistent with what we would expect for individuals recovering from surgery. Specifically, patients in the rehabilitation process would be expected to report improvements across items intended to measure how the prior injury impaired the previously measured constructs (i.e., Activity Levels, ADLs, and Athletic Activities) because the patient should experience health status improvement (e.g., decreased pain, increased ROM, increased strength) after surgery through a combination of treatment effectiveness, natural healing, and placebo. We would also expect to find that improvements in certain constructs (e.g., Activity Level) might not occur as quickly as other constructs (e.g., ADLs) because patients may have activity/rehabilitation restrictions or more substantial pathology that may slow improvements in specific constructs; however, the authors would also expect to then see significant improvements in those dimensions at later time points that are in line with the improvements found across the other constructs over time.

Thus, the current findings support the use of the 6- and 8-item IKDC-SKF short-form versions: sound measurement properties were demonstrated and theoretical support (e.g., patient-perceived improvements match expectations for the recovery process of the included patients in our study) was found. These results also provide support for clinicians who want to use the short form versions of the IKDC-SKF (i.e., 6- and 8-item versions) to measure change across time. The 9-item short form could be used with caution to assess change across time because it did not meet the strictest criterion for longitudinal invariance due to one problematic item. However, it is also important to note limitations with the 3-dimensional solutions: 1) parallel analysis better supported a 2-dimensional factor structure once problematic items had been removed; and 2) the Athletic Activities factor in the 8-item IKDC-SKF short form only contained two items and three to five items per factor has been recommended.20,46,60 Clinicians and researchers should consider summary of findings when deciding which version of the IKDC-SKF to use within their clinical practice or research; however, the 6-item IKDC-SKF likely has the greatest measurement support for its use across various research and clinical practice scenarios. Further scale development work is needed to develop items to accurately capture the desired information of the Athletic Activities factor and truly support a 3-dimensional IKDC-SKF factor structure.

LIMITATIONS AND FUTURE RESEARCH

While the current study included the use of contemporary analysis procedures on a large, diverse sample of patients, it does have limitations. First, the data set did not include information on the type of knee pathology or procedure performed. One of the preconditions for a viable IKDC-SKF instrument is that the model is stable over a variety of knee pathologies. Without the relevant demographic information, we were unable to conduct multi-group invariance tests by pathology or intervention type. Additionally, our sample had a small sample of patients classified in the older age (66 years or older) group, which prevented their inclusion in the multi-group age analysis. Further, responses to all 18-items were used to produce short-form versions; while the analysis processes used are common for instrument refinement, it is possible that respondents were influenced by items not included in the final models. Additional psychometric analyses could also be conducted; for example, the new models could be tested against a criterion standard scale to support validity, Rasch analysis (e.g., person differentiation) could be performed, and responsiveness (e.g., minimal clinically important difference [MCID] values) and test-retest reliability of the new models could be assessed. Finally, as this PROM was delivered via email, potential response biases could have affected results and it was not possible to examine if completion mode (i.e., paper or electronic) influenced results.

Future analysis should include multi-group invariance testing in older populations, while also examining the multi-group invariance properties across pathology or intervention groups. Further, researchers should examine the structural validity of the scale in different respondent groups who only answer the short form versions of the scale, while also incorporating additional items to measure the Athletic Activities factor more effectively. In addition to confirming the measurement properties of the short form versions, these analyses could provide insight into whether the 6-item short form may have other psychometric concerns (e.g., ceiling effects) when used in certain populations (e.g., competitive athletes) that could be resolved by developing an effective 3-dimensional scale. Finally, future research should also work to establish the test-retest reliability, responsiveness (e.g., MCIDs), and criterion validity of the short form versions.

CONCLUSION

The EFA and CFA resulted in short form versions of the IKDC-SKF that exceed contemporary fit recommendations. The identified models present as plausible alternatives to
the IKDC-SKF as the original item pool was reduced by more than 50%, but the short forms still accounted for most of the variance in participant responses on the IKDC-SKF. Further, the 6- and 8-item IKDC-SKF short forms met all criteria for applied multi-group and longitudinal invariance tests, which indicates the scales may be used to assess group differences or change across time. The overall analysis indicated the short form versions of the IKDC-SKF were structurally valid alternatives to the IKDC-SKF with improved measurement properties, reduced scale response burden, and evidence to support the assessment of patient improvement across time.

FINANCIAL DISCLOSURE
This publication was supported by an Institutional Development Award (IDeA) from the National Institute of General Medical Sciences of the National Institutes of Health under Grant #P20GM103408 and an Idaho WWAMI Research Training Support Award.

CONFLICT OF INTEREST
The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Submitted: March 10, 2023 CDT, Accepted: June 16, 2023 CDT
REFERENCES


International Journal of Sports Physical Therapy
APPENDICES

APPENDIX 1: INTERNATIONAL KNEE DOCUMENTATION COMMITTEE – SUBJECTIVE KNEE FORM: SHORT-FORM 9

INTERNATIONAL KNEE DOCUMENTATION COMMITTEE – SUBJECTIVE KNEE FORM

SHORT-FORM 9

1. What is the highest level of activity that you can perform without significant knee pain?
   - □ Very strenuous activities like jumping or pivoting as in basketball or soccer
   - □ Strenuous activities like heavy physical work, skiing or tennis
   - □ Moderate activities like moderate physical work, running or jogging
   - □ Light activities like walking, housework or yard work
   - □ Unable to perform any of the above activities due to knee pain

2. What is the highest level of activity you can perform without significant swelling in your knee?
   - □ Very strenuous activities like jumping or pivoting as in basketball or soccer
   - □ Strenuous activities like heavy physical work, skiing or tennis
   - □ Moderate activities like moderate physical work, running or jogging
   - □ Light activities like walking, housework, or yard work
   - □ Unable to perform any of the above activities due to knee swelling

3. What is the highest level of activity you can participate in on a regular basis?
   - □ Very strenuous activities like jumping or pivoting as in basketball or soccer
   - □ Strenuous activities like heavy physical work, skiing or tennis
   - □ Moderate activities like moderate physical work, running or jogging
   - □ Light activities like walking, housework or yard work
   - □ Unable to perform any of the above activities due to knee

4. How does your knee affect your ability to:

<table>
<thead>
<tr>
<th>Activity</th>
<th>Not difficult at all</th>
<th>Minimally difficult</th>
<th>Moderately difficult</th>
<th>Extremely difficult</th>
<th>Unable to do</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Go up stairs</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>b. Squat</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>c. Sit with your knee bent</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>d. Run straight ahead</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>e. Jump and land on your involved leg</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>f. Stop and start quickly</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

APPENDIX 2: INTERNATIONAL KNEE DOCUMENTATION COMMITTEE – SUBJECTIVE KNEE FORM: SHORT-FORM 8

INTERNATIONAL KNEE DOCUMENTATION COMMITTEE – SUBJECTIVE KNEE FORM

SHORT-FORM 8

1. What is the highest level of activity that you can perform without significant knee pain?
   - □ Very strenuous activities like jumping or pivoting as in basketball or soccer
   - □ Strenuous activities like heavy physical work, skiing or tennis
   - □ Moderate activities like moderate physical work, running or jogging
   - □ Light activities like walking, housework or yard work
   - □ Unable to perform any of the above activities due to knee pain

2. What is the highest level of activity you can perform without significant swelling in your knee?
   - □ Very strenuous activities like jumping or pivoting as in basketball or soccer
   - □ Strenuous activities like heavy physical work, skiing or tennis
   - □ Moderate activities like moderate physical work, running or jogging
   - □ Light activities like walking, housework or yard work
   - □ Unable to perform any of the above activities due to knee swelling

3. What is the highest level of activity you can participate in on a regular basis?
   - □ Very strenuous activities like jumping or pivoting as in basketball or soccer
   - □ Strenuous activities like heavy physical work, skiing or tennis
   - □ Moderate activities like moderate physical work, running or jogging

International Journal of Sports Physical Therapy
1 □ Light activities like walking, housework or yard work
0 □ Unable to perform any of the above activities due to knee

4. How does your knee affect your ability to:

<table>
<thead>
<tr>
<th>Activity</th>
<th>Not difficult at all</th>
<th>Minimally difficult</th>
<th>Moderately difficult</th>
<th>Extremely difficult</th>
<th>Unable to do</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Go up stairs</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>b. Squat</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>c. Sit with your knee bent</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>d. Jump and land on your involved leg</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>e. Stop and start quickly</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

APPENDIX 3: INTERNATIONAL KNEE DOCUMENTATION COMMITTEE – SUBJECTIVE KNEE FORM: SHORT-FORM

INTERNATIONAL KNEE DOCUMENTATION COMMITTEE – SUBJECTIVE KNEE FORM

SHORT-FORM 6

3. What is the highest level of activity you can participate in on a regular basis?

4. How does your knee affect your ability to:

<table>
<thead>
<tr>
<th>Activity</th>
<th>Not difficult at all</th>
<th>Minimally difficult</th>
<th>Moderately difficult</th>
<th>Extremely difficult</th>
<th>Unable to do</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Go up stairs</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>b. Squat</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>c. Sit with your knee bent</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Novice Inter-Rater Reliability on the Selective Functional Movement Assessment (SFMA) After a 4-Hour Training Session

Brent Harper¹, Adrian Aron²

¹ Physical Therapy, Chapman University, ² Physical Therapy, Radford University

Keywords: interrater reliability, movement, education, compensation patterns, students, teaching

International Journal of Sports Physical Therapy

Background
The Selective Functional Movement Assessment (SFMA) is a whole-body movement classification system that identifies non-optimal movement performance requiring further assessment. There needs to be more evidence specifying the training time required to obtain SFMA reliability for entry-level health care practitioners.

Purpose
The primary intent of this study was to determine SFMA inter-rater reliability between two third-year physical therapy students following an in-person three-hour training and one-hour follow-up training with a certified SFMA physical therapist. The secondary purpose was to compare rater scores of the composite criterion 50-point checklist and rater categorization using the top-tier movements in real-time assessments of healthy participants.

Study Design
Inter-rater reliability study.

Methods
Two novice raters received training on assessing movement using the SFMA. Participants included non-pregnant healthy adults screened for general exercise, participants were excluded for history of orthopedic surgery within the prior six months. Three independent raters, including two novice and one SFMA-certified rater, individually assessed the top-tier movements in separate rooms in real-time. Participants were randomly assigned a start location, and raters were blinded to each other’s criterion 50-point checklist and categorical scoring. Statistical analysis included a paired t-test, a repeated measures ANOVA, and a two-way, mixed absolute agreement ICC.

Results
Twenty-five participants (23.4 years ± 1.9; 72% female) completed the SFMA top-tier movements. Significant differences were identified with novice raters identifying fewer non-optimal movement patterns than the certified clinician. The intraclass correlation coefficient (ICC2,1) was moderate (0.60, p<0.001) for all three raters on the 50-point criterion checklist scoring.

Corresponding Author:
Brent A. Harper
Chapman University, Department of Physical Therapy,
9401 Jeronimo Road, Irvine, CA, 92618, USA
E-mail: brharper@chapman.edu
Phone: 714-516-5946
Fax: 949-206-0012
ORCID: 0000-0002-4705-5653
Conclusion

Third year physical therapy students were able to demonstrate moderate inter-rater reliability assessing healthy individuals using the 50-point criterion checklist. Variation between novice raters may reflect the amount of previous exposure assessing movement and suggests that some may require more time learning and practicing in order to identify non-optimal movement patterns that may require further assessment.

Level of Evidence

3b

©The Author(s)

INTRODUCTION

Musculoskeletal screening and movement-based assessment tests are used in physical therapy and sports medicine to identify injury risk, movement dysfunction, and sources of pain.\(^1,2\) Screening and assessment tests incorporating whole body kinetic chain movements may reveal underlying impairments contributing to a person’s chief complaint of pain in regions elsewhere in the body.\(^3\) Often, dysfunctional movement patterns and impaired motor control are exhibited in response to pain or limited joint motion.\(^3\) When aberrant movement patterns are present, other systems in the body compensate to complete functional movement patterns of daily life.\(^4-6\) These compensated functional movements are considered dysfunctions.\(^2-5\)

Regional interdependence (RI) is the theoretical construct which proposes that adjacent anatomical areas may contribute to or be the primary source of musculoskeletal symptoms. Thus, regardless of proximity to the anatomical site of symptoms, non-symptomatic dysfunction(s) in various body regions within the kinetic chain may directly or indirectly influence the clinical presentation.\(^1,2,7-9\) The Selective Functional Movement Assessment (SFMA) is based on RI.\(^10,11\) The pathoanatomical model may lead to misdiagnoses for the source of pain; in contrast, RI provides a more comprehensive approach to identify and assess multiple relevant body regions for their possible role in the clinical presentation.\(^9-11\) From the RI model, health care providers can analyze movements and postures and identify the specific movements that can lead to abnormal loads or abnormal pressure and adaptive changes resulting in changed kinematics.\(^12\)

CLASSIFICATION SYSTEMS

The SFMA is a musculoskeletal screening and diagnostic classification tool in which top-tier movements are scored at two levels. First, the composite criterion checklist is scored out of 50, then movements are assigned to a categorical identifier, filtering the movement into one of four categories.\(^1,2,7,8\) The SFMA begins with a foundational movement evaluation to identify impairments and limitations within tri-planar movements and to determine whether or not those movements provoke patient symptoms.\(^1,2,11\) Tri-planar movements are scored using a composite checklist and then each pattern is categorized into one of four categories: functional non-painful (FN), functional painful (FP), dysfunctional non-painful (DN), or dysfunctional painful (DP). Movements scored as DN are further assessed for mobility or motor control impairments contributing to the original site of pain and dysfunction.\(^1,2,7,8,11\)

Understanding the inter-rater reliability and validity of screening and diagnostic tools is an essential factor to consider before clinical utilization.\(^13,14\) The inter-rater reliability of the SFMA is not established for healthcare practitioners in training and with real-time assessment (a researcher physically evaluating the participant, one-on-one); thus, the substantiality of the test is still unknown.\(^15\)

To date, three studies\(^1,7,8\) have evaluated the reliability of the top-tier SFMA. In 2014, Glaws et al.\(^1\) performed top-tier SFMA inter-rater and intra-rater video assessments on healthy subjects using three SFMA-certified raters with varying levels of SFMA clinical application. In 2017, Dolbeer et al.\(^8\) assessed inter-rater reliability on subjects with pain with two raters assessing in real-time and one rater scoring by watching a video recording; however, all raters were SFMA-certified with over 400 hours of clinical application using this method. In 2019, Stanek et al.\(^7\) evaluated inter-rater and intra-rater reliability on healthy subjects using three raters in real-time comparing two SFMA-certified raters with varying levels of SFMA clinical application while one rater was a student with no formal SFMA top-tier training but who did have a summer clinical rotation utilizing this approach.

SIGNIFICANCE OF THE PROBLEM

Although prior research\(^1,7,8\) has demonstrated good reliability, healthcare practitioners in training have not been thoroughly studied, nor has the standard training time required to develop reliable movement pattern assessment and recognition. When two researchers are scoring the same participant simultaneously, only one can give the instruction to the patient, which, in itself, limits the inter-rater reliability results.\(^15\) Assessing movement live involves different viewing positions than a video assessment. Furthermore, watching video performance removes the real world setting in which clinical application occurs. To avoid these limitations, researchers need to individually perform the SFMA. The primary intent of this study was to determine SFMA inter-rater reliability between two third-year physical therapy students following an in-person three-hour training and one-hour follow-up training with a certified SFMA physical therapist, and the secondary purpose was to compare rater scores of the composite criterion 50-point checklist and rater categorization using the top-

International Journal of Sports Physical Therapy
Table 1. Subjects’ Descriptive Statistics

<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th>Male</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Participants</td>
<td>18</td>
<td>7</td>
<td>25</td>
</tr>
<tr>
<td>Age (years)</td>
<td>23.2 ± 8</td>
<td>24.0 ± 7</td>
<td>23.4 ± 1.9</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>164.7 ± 1.4</td>
<td>175.2 ± 2.5</td>
<td>167.6 ± 7.6</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>65.1 ± 2.9</td>
<td>78.5 ± 1.8</td>
<td>68.9 ± 12.3</td>
</tr>
<tr>
<td>BMI</td>
<td>23.9 ± 1.0</td>
<td>25.4 ± 7.1</td>
<td>24.4 ± 4.0</td>
</tr>
</tbody>
</table>

Note. Abbreviations: cm, centimeters; Kg, kilograms; BMI, Body Mass Index; Values are presented as Mean ± SD

tier movements in real-time assessments of healthy participants.

METHODS

PARTICIPANTS

A convenience sample of 29 healthy volunteers were recruited. Participants were excluded if they had any positive marks on the Physical Activity Readiness Questionnaire (PAR-Q) health assessment, had undergone orthopedic surgery within the prior six months, were currently pregnant or thought they might be pregnant, or were under the age of 18. Four participants reported pain during the testing procedures and were excluded to align with the inclusion criterion of healthy subjects and to mitigate error by maintaining a homogenous sample and similar to Glaws et al. and Stanek et al. which both included healthy participants without reported pain. The study had IRB approval. All subjects signed informed consent. The final sample consisted of 25 participants (7 male, 18 female). Participant demographics are provided in Table 1.

STUDY DESIGN

EQUIPMENT/ MATERIALS USED DURING DATA COLLECTION

Participants completed the top-tier SFMA movement patterns administered with verbal instruction, live and individually, by each rater (Appendix A). The standardized one-page scoring sheet was used to eliminate any discrepancy in criteria interpretation between researchers (Appendix B). No additional scripts, surveys, or specific software/equipment were used to conduct this study.

SFMA ADMINISTRATION TRAINING

The two student raters, rater 1 and rater 2, were third-year doctor of physical therapy students with no clinical experience performing the top-tier SFMA. Rater 1 completed an athletic training program but was not licensed and had never practiced as an athletic trainer. They participated in a three-hour in-person training course on conducting and scoring the SFMA with a certified SFMA physical therapist. The certified SFMA physical therapist serving as rater 3 had 15-hours of formal SFMA training, three-years of clinical application, and 18-years of overall clinical practice. The certified SFMA physical therapist provided the training using the SFMA training videos as the standard of instruction for teaching the top-tier movements. The instructor also demonstrated in-person to the raters the appropriate procedures for testing including movements, appropriate cues, and the importance of using multiple planes of view in order to fully appreciate the movement demonstrated. After the training and before data collection, a pilot study for internal validity was conducted on three participants concurrently assessed using the top-tier movements by all three raters. All raters recorded their scores on the participants’ movements and interpreted the results, identifying and discussing any areas of discrepancy between the raters. Variance was found during the scoring of the upper extremity and multi-segmental rotation movements. One hour of additional in-person training was administered in order to increase scoring consistency between the raters, resulting in a total of four hours of training.

TESTING PROCEDURE

The participants entered a room in the testing facility and were given a participant number. Participants completed the informed consent and HIPAA form and were screened for inclusion and exclusion criterion prior to testing. The researcher reviewed these forms with the participants and answered any questions. Participants then completed and signed consent forms.

Participants were randomly assigned to one of three examination rooms and then rotated through the other two rooms. Upon entrance, the administering rater would direct the participant to complete each of the desired top-tier movements of the SFMA. Each rater assessed the participants live and individually in separate rooms. After a brief demonstration of the desired motion, the rater assessed and scored movements as participants performed the SFMA top-tier movements. The participants were allowed three attempts to perform each specific movement, and the rater could observe the movements from any direction in order to obtain multiple planes of view.

Each participant was instructed to perform the top-tier movements in the specific order on the scoring sheet (Appendix B). This method is standardized through the SFMA in order to select the appropriate primary categorical pattern. Bilateral movements were standardized right side followed by left. After each movement, the participants were asked if they experienced any pain and their results were documented accordingly. The researchers were blinded to each other and each independently scored the SFMA with the 50-point criterion checklist scoring tool and the categorical scoring sheet. The participants were escorted to a different rater upon completion with their first rater in a sequential order after the initial randomization.

STATISTICAL ANALYSIS

Composite scores were derived from the 50-point criterion checklist and were compared between researchers. Data analysis was carried out using SPSS Version 27 (IBM Corporation, Armonk, NY). The inter-rater reliability was cal-
culated using intraclass correlation coefficients (ICC). A paired t-test assessed the difference in scores between rater 1 and 2 and a repeated measures ANOVA compared scores from raters 1 and 2 with rater 3. A two-way, mixed absolute agreement ICC analyzed results of the composite score to quantitatively measure the reliability and absolute agreement between the researchers. Normality was assessed using the Shapiro-Wilk test and Mauchly's test of sphericity. Partial eta squared effect size was interpreted as small 0.01, 0.06 medium, and 0.14 large. In order to control for type I error, a simple contrast correction was performed. Rater 3 was used as the reference category by which the other raters (e.g., rater 1 and 2) were compared. A Bonferroni correction established a new alpha of \( p = .017 \). ICC values less than 0.5 are indicative of poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.9 indicate good reliability, and values greater than 0.9 indicate excellent reliability.\(^{1,15}\)

Categorical classification of the top-tier functional movements for FN and DN used Cohen's Kappa to measure the agreement between researchers. Cohen's Kappa coefficient measured reliability between all three raters and to determine the likelihood their agreement was due to chance. Data were interpreted as statistically significant different with a \( p \)-value less than 0.05. Cohen's Kappa results were categorized as: 0.01-0.2 (1%-20%) representing slight agreement, 0.21-0.40 (20%-40%) representing fair agreement, 0.41-0.60 (40%-60%) corresponding to moderate agreement, 0.61-0.80 (60%-80%) accepted as substantial agreement, and 0.81-0.99 (80%-90%) considered almost perfect agreement, with 1.00 (100%) representing a perfect agreement.\(^{15}\)

**RESULTS**

Twenty-five volunteers, seven male and eighteen females with mean age of 23.4±1.9 years and BMI of 24.4±4.0, were analyzed and scored, subject demographic information is provided in Table 1. Normality was met.

There were significant differences in the top-tier 50-point criterion checklist between rater 1 and rater 2, \( t(24)=4.594, p<0.001 \) with a large effect size (Cohen’s \( d = 3.831 \)). Rater 1 identified more deviations from optimal top-tier movement performance using the checklist standard (Appendix B) than rater 2 (Table 2). Thus, rater 1 consistently identified more non-optimal movement patterns requiring further assessment during the 50-point criterion checklist assessment.

There were also significant differences in the 50-point criterion checklist between rater 1 and rater 3, \( F(1,24)=51.059, p<0.001 \) with large partial eta squared effect size (.680), as well as between rater 2 and rater 3, \( F(1,24)=111.484, p<0.001 \) with an even larger partial eta squared effect size (.823). Each rater identified a different total number of non-optimal movements out of the 50-points possible. After scoring all 25 participants using the 50-point criterion checklist scores, rater 1 identified a mean of 9.7 and SD of .9 non-optimal movement patterns, rater 2 identified a mean of 6.2 and SD of .7, and rater 3 identified a mean of 13.9 and SD of .8 (Table 2). Rater 1 and rater 2 had significant differences in the 50-point criterion checklist from each other and also when compared to the clinical expert, rater 3. Both rater 1 and 2 identified non-optimal movements from the top-tier screen; however, they did not identify as many when compared to the SFMA certified clinical expert. In fact, rater 1 identified 3.5 more non-optimal movements than rater 2. Rater 3 identified 7.7 more non-optimal movements than rater 2 and 4.2 more than rater 1.

The agreement between researchers for the 50-point criterion checklist results utilizing the intraclass correlation coefficient (ICC)\(_{2,1} \) measure was 0.60, \( p<0.001 \), demonstrating an overall moderate agreement between raters, \( F(24,48)=6.15, p<0.001 \) (Table 3). A higher composite score on the 50-point criterion checklist indicates more non-optimal movement patterns were identified (Appendix B, left column). Rater 3 identified the most non-optimal movement patterns and therefore had a higher scoring mean (13.92±4.17), while rater 2 had the lowest scoring mean (6.24±3.39). Rater 3 and rater 1 had a reliability of 0.78; however, reliability between rater 3 and rater 2 was 0.55 and reliability between rater 1 and rater 2 was 0.56 respectively. Cronbach’s alpha was used for internal consistency to examine each individual rater on how well they performed on scoring the 50-point criterion checklist. The item-to-total correlation for rater 2 was 0.59 suggesting that rater 2 was not similar to raters 1 and 3. When rater 2 scores were removed, the reliability increased between raters 3 and 1 with a Cronbach’s alpha of 0.87. This indicates that rater 2 was not assessing the movements as accurately as raters 3 and 1. Internal consistency should demonstrate a moderate correlation, somewhere between 0.70 and 0.90 respectively.\(^{17}\)

Categorical scoring results of the inter-rater reliability between all raters using Cohen’s kappa values ranged between slight and moderate depending on the movement pattern. Mean categorical kappa value scores for all raters

<table>
<thead>
<tr>
<th>Rater</th>
<th>Composite</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rater 1</td>
<td>9.7±.9</td>
<td>2</td>
<td>19</td>
</tr>
<tr>
<td>Rater 2</td>
<td>6.2±.7</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>Rater 3</td>
<td>13.9±.8</td>
<td>8</td>
<td>24</td>
</tr>
</tbody>
</table>

**Table 2. 50-Point Criterion Checklist Composite Score**

Note. Values are presented as Mean ± SD; Composite score is the number of non-optimal movements identified in the 50-point criterion checklist.

<table>
<thead>
<tr>
<th>Rater</th>
<th>ICC [2,1]</th>
<th>SEM</th>
<th>MDD</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Raters</td>
<td>0.6</td>
<td>4.1</td>
<td>11.4</td>
</tr>
<tr>
<td>Rater 3 to Rater 1</td>
<td>0.78</td>
<td>3.8</td>
<td>10.5</td>
</tr>
<tr>
<td>Rater 3 to Rater 2</td>
<td>0.55</td>
<td>3.8</td>
<td>10.5</td>
</tr>
<tr>
<td>Rater 1 to Rater 2</td>
<td>0.56</td>
<td>4</td>
<td>11.1</td>
</tr>
</tbody>
</table>

**Table 3. Inter-rater Reliability 50-Point Criterion Checklist Score**

Note. Values are presented as Mean ± SD; SEM=Standard Error of the Mean; MDD=Minimal Detectible Difference

---

*International Journal of Sports Physical Therapy*
was only fair; however, the mean percent agreement in the ability to identify DNs, which is important to identify since they warrant further assessment, was substantial (Table 4).

**DISCUSSION**

There is currently limited research available to describe the inter-rater reliability of raters scoring the top-tier SFMA. The primary intent of this study was to determine SFMA in-ter-rater reliability between two third-year physical therapy students following an in-person three-hour training and one-hour follow-up training with a certified SFMA physical therapist, and the secondary purpose was to compare rater scores of the composite criterion 50-point checklist and rater categorization using the top-tier movements in real-time assessments of healthy participants.

Three studies\(^1\),\(^7\),\(^8\) have previously evaluated the reliability of the SFMA. For consistency, this study followed the same top-tier SFMA movements (Appendix A) as described by Glaws et al.,\(^1\) Dolbeer et al.,\(^8\) and Stanek et al.\(^7\). Furthermore, the standardized one-page scoring sheet used in this study (Appendix B) was the same presented by Glaws et al.\(^1\) and Dolbeer et al.\(^8\) in their appendices and as described by the Functional Movement Systems, Inc. training materials. For this study for rater convenience when assessing and scoring the movements, the Criterion Checklist and Categorical scoring forms were combined on a single page for easier data completion (Appendix B).

The results of the current study indicate a significant difference with large effect size between raters 1 and 2. On average, rater 1 identified 3.43 more non-optimal movements than rater 2. The current study differed from prior SFMA reliability studies\(^1\),\(^7\),\(^8\) in that this study sought to assess if novice raters with only four hours of training and no clinical experience could consistently score in real-time, independently, using the SFMA 50-point criterion checklist. Stanek et al.\(^7\) included an undergraduate athletic training student who had fair reliability compared with two SFMA certified clinicians after completing a clinical rotation utilizing this method. In this study, rater 1 completed an athletic training program prior to directly entering the physical therapy program. It may be that movement assessment practice gained through the athletic training program enabled rater 1 to identify more non-optimal movement than rater 2. Therefore, despite having the same amount of in-person training and practice with the SFMA top-tier prior to study initiation, rater 1 with more movement assessment exposure performed better. Another possible explanation may relate to the additional experience rater 1 may have gained following assessment protocols during the athletic training program, allowing the rater to engage more familiarly with a structured, systematic format such as the SFMA assessment procedure.

The results of the present study demonstrated a significant difference with large effect size between novice raters when compared to the SFMA certified clinician. Although both novice raters identified non-optimal movements, they were not able to identify the same number as the clinical expert. Stanek et al.\(^7\) included a student, but all raters were in the room when SFMA scoring took place and the expert rater provided all verbal instructions, relieving the student...
of any responsibility or direct interaction with the participants. In contrast, student raters in this study were alone in the examination rooms and were individually responsible for verbally directing (Appendix A), assessing, and recording results in real-time. Thus, rather than simply observing and recording, these two student raters were responsible for the entire process. This may have increased the cognitive load as the raters attempted to identify as many items on the 50-point criterion checklist while also remembering to observe the quality of the movement (Appendix B, SFMA 50-Point Criterion Checklist). This distraction by the process may explain why student raters had difficulty assessing for and marking specifically articulated non-optimal movement patterns. Therefore, if the students failed to mark non-optimal movement on the criterion checklist, it would lead to inappropriate categorization.

The findings of the current study indicate that the top-tier SFMA inter-rater reliability between the three raters yielded poor to moderate agreement (ICC=0.60, CI[95%] 0.02-.84, p<0.001) of the composite scoring for all subjects (Table 3). Dolbeer et al.6 found similar inter-rater reliability (ICC=0.61, CI[95%] 0.45-.73) of the criterion checklist composite score between three certified SFMA raters with over 400 hours of application. The similarity in inter-rater reliability despite more SFMA clinical experience in the Dolbeer et al.6 study may be due to the fact that their raters assessed individuals with pain while this study focused on healthy participants. The poor to moderate inter-rater reliability (ICC=0.43, CI[95%] 0.12-.67) of the study by Glaws et al.1 may have been related to the use of video, rather than live, assessment. Allowing the rater to examine the participant with control over the instructions and the ability to move around the participant, as was done in the current study, may have allowed the student raters to demonstrate a greater degree of inter-rater reliability than the more experienced raters in Glaws et al.1

There are a limited number of studies examining the reliability of the top-tier SFMA. Prior research assessed reliably using video and real-time concurrently with varying levels of SFMA training and clinical practice resulting in varying levels of agreement, ICC=0.43, CI[95%] 0.12-.671 and ICC=0.61, CI[95%] 0.45-.736 respectively. The current study obtained similar or better reliability (ICC=0.60, CI[95%] 0.02-.84, p<0.001) from a four-hour educational training of health care providers in their third year of training, with minimal clinical exposure, and no use of SFMA in the clinic. This adds significantly to the research for a health care profession which continually assesses movement in the clinic since it is the first study to assess SFMA reliability in a real-time, clinical situation and in which the raters performed the SFMA evaluation separately.

In summary, individuals in an entry-level physical therapy educational program may not be able to identify all non-optimal movement patterns using the SFMA top-tier after limited training. Inexperience, combined with limited training time, may explain the lack of agreement between the novice and certified raters in the categorical scores. Novice student raters may have an inability to distinguish the degree of effort and asymmetry during the movement scoring when using the 50-point criterion checklist. The novice raters seemed to be more concrete, focusing on the completion of the movement without evaluating the overall quality of the movement pattern. Thus, the novice might mark a movement performed with excessive effort as complete without marking the excessive effort or lack of motor control. This, in turn, would cause the student rater to classify an effortful or uncontrolled movement as FN rather than DN. This lack of attention to details related to the quality of the movement pattern was noted in the pilot study and was a primary reason for the additional hour of training. Although the hour appeared to be sufficient in the pilot study, this gain was not retained several days later during the data collection. Those with more clinical experience and those who have completed other movement-based education with clinical rotations (i.e., athletic training program) may be more proficient at assessing movement patterns. Although four hours of training may not be sufficient to allow novice practitioners to identify musculoskeletal impairments requiring further clinical assessment, some novice third year doctor of physical therapy student practitioners demonstrate more ability to perform movement assessment, perhaps due to greater exposure to movement assessments, to identify potential clinical regions that require further assessment.

LIMITATIONS

Since this study involved live examinations in which the participants were examined three times in a row with a few minutes between each researcher, participants may have experienced a learning effect. However, this effect was controlled through the randomization of the order in which participants were seen by the researchers. Furthermore, the participants were not coached on strategies to improve their movement during the assessment as the raters were observing the instinctive, non-guided movement patterns of the individuals. Future real-time reliability studies might consider video recordings of each rater’s scoring to control for the possibility of subjects modifying or changing categories with a few repetitions of a movement pattern. Since the two novice raters were third year doctor of physical therapy students, as opposed to seasoned clinicians, they had minimal experience with both the SFMA and the evaluation of participant movement patterns within a clinical setting. However, this could be considered an accurate representation of new graduates who want to incorporate the SFMA into their assessments upon initiation of their careers. The target population of this study consisted of healthy participants that produced no positive marks on their PAR-Q form, while the SFMA is designed to be utilized on clinical patients that are experiencing musculoskeletal pain.

CONCLUSION

The results of this study indicate third year physical therapy students were able to demonstrate moderate inter-rater reliability in assessing healthy individuals using the
50-point criterion checklist for top-tier SFMA. There were differences between student raters. Variation between novice raters may reflect the amount of time accrued assessing movement and suggests that some students may require more time learning the steps involved and practicing movement assessment in order to identify non-optimal movement patterns that may require further assessment.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

ACKNOWLEDGMENTS

Authors would like to thank Daniel Harrell, Zachary Harrel-son, Chelsea Hermann, and Alison Phillips with their assistance in this research project and Sonya Harper for editing assistance.

SOURCE OF FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Submitted: March 17, 2023 CDT, Accepted: May 16, 2023 CDT

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-NC-4.0). View this license’s legal deed at https://creativecommons.org/licenses/by-nc/4.0 and legal code at https://creativecommons.org/licenses/by-nc/4.0/legalcode for more information.

International Journal of Sports Physical Therapy
REFERENCES


SUPPLEMENTARY MATERIALS

Appendix A

Appendix B
Original Research

Does High Medial Elbow Stress During Pitching Compromise the Dynamic Stabilizers of The Elbow?

Malachy P. McHugh1*, Michael J. Mullaney1

1 Nicholas Institute of Sports Medicine and Athletic Trauma

Keywords: Baseball pitching, fatigue, Motus sleeve, recovery, elbow valgus torque

https://doi.org/10.26603/001c.81596

International Journal of Sports Physical Therapy

Background

The flexor carpi ulnaris (FCU) and flexor digitorum superficialis (FDS) are thought to provide dynamic stability to the medial elbow, with a lesser contribution from the pronator teres (PT).

Hypothesis/Purpose

The purpose of this study was to determine if baseball pitchers with higher valgus elbow torque experience greater FCU and FDS strength loss.

Study Design

Controlled Laboratory Study

Methods

A pilot study was performed to determine if middle and ring finger flexion strength tests preferentially activated the FCU and FDS versus the PT (10 men age 36±12 yr). EMG amplitudes, expressed as percent of maximal voluntary contraction (MVC) were compared between tests and muscles. In a field study of college baseball pitchers, middle finger, ring finger and grip strength were tested prior to, immediately after, and one day after 14 pitching performances in 10 pitchers (21±2 yr). Elbow valgus torque was measured from an inertial measurement unit, housed in a compression sleeve and pitchers were categorized as having high or low valgus torque.

Results

For the pilot study EMG activations were 74% FDS, 66% FCU and 35% PT for the middle finger test (muscle effect p=0.032) and 95% FCU, 61% FDS and 23% PT for the ring finger test (muscle effect p=0.005). In the field study, pitchers with high valgus torque showed marked post-game middle finger fatigue (88% of baseline) and incomplete recovery the following day (95%), while pitchers with low valgus torque showed no strength loss (107% post game, 106% a day later; group x time p=0.022). Results were similar for ring finger strength (high torque: 94% post game 96% a day later; low torque: 114% post game 107% a day later; group x time p=0.048). By contrast, grip strength was not different between pitchers with high versus low valgus torque (p=0.145).

Conclusion

High medial elbow stress during pitching fatigues the dynamic stabilizers of the medial elbow.

Corresponding Author:
Malachy P. McHugh PhD
Nicholas Institute of Sports Medicine and Athletic Trauma, at Manhattan Eye Ear & Throat Hospital (A division of Lenox Hill Hospital)
210 East 64th Street, 5th Floor, New York, NY 10065 USA
E-mail address: mchugh@nismat.org

*a Corresponding Author:
INTRODUCTION

Elbow injuries in general, and ulnar collateral ligament (UCL) injuries in particular, have become an increasing problem in baseball. Over the last 25 years the predominant injury in major league baseball (MLB) has shifted from the shoulder to the elbow. The vast majority of UCL injuries are seen in pitchers at both the professional and collegiate level.

The muscles of the flexor pronator mass can provide dynamic stability to the medial elbow. Specifically, the flexor carpi ulnaris (FCU) and the flexor digitorum superficiales (FDS) have been shown to provide dynamic stability, with a lesser contribution from the pronator teres (PT). In line with these findings, exercise-induced wrist flexion fatigue has been shown to increase medial elbow gapping under valgus stress. The FDS and FCU are highly active during the acceleration and deceleration phases of the pitching motion (FDS 80%-71% MVC, FCU112%-77% MVC) and likely provide dynamic elbow stability. In support of this, FCU and FDS elasticity have been shown to increase after repeated pitches and the change was correlated with an increase in medial elbow joint space.

In response to the increased prevalence of elbow injuries in baseball there has been increased interest in measuring the stress in the medial elbow during pitching. To this end, a wearable inertial sensor has been developed to quantify medial elbow torque during pitching. Medial elbow valgus torque, an index of valgus stress on the elbow, was shown to be higher for fastballs and changeups compared with curveballs. Additionally, valgus torque increased in later innings of a simulated game, despite a decline in pitch velocity, and this effect was attributed to fatigue. However, the association between valgus torque at the elbow during pitching and muscle fatigue in the dynamic stabilizers of the elbow has not been previously examined. It is possible that pitchers with high valgus torque during pitching experience greater fatigue in the dynamic stabilizers of the elbow and thus are at increased risk of injury.

The first step in studying a possible relationship between high valgus elbow torque during pitching and excessive fatigue in the dynamic stabilizers of the elbow is to develop a strength test for the dynamic stabilizers. Previous work in college pitchers has shown that grip strength is minimally impaired post game or the next day. It remains unclear whether grip strength is a good indicator of fatigue in the stabilizers of the medial elbow. Developing a strength test specific to the FCU and FDS could provide important information on dynamic stability of the medial elbow in pitchers. Middle finger extension strength measured with a hand-held dynamometer has been shown to provide a good measure of impairment in patients with lateral elbow tendinopathy. Applying the same methodology to finger flexion strength might provide an indicator of medial elbow function.

Therefore, the purposes of this study were twofold: (1) initially determine if finger flexion strength testing with a hand-held dynamometer selectively activates the FCU and FDS, (2) determine if baseball pitchers with higher valgus elbow torque experience greater FCU and FDS strength loss. It was hypothesized that the FCU and FDS would be preferentially activated during middle and ring finger flexion MVC testing, and that pitchers with higher elbow valgus torque would experience greater finger flexion fatigue and slower recovery.

MATERIALS AND METHODS

This study was divided into two sections: (1) a pilot study in healthy control subjects to develop FCU and FDS strength tests and (2) a field study in college baseball pitchers to examine fatigue in the FCU, FDS and grip tests and the relationship to valgus elbow torque during pitching. The pilot study and the field study were approved by institutional review board and all participants gave written informed consent.

PILOT STUDY: DEVELOPMENT OF FCU AND FDS STRENGTH TESTS

Electromyograms (EMG) were recorded from the FCU, FDS and PT during manual muscle testing of middle finger and ring finger flexion strength on the dominant side in 10 healthy men (age 36±12 yr) using a 16-channel BTS FREEEMG 500 system, CMRR: >110 dB at 50–60 Hz; input impedance: >10 GΩ (BTS Bioengineering, Milan, Italy). After the skin of each participant was shaved, cleaned, and lightly abraded, disposable silver/ silver chloride passive dual electrodes (2.0-cm interelectrode distance) (Noraxon, Scottsdale, AZ) were applied. Electrodes were placed over muscle belly at the proximal median forearm (FCU), one quarter the distance from medial humeral epicondyle to the skin fold at the wrist (FDS) and diagonally over muscle belly, slightly distal to skin fold in cubital fossa (PT). Maximum voluntary contraction (MVC) tests were performed for each muscle: maximal resisted wrist flexion and ulnar deviation for FCU testing; maximal resisted four-finger flexion with opposition applied across the middle phalanges for FDS testing; and maximal resisted forearm pronation with slight elbow flexion for PT. Since these MVC tests are not conducive to testing with a hand-held dynamometer for quantifying strength, the middle finger and ring finger flexion tests were chosen for strength testing.

For both the pilot study and field study finger flexion strength was performed with a hand-held dynamometer (Lafayette Instruments, Lafayette, IN). Subjects were seated for testing with the elbow flexed approximately 40°, with the forearm supinated and rested on a flat surface with the wrist in neutral position (Figure 1). The subject stabi-
lized their forearm with their contralateral hand, pressing it firmly against the table while avoiding the surface electrodes. The subject then flexed the test finger (middle or ring finger) while the tester stabilized the other three fingers against the table. Then the tester placed the 1 cm diameter probe of the dynamometer distally on the middle phalanx just proximal to the distal interphalangeal joint. The subject maximally flexed the test finger against the dynamometer while the tester extended the finger (break test). Two trials were performed for each test. A third trial was performed if there was a marked difference between the first two trials (>50%) and subsequently the outlier was discarded and the mean of two was recorded. Index finger flexion break test strength has previously been shown to be reliable.18 Additionally, a hand-held dynamometer test set up has previously been used to detect middle finger extension weakness and treatment training effects in patients with lateral epicondylosis.14

PILOT STUDY STATISTICS

Middle versus ring finger strength were compared using a paired t test. Differences in FCU, FDS and PT activation between the middle and ring fingers were compared using Test (middle versus ring finger) by Muscle (FCU, FDS, PT) repeated measures analysis of variance. The dependent variable was EMG expressed as a percentage of MVC. Pairwise comparisons between tests and between fingers were made using the least significant difference approach.

FIELD STUDY IN PITCHERS

PROCEDURES

Middle and ring finger flexion strength and grip strength were measured immediately prior to a game, immediately after the pitcher was removed from the game, and on the day after the game. The study participants were 10 NCAA division 3 baseball pitchers (3 freshmen, 1 sophomore, 2 juniors, 4 seniors; age 21±2 yr, height 1.83±0.06 m, body mass 85.4±9.2 kg). Data were collected during intra-squad scrimmages and preseason games in the Fall season, with 6 pitchers having data from one performance and 4 pitchers having data from two performances (total 14 game data sets). Elbow valgus torque was measured from an inertial measurement unit (Motus Global, Massapequa, NY), housed in a compression sleeve, worn on the elbow during pregame bullpen pitches. The inertial measurement unit was removed prior to the game (the pitchers preferred not to wear the device). The elbow valgus torque from the bullpen pitches was averaged for each pitcher. Pitch velocities were not measured for the games used in this study. However, each player had his fast ball velocity measured earlier in the season which was used during analysis. All participants gave written informed consent, and the study was approved by institutional review board.

Middle and ringer finger flexion strength were measured as described in the methods for the pilot study. Grip strength measurements were taken in a standing position using a hydraulic hand dynamometer (Jamar; Performance Health, Warrenville, IL). Pitchers were instructed to have their shoulder adducted and neutrally rotated, elbow flexed at 90°, and forearm in neutral position during the grip test.

Figure 1. (A) Middle finger flexion strength test. (B) Ring finger flexion strength test. The dynamometer was placed just proximal to the distal interphalangeal joint.
Pitchers were instructed to squeeze the dynamometer as hard as they could (isometric test). The average of two trials was recorded and used for data analysis.

The inertial measurement unit was aligned with the medial aspect of the ulna approximately 5 cm distal to the medial epicondyle of the humerus and held in place with a commercial sleeve. In addition to elbow torque (peak elbow valgus torque) the sensor calculates arm slot angle (angle of forearm relative to the ground at ball release), arm speed (peak forearm angular velocity in degrees per second) and shoulder rotation angle (angle between forearm and ground at maximum external rotation). Pitchers were classified as high elbow valgus torque (above the mean for the group) or low elbow valgus torque (below mean for the group).

**FIELD STUDY STATISTICS**

Pitchers were categorized as having high or low valgus torque (see results section for details). Effect of valgus elbow torque on fatigue and strength recovery was assessed using mixed-model ANOVA. Based on the pilot study, it was estimated that with 14 pitching performances there would be 80% power to detect a 11% change in middle finger strength after a pitching performance (p<0.05). Pearson correlation coefficients and stepwise multiple regression were used to identify factors associated with post game fatigue.

**RESULTS**

**PILOT STUDY**

Finger flexion force was significantly greater for the middle finger test versus the ring finger test (75±17 N vs. 52±9 N, P<0.001). Muscle activations varied significantly between muscles and between tests (Table 1). FDS activity was greater than PT activity for both the middle finger test (p<0.005) and the ring finger test (p=0.002). FCU activity was higher than PT activity for the ring finger test (0.010). Therefore, the finger flexion tests preferentially activated the FDS and FCU over the PT.

**FIELD STUDY**

The pitchers threw 58±13 pitches in the 14 games (range 42 to 82 pitches) not including bullpen warm up throws. For the bullpen pitches valgus elbow torque was 50.7±14.5 Nm with an arm slot angle of 48±9°, an arm speed of 912±116°/s and a shoulder rotation angle of 145±10°. The number of bullpen pitches varied between pitchers (range 20-29), but elbow valgus torque for a given pitcher did not vary markedly between pitches, with an average coefficient of variation of 14%. Elbow torque values were very similar between games for the four pitchers with data from two games, varying by 6.6%, 9.5%, 0.2%, and 5.4% (Figure 2).

Four pitchers (6 games) had elbow torques below 50 Nm (35.9±3.2 Nm, range 30.9-40.2 Nm). Six pitchers (8 games) had elbow torques above 50 Nm (61.7±6.5 Nm, range 52.5-68.8 Nm). The total number of game pitches was not different between the low and high elbow valgus torque groups (53±11 vs 61±15, p=0.263). The high torque group comprised pitchers with higher (p=0.010) pitch velocity (87±2 mph) than the pitchers in the low torque group (82±2 mph).

Pregame strength was 77.1±11.3 N for middle finger flexion, 58.5±11.2 N for ring finger flexion and 552±76 N for grip strength. Middle and ring finger flexion fatigue, and subsequent recovery, were worse for pitchers with high valgus elbow torque compared with pitchers with low valgus elbow torque (Group by Time p=0.022 middle finger, Figure 3, p=0.031 ring finger, Figure 4). Pitchers with high torque had greater post-game fatigue than pitchers with low torque for middle finger flexion (88±10% of baseline strength vs. 107±18%, p=0.026) and ring finger flexion (94±15 vs. 114±13%, p=0.016). By contrast, changes in grip strength were unaffected by elbow torque (Group by Time p=0.145, Figure 5).

Post-game middle finger flexion fatigue was correlated with the number of pitches thrown (r=0.617, p=0.019) and elbow valgus torque (r=0.597, p=0.024). The combination of pitch count and elbow valgus torque explained 57% of the variance in middle finger flexion fatigue (R=0.755, p=0.010). Post-game ring finger flexion fatigue was correlated with the number of pitches thrown (r=0.352, p=0.041) and elbow valgus torque (r=0.652, p=0.015). The combination of pitch count and elbow valgus torque did not further improve the prediction of ring finger fatigue. Neither pitch count (r=0.058, p=0.845) nor elbow torque (r=0.426, p=0.129) were correlated with grip fatigue.

**DISCUSSION**

In the pilot study the EMG data from the FCU, FDS and PT indicated that middle and ring finger flexion tests activated the FCU and FDS substantially more than the PT. Previous literature showed that the FCU and FDS provide dynamic stability to the medial elbow.3-7 In the field study pitchers with greater elbow valgus stress during pitching experienced greater fatigue in the finger flexion tests. This indicates that pitchers with high valgus elbow stress place greater demand on the dynamic stabilizers and thus exhibit greater fatigue in these muscles. By contrast, change in grip strength, which is a more general test of forearm fatigue, was unaffected by elbow valgus stress.

It was anticipated that the middle and ring finger flexion tests would activate the FDS since the MVC test for the FDS is resisted four finger flexion. A hand-held dynamometer

---

**Table 1. Muscle activation (%MVC) during middle and ring finger flexion strength tests (mean±SD)**

<table>
<thead>
<tr>
<th></th>
<th>MIDDLE FINGER</th>
<th>RING FINGER</th>
</tr>
</thead>
<tbody>
<tr>
<td>FCU</td>
<td>66±44%</td>
<td>93±64%</td>
</tr>
<tr>
<td>FDS</td>
<td>74±25%</td>
<td>61±31%</td>
</tr>
<tr>
<td>PT</td>
<td>35±22%</td>
<td>23±13%</td>
</tr>
</tbody>
</table>

FCU, flexor carpi ulnaris; FDS, flexor digitorum superficialis; PT, pronator teres. Muscle Effect p=0.008; Test Effect p=0.874; Muscle by Test p=0.01. Middle Finger muscle effect p=0.032; FDS>PT p=0.005. Ring Finger muscle effect p=0.003; FCU>PT p=0.010; FDS>PT p=0.002.
Figure 2. Elbow valgus torque for the 10 pitchers. Players were classified as high (n=6) or low (n=4) valgus torque pitchers. Mean±SD displayed.

Figure 3. Middle finger flexion strength immediately post game and one day later. Effect of Time p=0.587, Effect of Elbow Torque p=0.048, Torque by Time p=0.022. * Significant difference between low and high torque groups (p<0.05). Mean±SE are displayed.

test of four finger flexion was not attempted because there would be no way to uniformly apply resistance across the four fingers to allow comparable contributions from each finger. Since the FCU tendon inserts on the pisiform and is not attached to the carpal bones, it is surprising that it was so highly active during ring finger flexion. This may reflect synergistic action between finger and wrist flexion and the role of the pisiform. In fact, ring finger flexion has been shown to contribute to medial stability of the elbow. However, both middle finger flexion and index finger flexion have been shown to provide greater medial stability that
Figure 4. Ring finger flexion strength immediately post game and one day later. Effect of Time p=0.624, Effect of Elbow Torque p=0.048, Torque by Time p=0.031. * Significant difference between low and high torque groups (p<0.05). Mean±SE are displayed.

Figure 5. Grip strength immediately post game and one day later. Effect of Time p=0.013, Effect of Elbow Torque p=0.077, Torque by Time p=0.143. Mean±SE are displayed.

Future work should examine index finger strength changes in baseball pitchers.

While the focus of this study was middle and ring finger flexion strength in relation to medial elbow stability, previous work has shown that full grip contraction (mass grasp) can stabilize the medial elbow. However, changes in grip strength in the present study were not associated with the valgus torque on the medial elbow during pitching. Therefore, these results support finger flexion strengthening exercises for baseball pitchers as opposed to grip strengthening exercises. Similarly, others have proposed that index
and middle finger flexion exercises be added to throwers’ injury prevention training programs.\(^{19}\)

Higher pitch velocities are associated with greater valgus torque at the elbow, however, this relationship is more pronounced within pitchers than between pitchers.\(^{21}\) Pitch velocity explained 96% of the variance in elbow torque when looking at individual professional pitchers throwing at a range of different velocities.\(^{21}\) By contrast, pitch velocity only explained 8% of the variability in elbow torque when comparing data across a group of professional pitchers.\(^{21}\) Thus, variation in individual pitchers’ stature, anthropometrics and biomechanics likely contribute markedly to the elbow torque.

The wearable inertial sensor used in this study can provide useful information for monitoring and managing pitchers. Part of the rationale for the study was the finding that valgus elbow torque measured with this sensor increased progressively after three innings of a simulated game despite a decline in pitch velocity.\(^{11}\) This increased medial elbow stress was attributed to fatigue, but the only measure of fatigue was a visual analog scale used to capture the pitcher’s perception of fatigue with each inning pitched. There was clear heterogeneity in valgus elbow torque between the pitchers studied here, and homogeneity between games for pitchers with data from two games. Given the individual nature of valgus torque profiles in pitchers\(^ {10,11}\) a prospective study of valgus elbow torque and injury risk across a large sample of pitchers is warranted.

One limitation was the relatively small number of pitches (\(58\pm15\)) thrown by the pitchers in this study. Because the games were intrasquad scrimmages and pre-season games, pitchers were scheduled to pitch a set number of innings in each game. Most pitchers were restricted to four innings thus limiting the total number of pitches. Middle and ring finger fatigue were correlated with pitch count. For the middle finger, pitch count and elbow torque combined to explain 57% of the variance in fatigue. For the ring finger elbow torque explain 40% of the variance in fatigue and pitch count explained 30%, but the combination did not further explain the variance.

While the sample size was sufficient to demonstrate an effect of elbow valgus torque on finger flexion fatigue the data should be viewed as preliminary. Data on a greater number of pitchers throwing a greater number of pitches is needed to fully understand the role of elbow valgus torque in fatigue of the dynamic stabilizers of the medial elbow. Additionally, the sample size was insufficient to fully examine recovery of finger flexion strength on the day after the game. The data were equivocal with regards to the role of elbow torque and the extent of recovery.

Ideally pitch velocity would have been measured in the games studied here (instead of at an independent time during the season) since higher velocities are associated with greater elbow valgus torque.\(^{21}\) However, the relationship between fastball velocity and elbow torque was not strong in professional pitchers.\(^{21}\) Furthermore, weighted-ball throwing training was sufficient to increase pitch velocity with no change in elbow torque.\(^ {22}\) It remains to be determined whether high elbow valgus torque independent of pitch velocity leads to earlier fatigue of the dynamic stabilizers of the medial elbow. If so, identifying high velocity pitchers that pitch with low valgus elbow torque may be beneficial.

It would have been preferable if the EMG measurements to validate the tests of FCU and FDS function were performed on the baseball pitchers. However, this would have involved significantly more time with the pitchers, who were in season at the time of data collection. The team did not provide additional time to conduct EMG testing on the players. Using a healthy control group in a laboratory setting provided a greater degree of control than would have been possible with the pitchers and was able to confirm the recruitment of key muscles during various tests.

While the results of this study point to the potential utility of finger flexion strength measures in baseball pitchers it is important to establish the reliability of this measure and the degree of tester expertise required to obtain reliable measures. The tester in this study had more than 20 years of experience using this hand-held dynamometer to test baseball pitchers. Had there been substantial measurement variability it would not have been possible to detect the effects of elbow valgus torque on finger flexion strength. However, since this is not a widely used strength test it remains to be seen how useful the measurement is in the hands of other clinicians.

CONCLUSION

The results of this study indicate that pitchers with higher valgus elbow torque during pitching experience greater fatigue in the dynamic stabilizers of the medial elbow. This may lead to excessive stress on the UCL. These pitchers may benefit from finger flexion strength training to maintain dynamic stability of the medial elbow during a pitching performance.

ACKNOWLEDGEMENTS

This study was presented at the IJSPT Research Symposium at the Orthopedic Summit. Dec 11-14, 2021, Las Vegas, NV.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to report.

Submitted: February 15, 2023 CDT, Accepted: May 06, 2023 CDT
REFERENCES


Advancements for the Future: A National Survey of Fastpitch Softball Coaches’ Perspectives on Injury Prevention Programming

Jaime Nurrenbern¹, Courtney Klueppel¹, Jeff Tilly², Kyle Matsel¹²³@¹

¹ Physical Therapy, University of Evansville, ² Athletic Training, University of Evansville, ³ ProRehab, PC

Keywords: injury prevention programs, risk factors, softball coaches, survey

https://doi.org/10.26603/001c.83262

International Journal of Sports Physical Therapy

Background
Approximately 50 percent of softball injuries are the result of overuse or chronic conditions. However, research exploring preventative measures for softball players is limited and usage of injury prevention strategies among softball coaches is unknown.

Hypothesis/Purpose
This survey aimed to investigate if softball coaches are implementing injury prevention programs to reduce injury and improve the performance of their players. The secondary purpose was to identify barriers to the implementation of injury prevention programs. Finally, this survey explored the coaches’ knowledge of injury risk factors and their views on design and usage of preventative programs.

Study Design
Descriptive cross-sectional survey

Methods
A 35-item survey was sent to approximately 14,000 high school and collegiate fastpitch softball coaches throughout the United States. Data were collected over a three-month period with an overall response rate of 1.2%.

Results
Among responding coaches, 45.9% (n=79/172) reported implementing injury prevention programs. Coaches who implement injury prevention strategies most frequently utilize team-based programs (68.8%, n=52/95) compared to group-based (19.0%, n=15/95) or individualized programs (15.2%, n=12/93). Coaches who do not use preventative programming reported that being unsure of what program to perform (53.8%, n=50/93) and not having enough staff (20.4%, n=19/93) were the greatest barriers to implementation. Although over 50% of coaches recognized arm fatigue/overuse (27.9%, n=48/172) and decreased core strength (22.7%, n=39/172) were important risk factors, 36% (n=94/172) "disagree" that softball pitchers should adhere to pitch counts and 90% (n=85/92) believe that preventative programming for pitchers and position players should be similar.

Conclusion
Less than 50 percent of softball coaches implement exercise programs to prevent injury. Limited familiarity with effective program design, inadequate staffing, and inconsistent risk factor awareness are the major contributors to lacking implementation.

Corresponding Author:
Kyle Matsel PT, DPT, PhD
University of Evansville
515 Bob Jones Way
Evansville, IN 47708
Phone: (812) 488-5105
Email: km218@evansville.edu

a Corresponding Author:
Collaboration between rehabilitation professionals and softball coaches regarding preventative programming should be considered.

Level of Evidence
Level 3
©The Author(s)

INTRODUCTION

Softball is a rapidly growing and evolving sport, drawing the attention of millions across the world. With the amplifying intensity of softball games, there has been an associated increased prevalence of musculoskeletal injury in the sport. The injury incidence for high school and college fast pitch softball (FPS) players is 1.16 and 3.19 per 1000 athletic exposures (AE), respectively. Although several injuries occur from trauma associated with sliding, being hit by a pitch, or sustaining a concussion, approximately fifty percent of FPS injuries are from overuse or chronic conditions. Due to similarities between softball and baseball, such as dominant shoulder retrotorsion and comparable shoulder distraction torque during pitching, most research has focused primarily on the injury risk prevention of baseball players. As a result, there is a paucity of research exploring injury rates and preventative measures for FPS players.

Musculoskeletal injuries can affect various body locations and occur frequently in both softball pitchers and position players. Interestingly, position players tend to more commonly experience lower extremity injuries whereas pitchers sustained a more equal distribution of injuries between the upper and lower extremities. FPS pitchers appear to be more vulnerable to time-loss injuries compared to position players. In a systematic review, Paul et al reported that 37.5% of injuries in FPS pitchers were directly attributed to pitching with 61% of those involving the shoulder. These findings are suggestive that the biomechanics behind pitching, especially in the shoulder, and pitching volume could expose pitchers to throwing-related injuries. Furthermore, the prevalence of injury increases as the experience of the softball pitcher increases. For example, college softball pitchers tend to demonstrate less upper extremity range of motion when compared to high school-level pitchers. This decrease in mobility could be a contributing factor to increased injury rates and has led coaches to consider implementing dynamic warm-ups and preventative exercise programs.

Injury prevention programs generally consist of stretching or strengthening exercises that are designed to maintain adequate mobility and strength throughout the season. Although much attention has been focused on shoulder strengthening, declines in the function of the lower body are common. Specifically, the microtrauma from the stress endured throughout the softball season results in declines in strength, altered shoulder total arc of motion, and decreased hip rotational range of motion. Due to the diverse nature of softball injuries, many researchers have suggested that injury prevention programs should be individualized to address the unique limitations of the players. However, current perspectives from high school baseball coaches reflect a preference toward the utilization of team-based injury prevention to maintain arm health. Unfortunately, established evidence-based injury prevention programming does not currently exist for softball players. Furthermore, it is unknown to what extent fastpitch softball coaches are using novel injury prevention programming and how those programs are structured.

The authors are unaware of another study that has explored FPS coaches’ views on injury prevention programs and their willingness to implement an evidence-based program. Therefore, the purpose of this study was to investigate if FPS coaches are implementing injury prevention programs to reduce injury and improve the performance of their players. It was hypothesized that less than 50% of responding FPS coaches are implementing injury prevention programs. The secondary purpose was to identify common barriers to the implementation of injury prevention programs. It was hypothesized that a lack of access to effective injury prevention programming and resources would be identified as the greatest barrier to implementation. Finally, this survey explored the coaches’ knowledge of injury risk factors and their views specific to the structural design and usage of injury prevention programs. It was hypothesized that most coaches who are performing injury prevention will be using team or group-based programming and will demonstrate a limited understanding of risk factor management.

MATERIALS AND METHODS

SURVEY DEVELOPMENT

This was a descriptive, cross-sectional study using an electronic survey to evaluate the current usage and effectiveness of injury prevention programs among softball coaches in the United States. The electronic survey (Appendix 1) was developed in Qualtrics (electronic data capture tools hosted at the Rocky Mountain University of Health Professions) based on the current literature involving injury prevention exercise programs in softball players. The survey was created by two physical therapy students (JN and CK) who modified it from a previous survey conducted by Matsel et al which explored high school baseball coaches’ perspectives on injury prevention implementation. The students collaborated with an associate professor (KAM) with multiple years of clinical and research experience working with softball athletes during survey development.

SURVEY SAMPLING

Approximately 14,000 FPS coaches were contacted through email to participate in the online survey. Distribution of the survey was administered by the Clell Wade Coaches On-
line Directory. This organization has databases containing the email addresses of several thousands of high school and collegiate FPS coaches. The Clell Wade organization contacted the participants directly through their email database and provided them with the electronic link to the survey. In addition to using Clell Wade Coaches Online Directory, the survey was posted on multiple social media platforms to reach FPS coaches who are not represented in this database.

The investigators of this study did not have access to participants’ emails or any other personally identifiable information. Participation in the survey was voluntary and all responses to the survey questions were anonymous. Prior to beginning the survey, participants read the consent form and checked the “I agree” option if they wished to consent to the study. Participants were included in the study if they were willing and able to complete the online survey. Participants were excluded if they did not complete all the required survey questions, could not read English, or selected the “I do not agree” to consent option prior to filling out the survey. Responses to the online survey were prospectively collected for three consecutive months from June to August 2022. Approval from the institutional review board at the University of Evansville was obtained prior to data collection for this descriptive survey study.

STATISTICAL METHODS

Descriptive statistics for nominal and ordinal data were summarized through frequencies and percentages and analyzed for differences with a chi-square test. Cross tables and chi-square tests of independence were used to consider associations between the use of injury prevention programs and coaching experience, age, and education level. An alpha level of $p < 0.05$ was considered statistically significant for all tests. All data analyses were performed with R for Mac OS 4.1.2 statistical software (RStudio for Mac, Version 2022.12.0+353).

RESULTS

Demographic information for the level and age of softball players coached and the responding coaches’ years of coaching experience is presented in Table 1. A total of 218 FPS coaches throughout the United States responded to the survey between June 1, 2022 and August 30, 2022. Of these 218 surveys, 46 surveys were excluded due to being incomplete or failing to meet the inclusion criteria. Therefore, the remaining 172 surveys represented an inclusion rate of 79% (172/218) and were used for data analyses.

INJURY PREVENTION PROGRAMMING

Among softball coaches who responded to the survey, 45.9% (n=79/172) reported that their players participate in an injury prevention program (Table 2). Coaches who implement injury prevention strategies most frequently utilize team-based programs (68.8%, n=52/79) compared to group-based (19.0%, n=15/79) or individualized programs (15.2%, n=12/79) which focus on specific demands or limitations unique to each player ($p=0.001$). Interestingly, 85% (n=146/172) of responding coaches would be interested in a screening tool to better inform individualized injury prevention initiatives (Table 2).

Collegiate-level softball coaches are more likely to implement injury prevention programs ($p<0.005$) and more likely to individualize their programs ($p<0.001$) to meet the specific needs of their players compared to high school-level coaches. However, there was no significant difference between softball coaches’ usage of injury prevention programming and years of coaching experience ($p=0.22$) (Table 3). Over 50% (50.6%, n=40/79) of coaches implementing injury prevention programs report that preventative exercises should be performed throughout the season and completed at least two to three times per week (64.5%, n=51/79). Furthermore, 44% (n=41/93) of coaches consider 5-10 minutes of practice time a reasonable amount of time to dedicate to injury prevention exercises (Table 2).

BARRIERS TO INJURY PREVENTION PROGRAMMING

The coaches’ perceptions of the greatest barriers to implementing injury prevention programs are summarized in Figure 1. Responding softball coaches who did not perform injury prevention programs reported that being unsure of what program to perform (53.8%, n=50/93) and not having enough staff to assist with program design and execution (20.4%, n=19/93) were the greatest barriers to implementation. Only 4.4% of 93 softball coaches reported not implementing preventative exercises due to previous experience with ineffective programs (2.2%, n=2/93) or beliefs that injury prevention programs typically do not affect injury occurrence (2.2%, n=2/93). Further, if an evidence-based injury prevention program were developed that showed promise at reducing injury incidence in softball players, 73% (n=68/93) of the coaches currently not implementing preventative programming would reconsider inclusion of programming.

RISK FACTOR MANAGEMENT AWARENESS

Most responding FPS coaches (64.6%, n= 51/79) who have their players perform exercise programs reported that the program’s main goal is the prevention of injuries (Figure 2). Among FPS coaches, 83.7% (n=144/172) "strongly agree" that injury risk factors should be monitored throughout the entire season. Participating coaches reported that arm fatigue/overuse (27.9%, n=48/172) and decreased core strength (20.7%, n=39/172) were the most important risk factors affecting player availability during the season. Interestingly, 36% (n=64/172) of surveyed coaches "disagree" or "strongly disagree" that softball pitchers should have a pitch count limitation. Furthermore, 90% (90.2%, n=83/92) of responding coaches reported that injury prevention programming should be similar for pitchers and position players. Overall, FPS coaches reported that the players (50%, n=86/172) and coaches (35.5%, n=61/172) play the largest role in preventing softball injuries (Figure 3).
Table 1. Demographics of Fastpitch Softball Coaches

<table>
<thead>
<tr>
<th>What level of team do you coach?</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High school</td>
<td>107 (62)</td>
</tr>
<tr>
<td>Travel</td>
<td>38 (22)</td>
</tr>
<tr>
<td>10U</td>
<td>6 (15.8)</td>
</tr>
<tr>
<td>12U</td>
<td>8 (21.1)</td>
</tr>
<tr>
<td>14U</td>
<td>14 (36.8)</td>
</tr>
<tr>
<td>16U</td>
<td>7 (18.4)</td>
</tr>
<tr>
<td>18U</td>
<td>3 (7.9)</td>
</tr>
<tr>
<td>College</td>
<td>27 (15.7)</td>
</tr>
<tr>
<td>Division I</td>
<td>5 (18.5)</td>
</tr>
<tr>
<td>Division II</td>
<td>3 (11.1)</td>
</tr>
<tr>
<td>Division III</td>
<td>6 (22.2)</td>
</tr>
<tr>
<td>NAIA</td>
<td>8 (29.6)</td>
</tr>
<tr>
<td>NJCAA</td>
<td>5 (18.5)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>How many years have you been involved in coaching softball?</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 1 years</td>
<td>12 (6.98)</td>
</tr>
<tr>
<td>1-3 years</td>
<td>28 (16.3)</td>
</tr>
<tr>
<td>4-6 years</td>
<td>29 (16.9)</td>
</tr>
<tr>
<td>7-10 years</td>
<td>28 (16.3)</td>
</tr>
<tr>
<td>&gt;10 years</td>
<td>75 (43.6)</td>
</tr>
</tbody>
</table>

* NAIA = National Association of Intercollegiate Athletes
* NJCAA = National Junior College Athletic Association

DISCUSSION

The overarching aim of this survey was to explore the practice and implementation of injury prevention programs by FPS coaches. The results of the survey support the primary hypothesis that less than 50% of responding FPS coaches are implementing injury prevention programs. Overall, most coaches who implement injury prevention programs utilized team-based or group-based programs. However, college-level coaches implement injury prevention programs and individualize those programs more frequently compared to high school coaches. The coaches agree that injury prevention programs are effective at reducing injuries, but largely fail to implement them due to uncertainty of what programs to perform.

There is limited research specific to structured injury-preventative programming for softball players which makes evidence-based program development decisions difficult for coaches. However, there are consistencies between the results of this current survey of fastpitch softball coaches and previous research on perceived vital programming components. In a recent systematic review, Paul et al.1 highlighted the most common injury risk factors in softball players were associated with the shoulder complex, core stabilization, and hip strength. In the current survey, coaches highlighted the importance of shoulder and hip strengthening to improve durability throughout the season. However, coaches failed to recognize the importance of core stability as a risk factor by ranking the shoulder (97.83%, n=90/451), knee (77.17%, n=71/451), and hip (70.65%, n=65/451) ahead of the core as the three most important body regions to include in injury prevention programming. Although a direct relationship between core stability and injury incidence has not be established in softball players,19 core stability is a foundational component of overall athlete performance especially in rotational sports such as softball.20

Although rotational sports such as softball and baseball are commonly referred to as overhead sports which suggests a desire to strengthen the throwing shoulder, the entire kinetic chain including the lower body plays a role in player durability and performance. In the current survey, responding coaches suggest that hip and knee exercises should be included in programming, but lower body (hip, knee, ankle, foot) mobility and strength were not rated as major injury risk factors (9.88%, n=17/172). One possible explanation could be that coaches consider the lower body important for performance enhancement but not as critical for injury prevention. Guy et al.15 reported that hip range of motion and strength appears to decrease throughout the softball season in both pitchers and position players. Hip mobility declines have also been demonstrated in baseball players throughout the season21,22 and have been associated with increased stress on the throwing arm.23 To date, no studies have specificity associated hip mobility declines with increased risk of arm injury in softball players; however, it seems logical that maintaining hip mobility and strength throughout the season could be advantageous.
Table 2. Injury Prevention Programming Application and Usage

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do your players perform an injury prevention program?</td>
<td>79 (45.9)</td>
<td>93 (54.1)</td>
<td></td>
</tr>
<tr>
<td>Which of the following best describes your injury prevention program?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Team-Based – every athlete performs the same routine</td>
<td>52 (68.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group Based – different program based on playing position</td>
<td>15 (19.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Individualized – program based on individual limitations</td>
<td>12 (15.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Would you be interested in a 3-minute screen to help individualize your</td>
<td>146 (84.9)</td>
<td>26 (15.1)</td>
<td></td>
</tr>
<tr>
<td>injury prevention program?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What is a reasonable amount of time to dedicate to injury prevention during</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>practice?</td>
<td>≤ 5 minutes</td>
<td>7 (7.61)</td>
<td></td>
</tr>
<tr>
<td>6-10 minutes</td>
<td>41 (44.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11-15 minutes</td>
<td>33 (35.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16-20 minutes</td>
<td>11 (12.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>How many times a week do your players perform your injury prevention</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>program?</td>
<td>1x/week</td>
<td>3 (3.8)</td>
<td></td>
</tr>
<tr>
<td>2x/week</td>
<td>17 (21.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3x/week</td>
<td>34 (43.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4x/week</td>
<td>6 (7.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5x/week</td>
<td>12 (15.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6x/week</td>
<td>5 (6.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7x/week</td>
<td>2 (2.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>What time of year do you have your players perform your injury prevention</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>program?</td>
<td>Year-round</td>
<td>26 (32.9)</td>
<td></td>
</tr>
<tr>
<td>Preseason only</td>
<td>7 (8.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>During the season only</td>
<td>5 (6.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Offseason only</td>
<td>1 (1.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-season through end of season</td>
<td>40 (50.6)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Relationship Between Player Level or Coaching Experience & Injury Prevention Usage

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is coaching experience related to injury prevention implementation?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-6 years of experience</td>
<td>38.0%</td>
<td>62.0%</td>
<td>0.22</td>
</tr>
<tr>
<td>&gt;7 years of experience</td>
<td>48.4%</td>
<td>51.6%</td>
<td></td>
</tr>
<tr>
<td>Is coaching experience related to use of individualized injury prevention</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1-6 years of experience</td>
<td>86.7%</td>
<td>13.3%</td>
<td>0.97</td>
</tr>
<tr>
<td>&gt;7 years of experience</td>
<td>83.7%</td>
<td>16.3%</td>
<td></td>
</tr>
<tr>
<td>Is the level of softball players related to injury prevention usage?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High School/Travel team level players</td>
<td>41.4%</td>
<td>58.6%</td>
<td>0.005*</td>
</tr>
<tr>
<td>College level players</td>
<td>70.4%</td>
<td>29.6%</td>
<td></td>
</tr>
<tr>
<td>Is the level of player related to use of individualized injury prevention</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High School/Travel team level players</td>
<td>17.2%</td>
<td>82.8%</td>
<td>&lt;0.001*</td>
</tr>
<tr>
<td>College level players</td>
<td>93.3%</td>
<td>6.7%</td>
<td></td>
</tr>
</tbody>
</table>

* indicates a statistically significant association

International Journal of Sports Physical Therapy
Figure 1. The most significant factor why fastpitch softball coaches don’t implement injury prevention programs

Figure 2. Fastpitch Softball coaches’ primary goal of Injury Prevention Programming

Although the entire kinetic chain contributes to throwing performance and durability, much of the mechanical stress from throwing frequently does manifest in the shoulder and elbow of overhead athletes. Interestingly, the throwing shoulder of softball pitchers is subjected to similar distraction forces as seen in a baseball pitcher. However, softball pitchers have more relaxed pitch counts and recovery guidelines compared to baseball players. Current research has identified a direct relationship between seasonal throwing volume and development of arm pain in high school softball pitchers. In the current survey, less than half of surveyed coaches (45.4%, n=78/172) were in favor of standardized pitch count guidelines in softball, even though multiple studies have discussed the concept of overuse leading to increased incidence of softball injuries. Pitch count adherence is an emerging concept for softball coaches largely due to the biomechanical and tissue stresses of the windmill style pitch being understudied. This could explain why the majority of participating coaches in this survey did not recognize pitch counts as a strong injury prevention priority.

Among the responding coaches who implement injury prevention, most utilize group or team-based programs instead of individualized programs. This is consistent with available literature which has explored injury prevention programs for softball players. Individualized prevention programs can be more challenging to implement and requires frequent testing to establish the athlete's baseline strength, range of motion, and balance. Likewise, 90% (n=83/92) of responding coaches reported that pitchers and position player should perform similar injury prevention programs throughout the season. The softball pitch requires unique muscular demands and joint motions of the upper and lower body which suggests that softball
pitchers could benefit from a more individualized injury prevention program which differs from programming for position players. For example, the overhand thrower needs to absorb the vertical force in the stride leg while the windmill pitcher needs to post-up on the stride foot to transfer force into the ball.25 As a result, softball coaches may need to assume a more active role in developing specific injury prevention programs based on the unique demands of the player’s position and physical function. In the current survey, 50% (n=80/172) of coaches indicated that the players have the biggest role in overall player health and preventing injuries. In comparison to a similar survey by Matsel et al.,18 57% of baseball coaches believed they played the largest role in player health, while only 38% of coaches responded that athletes are the most responsible. These two studies highlight the self-advocacy and autonomy of coaches and athletes, demonstrating they can reduce injury rates without dependence on parents, teammates, and other medical providers (physical therapists, athletic trainers, physicians).

As more research surfaces about fastpitch softball and the increasing injury rates, better educational accessibility and applicability for coaches and players is warranted. Those who design injury prevention routines should have access to the knowledge needed to create programs focused on targeting body regions known to contribute to injury. An apparent gap exists between the recommendations established in the medical literature and the practical utilization of injury prevention programs among softball coaches. In the future, better collaboration between rehabilitation professionals and softball coaches could result in more effective programming to mitigate the growing injury rates in fastpitch softball. Future research should explore common physical risk factors which develop throughout the softball season as a basis for informing a comprehensive injury prevention exercise program.

LIMITATIONS
The results of this survey should be interpreted conservatively as the authors recognize some significant limitations. First, the sample size of the responding coaches who completed the survey was smaller than anticipated resulting in underpowered results. The authors attempted to reach as many FPS coaches throughout the United States as possible through access to mailing lists from professional organizations and private databases, but many of these mailing lists may have been outdated with discontinued emails or coaches who are no longer actively engaged in coaching. The total number of active FPS coaches is difficult to estimate but the small response rate of 1.2 percent may not be reflective of the perspectives of the entire population of softball coaches. Second, the authors used email to survey a geographically diverse sample of FPS coaches throughout the United States. However, technological restrictions such as limited access to a computer or the internet may have resulted in a selection bias. Finally, there were a limited number of college-level coaches who responded to the survey. The authors sought to survey FPS coaches from both college and high school levels to determine similarities and differences, however, college coaches only represented 15% of the sample thus limiting the conclusions.

CONCLUSION
The results of this survey suggest that less than 50% of responding fastpitch softball coaches implement injury prevention programs to prevent injury. Limited familiarity with effective program design, inadequate staffing, and inconsistent risk factor awareness are the major contributors to lacking implementation. Educational collaboration between rehabilitation professionals and softball coaches regarding preventative programming strategies and injury risk factor management should be considered.
ACKNOWLEDGMENTS

The authors would like to thank the University of Evansville Department of Physical Therapy and the Ridgeway Student Research Endowment for funding this project.

Submitted: February 11, 2023 CDT, Accepted: May 24, 2023 CDT
REFERENCES


doi:10.1177/2325967119867426


SUPPLEMENTARY MATERIALS

Appendix 1

Original Research

Reliability Analysis of In-person and Virtual Goniometric Measurements of the Upper Extremity

Tracy Spigelman1,2, Leah Simpkins1, Casey Humphrey3, Yehor Vitel4, Aaron Sciascia4

1 Parks and Recreation Exercise and Sports Science, Eastern Kentucky University, 2 Department of Occupational Science and Occupational Therapy, Eastern Kentucky University, 3 Exercise and Sport Science, Eastern Kentucky University, 4 Institute for Clinical Outcomes and Research, Lexington Clinic

Keywords: goniometer, range of motion, reliability, virtual

https://doi.org/10.26603/001c.81065

International Journal of Sports Physical Therapy

Background

Virtual healthcare has forced clinicians to modify or eliminate parts of the musculoskeletal evaluation such as motion assessment. Although acceptable to excellent levels of in-person goniometric reliability is achievable, reliability of virtual assessments is unknown.

Purpose

To determine if similar upper extremity goniometric measurements could be obtained in-person and virtually.

Study Design

Reliability study; classroom setting

Methods

Publicly recruited sample over 18 years of age with no upper extremity injuries. Each subject was tested in a standing position with dominant arm facing the clinicians to visualize the landmarks for goniometer placement. Flexion and extension of the shoulder, elbow and wrist were measured. Prior to performing in-person goniometric measurements for each joint, an image was captured of each pre-determined joint position using a mobile device with a camera. This image represented the screenshot on a virtual platform. Four clinicians performed in-person measurements twice during the same session on each subject. The following week clinicians measured virtual images using the same techniques. Inter-rater and intra-rater reliability were determined via Intraclass correlation coefficients (ICC).

Results

Inter-rater reliability for five of the six in-person (ICC>0.81) and virtual measurements (ICC>0.78) were classified as excellent. In-person wrist extension (ICC=0.60) and virtual wrist flexion (ICC=0.65) were classified as good. Intra-rater reliability for individual clinicians were between good and excellent for the in-person measurements (ICC=0.61-0.96) and virtual measurements (ICC=0.72-0.97). There were a greater number of excellent ICC values for the virtual measurements (90%) compared to in-person measurements (70%). There were statistically significant differences between in-person and virtual sessions for five of six measurements (p<0.006). Only elbow extension did not differ between sessions (p=0.966).

Corresponding author:

Tracy Spigelman, PhD, ATC
Department of Parks and Recreation Exercise and Sports Science,
Eastern Kentucky University, Richmond,
Kentucky 40457, USA
Phone 859-622-2134
Email Tracy.Spigelman@eku.edu
Conclusion

Virtual assessment compared to goniometric measurements showed good to excellent inter- and intra-rater reliabilities (ICC > 0.60), which suggests clinicians can utilize goniometry either in person or on a virtual platform.

Level of Evidence

3b
©The Author(s)

INTRODUCTION

One of the primary components of the musculoskeletal physical examination is the assessment of joint and limb motion. Clinicians can obtain the amount of joint motion a person has with a variety of devices including three-dimensional analysis motion tracking systems, manual and digital inclinometers, and most commonly, manual goniometers. Also, clinicians have often been shown to have acceptable to excellent levels of inter-rater and intra-rater reliability when utilizing these devices in-person.1–4

The rise of virtual healthcare visits (often termed telehealth or telemedicine) has forced clinicians to either modify, or in some cases eliminate, components of the physical examination due to the logistics and barriers of administering a virtual visit.5–7 When using a virtual platform to administer a telehealth examination, motion assessments have either been eliminated or reduced to visual qualitative assessments. For example, a patient may be asked to elevate the arm in front of the body to determine how much shoulder flexion is able to be performed. However, without the use of a measuring device, clinicians would be relegated to using visible landmarks or categories to document the motion achieved i.e., patient was able to elevate arm to just below the ear. This raises two concerns: 1) the qualitative nature of the modified assessment is subjective and not exact and 2) previous literature has demonstrated that motion assessment reliability is more consistent with instrumentation compared to only using visual means.8 Furthermore, it has been demonstrated that patients tend to over-estimate the amount of motion they can perform when verbally asked to quantify their joint motion suggesting the elimination of objective motion assessments should not be considered.9 The non-objective assessments could in turn negatively affect clinical decision making for providing an accurate diagnosis, determining proper treatment, and properly monitoring patient progress across treatment.

Although it is possible these types of issues could arise during a telehealth examination, it is also possible that joint motion could be assessed with acceptable reliability in the virtual environments by utilizing a simple image capturing technique such as a screen shot and goniometer during the telehealth session. Therefore, the purpose of this study was to determine if similar goniometric measurements of the upper extremity could be obtained in-person and virtually. It was hypothesized that inter-rater and intra-rater test/re-test reliability for both in-person and virtual measurements would reach an acceptable level of reliability defined as an intraclass correlation coefficient ≥0.60.

METHODS

SUBJECTS

A publicly recruited sample of subjects volunteered to participate in this study. Inclusion criteria included: age between 18–60 years; able to actively elevate the arm to ear level (approximately 150°); actively move the elbow from an extended to flexed position within an approximate range of 0°–90°; and actively move the wrist into a flexed and extended position from a starting position of neutral (0°) to a non-specific range of motion in both directions. Subjects were excluded if age was <18 years and >60 years, could not move the shoulder, elbow, and wrist as noted in the inclusion criteria, had a Disabilities of the Arm, Shoulder, and Hand (DASH)10 disability score ≥40%,11 or had neurological compromise that would prevent joint/limb motion from occurring.

PROCEDURES

After reading and signing the informed consent packet, demographic information including age, sex, height, weight, and arm dominance were obtained. Following completion of the demographic obtainment, subjects completed the DASH.10

IN-PERSON MEASUREMENTS

Prior to performing the in-person goniometric measurements for each joint, an image was captured of each pre-determined joint position using a mobile device with a camera (iPad Air 2, Apple, Inc, Cupertino, CA). This still shot image represented an image that could be captured via screenshot on a virtual platform. Next, serial in-person measurements were obtained by each of four clinician research team members. The clinicians were comprised of two certified athletic trainers and two occupational therapists. All clinicians had a minimum of 10 years of clinical experience. Each clinician member of the research team performed all six measurements on each subject consecutively. This process continued until all team members performed all measurements twice in the same session. This was necessary for determining the test/re-test reliability for each clinician (intra-rater reliability). The goniometer dial was covered with paper so the team member obtaining the measurement could not see the values. To reduce the potential for recording bias by the team member performing the measurements, a second team member read and recorded the range of motion to the nearest degree mark.
The dominant arm of each subject was utilized for all measurements unless the dominant arm did not meet the inclusion criteria. Each subject was tested in a standing position facing sideways with their dominant arm facing the camera. This was necessary for both the in-person and virtual assessments to clearly visualize the anatomical landmarks for goniometer placement (Table 1). The testing positions for each measurement occurred as follows: Shoulder flexion: humerus at 150° flexion (approximately ear level); Shoulder extension: humerus at maximal extension without altering erect trunk position; Elbow flexion: humerus in line with trunk, elbow at 90° flexion, and forearm in supination; Elbow extension: humerus in line with trunk, elbow at 0° extension, and forearm supinated; Wrist flexion: humerus in line with trunk, elbow at 90° flexion, forearm pronated, and wrist maximally flexed; and Wrist extension: humerus in line with trunk, elbow at 90° flexion, forearm pronated, and wrist maximally flexed.

VIRTUAL MEASUREMENTS

Approximately one week (7-10 days) after the in-person measurements were completed, the research team members measured the captured images using the same goniometric techniques. This step also utilized two team members, with one member performing the in-person measurement for each image and another member reading the goniometer (and vice versa). The images were placed on a cloud-based shared drive, for all team members to be able to access the images at each person’s personal computer. Team members were not permitted to alter the image characteristics (brightness, contrast, resolution, etc) but were permitted to use the zoom function contained within the computer’s image viewing software to enlarge each image for more accurate placement of the goniometer. The average of the two trials was calculated for both in-person and virtual sessions and used for statistical analysis. Intra-rater reliability for each joint measurement was determined for each clinician per each session (i.e. clinician #1 trial 1 versus trial 2, clinician #2 trial 1 versus trial 2, etc.) while inter-rater reliability for each joint measurement was determined by comparing all results for trial 1 versus trial 2 for all four clinicians combined for each session.

STATISTICAL ANALYSIS

Summary statistics for demographic items were calculated and reported as means and standard deviations for continuous variables and frequencies with percentages for categorical variables. The distribution of data for each variable was assessed for normality using the Shapiro-Wilk test. Using a two-way random with absolute agreement design for inter-rater (2, k) and intra-rater (2,1) test/re-test reliability, intraclass correlation coefficients (ICC) were calculated for both in-person and virtual testing sessions. Once the ICs were determined, standard error of measurement (SEM) and minimal detectable change at the 90% (MDC90) and 95% (MDC95) confidence level were calculated. An ICC greater than 0.75 was interpreted as excellent, 0.40-0.60 was good, 0.59-0.40 was fair, and <0.40 was considered poor. Finally, a between session comparison of measurement values was conducted using paired t-tests or Wilcoxon sign rank tests (based on normality results) for the overall comparisons (in-person versus virtual) and one-way analyses of variance with Bonferroni correction for between examiner comparisons.

Using previously established criteria for sample size estimation, it was determined that 20 subjects would be needed to achieve a minimum intraclass correlation coefficient of 0.60 at an alpha level of 0.05 and beta level of 0.90.

RESULTS

Twenty subjects (Age: 30.8±12.8 years; height: 169.8±10.2 centimeters; weight: 76.8±18.9 kilograms; DASH: 26.0±2.4%; Sex: 85% female) participated in the study.

INTER-RATER RELIABILITY

The ICs for five of the six in-person measurements were classified as excellent (ICC>0.81) (Table 2). In-person wrist extension was classified as good (ICC=0.60). Similarly, the ICs for five of the six virtual measurements were classified as excellent (ICC>0.78). Virtual wrist flexion was classified as good (ICC=0.65).

INTRA-RATER RELIABILITY

Overall, the ICs for the individual clinicians were between good and excellent for the in-person measurements (range: 0.61-0.96) and virtual measurements (range: 0.72-0.97) (Table 3). When examining the individual measurement results, the ICs for both in-person (ICC>0.84) and virtual (ICC>0.95) shoulder extension and in-person (ICC>0.89) and virtual (ICC>0.94) elbow extension were all classified as excellent. There were a greater number of excellent ICC val-

<table>
<thead>
<tr>
<th>Table 1. Goniometer Placement Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Axis</strong></td>
</tr>
<tr>
<td>Shoulder flexion/extension</td>
</tr>
<tr>
<td>Elbow flexion/extension</td>
</tr>
<tr>
<td>Wrist flexion/extension</td>
</tr>
</tbody>
</table>
ues for the virtual measurements (90%) compared to the in-person measurements (70%).

**BETWEEN SESSION COMPARISONS**

When combining all clinician measurements, there were statistically significant differences between in-person and virtual sessions for five of the six measurements (p≤0.006) (Table 4). Only the measurement of elbow extension did not differ between sessions (p=0.966).

**INTER-RATER RELIABILITY**

Upon review of the inter-rater reliability, Examiner 1 recorded significantly lower amounts of in-person shoulder flexion compared to Examiner 3 (p=0.010) and Examiner 4 (p<0.001) (Table 5). Similarly, Examiner 1 recorded significantly lower amounts of in-person wrist extension compared to the other three examiners (p<0.001). Regarding the virtual measurements, Examiner 4 recorded significantly greater amounts of shoulder flexion compared to Examiners 1 and 2 (p<0.001). Examiners 3 and 4 recorded significantly greater amounts of wrist flexion compared to Examiner 1 (p<0.010) while Examiner 3 also recorded significantly greater wrist flexion compared to Examiner 2 (p=0.005). Finally, Examiner 1 recorded significantly lower amounts of wrist extension compared to Examiners 2 (p=0.051) and 3 (p=0.025).

**DISCUSSION**

Clinicians routinely utilize range of motion measures to predict the development of and to diagnose certain pathologies as well as to determine function of a body part. Following COVID-19, virtual patient evaluations became more common raising concern about the inclusion of and reliability of motion measurements. This study aimed to determine the reliability of measuring virtual range of motion in the shoulder, elbow and wrist using a goniometer. Virtual assessment compared to in-person goniometric measurements showed good to excellent inter- and intra-rater reliabilities (ICC=0.60).

Past researchers have attempted to examine range of motion in the shoulder, elbow, and knee using methods such as radiographs, visual estimation, inclinometer, smart phone applications and goniometry. Blonna et al. reported excellent to good reliability between surgeons and physician assistants when comparing visual observation of elbow flexion/extension to goniometry but noted the highest ICC’s using a goniometer. Similarly, van de Pol et al. found a wide range of inter-rater reliability depending on the method utilized but concluded devices such as goniometers or inclinometers should be utilized over visual observation due to more consistent and higher ICC values. The current results agreed with these findings where the reliability metrics for all measurements ranged from good to excellent . When employing visual observation alone, Hickey et al. reported limited agreement to observe and define asymptomatic versus symptomatic scapular motion via video cassette tapes between experienced and novice clinicians. This suggests that more than a trained eye should be used for shoulder evaluations and although their study did not use a goniometer, it does support the need for a more quantitative form of measurement for shoulder evaluation in a virtual medium.

The most important finding of this study is that although both in-person and virtual measurements ranged from good to excellent test/re-test reliability, there were a higher number of excellent ICC’s for the virtual measurements. This is most likely due to the lack of movement between trials for the virtual measurements. These data suggest using a screen shot and goniometer during a virtual examination for assessing flexion and extension of the shoulder, elbow, and wrist. Recent literature, supporting the virtual examination, has focused on camera positioning and placement in addition to clothing to ensure the most accurate measures. These variables are causes for possible differences in the measurement values between in-person and virtual sessions of the current study. It is important to set up a consistent space for the clinician to perform the best evaluation of a patient, however, these results point to-

---

**Table 2. Inter-rater Reliability**

<table>
<thead>
<tr>
<th></th>
<th>ICC</th>
<th>95% CI</th>
<th>SEM (°)</th>
<th>MDC90 (°)</th>
<th>MDC95 (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IP</td>
<td>V</td>
<td>IP</td>
<td>V</td>
<td>IP</td>
</tr>
<tr>
<td>Shoulder Flexion</td>
<td>0.88</td>
<td>0.86</td>
<td>0.78,0.94</td>
<td>0.75,0.94</td>
<td>3.01</td>
</tr>
<tr>
<td>Shoulder Extension</td>
<td>0.96</td>
<td>0.95</td>
<td>0.93,0.98</td>
<td>0.92,0.98</td>
<td>1.86</td>
</tr>
<tr>
<td>Elbow Flexion</td>
<td>0.81</td>
<td>0.91</td>
<td>0.66,0.91</td>
<td>0.83,0.96</td>
<td>1.70</td>
</tr>
<tr>
<td>Elbow Extension</td>
<td>0.96</td>
<td>0.98</td>
<td>0.92,0.98</td>
<td>0.97,0.99</td>
<td>1.48</td>
</tr>
<tr>
<td>Wrist Flexion</td>
<td>0.89</td>
<td>0.65</td>
<td>0.81,0.95</td>
<td>0.36,0.84</td>
<td>2.49</td>
</tr>
<tr>
<td>Wrist Extension</td>
<td>0.60</td>
<td>0.78</td>
<td>0.27,0.82</td>
<td>0.59,0.90</td>
<td>6.20</td>
</tr>
</tbody>
</table>

ICC=intraclass correlation coefficient; 95%CI=95% confidence interval; SEM=standard error of measurement; MDC=minimal detectable change; IP=in-person; V=virtual; °=degrees
Table 3. Intra-rater Reliability

<table>
<thead>
<tr>
<th>Examiner</th>
<th>95% CI</th>
<th>SEM (°)</th>
<th>MDC90 (°)</th>
<th>MDC95 (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Examiner 1</td>
<td>0.80</td>
<td>0.94</td>
<td>0.49,0.92</td>
<td>0.84,0.97</td>
</tr>
<tr>
<td>Examiner 2</td>
<td>0.86</td>
<td>0.90</td>
<td>0.65,0.95</td>
<td>0.74,0.96</td>
</tr>
<tr>
<td>Examiner 3</td>
<td>0.70</td>
<td>0.94</td>
<td>0.22,0.88</td>
<td>0.85,0.98</td>
</tr>
<tr>
<td>Examiner 4</td>
<td>0.72</td>
<td>0.92</td>
<td>0.30,0.89</td>
<td>0.80,0.97</td>
</tr>
</tbody>
</table>

Table 4. Between Session Comparison of In-Person and Virtual Measurements Reported as Mean (Standard Deviation)

<table>
<thead>
<tr>
<th></th>
<th>In-Person</th>
<th>Virtual</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shoulder Flexion</td>
<td>146.8° (8.0°)</td>
<td>150.5° (9.6°)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Shoulder Extension</td>
<td>41.4° (8.9°)</td>
<td>45.5° (8.1°)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Elbow Flexion</td>
<td>83.4° (3.9°)</td>
<td>85.1° (4.0°)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Elbow Extension</td>
<td>0.6° (7.2°)</td>
<td>-0.2° (6.7°)</td>
<td>0.966</td>
</tr>
<tr>
<td>Wrist Flexion</td>
<td>72.8° (7.5°)</td>
<td>68.0° (9.5°)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Wrist Extension</td>
<td>64.8° (9.8°)</td>
<td>62.1° (8.0°)</td>
<td>0.006</td>
</tr>
</tbody>
</table>

LIMITATIONS

The findings of this study show virtual range of motion measures had a high level of ICC's (excellent) which suggests clinicians can obtain quantitative measurements even if the patient is not directly in front of the clinician. How-
Table 5. Between Examiner Comparisons for In-Person and Virtual Measurements Reported as Mean (Standard Deviation)

<table>
<thead>
<tr>
<th></th>
<th>Examiner 1</th>
<th>Examiner 2</th>
<th>Examiner 3</th>
<th>Examiner 4</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shoulder Flexion</td>
<td>140.9° (6.5°)</td>
<td>146.4° (7.8°)</td>
<td>148.2° (7.4°)</td>
<td>151.6° (6.9°)</td>
<td>1 &lt; 3 (p=0.01) and 4 (p&lt;0.001)</td>
</tr>
<tr>
<td>Shoulder Extension</td>
<td>39.4° (8.3°)</td>
<td>41.3° (8.4°)</td>
<td>43.5° (10.7°)</td>
<td>41.3° (8.1°)</td>
<td></td>
</tr>
<tr>
<td>Elbow Flexion</td>
<td>84.5° (3.6°)</td>
<td>81.9° (4.6°)</td>
<td>83.2° (3.7°)</td>
<td>84.2° (3.6°)</td>
<td></td>
</tr>
<tr>
<td>Elbow Extension</td>
<td>1.3° (6.5°)</td>
<td>1.7° (5.5°)</td>
<td>0.1° (9.0°)</td>
<td>-0.7° (7.6°)</td>
<td></td>
</tr>
<tr>
<td>Wrist Flexion</td>
<td>70.7° (7.6°)</td>
<td>71.6° (7.5°)</td>
<td>73.8° (7.7°)</td>
<td>75.2° (6.9°)</td>
<td></td>
</tr>
<tr>
<td>Wrist Extension</td>
<td>52.0° (9.0°)</td>
<td>71.5° (4.7°)</td>
<td>67.1° (5.6°)</td>
<td>68.8° (5.2°)</td>
<td>1 &lt; 2-4 (p&lt;0.001)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Examiner 1</th>
<th>Examiner 2</th>
<th>Examiner 3</th>
<th>Examiner 4</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shoulder Flexion</td>
<td>145.1° (8.4°)</td>
<td>146.4° (8.2°)</td>
<td>152.0° (8.7°)</td>
<td>158.5° (7.3°)</td>
<td>4 &gt; 1 and 2 (p&lt;0.001)</td>
</tr>
<tr>
<td>Shoulder Extension</td>
<td>47.7° (7.8°)</td>
<td>46.8° (7.3°)</td>
<td>42.7° (8.7°)</td>
<td>45.0° (8.3°)</td>
<td></td>
</tr>
<tr>
<td>Elbow Flexion</td>
<td>85.0° (4.3°)</td>
<td>83.8° (4.4°)</td>
<td>86.6° (3.0°)</td>
<td>85.0° (4.0°)</td>
<td></td>
</tr>
<tr>
<td>Elbow Extension</td>
<td>0.1° (6.8°)</td>
<td>-0.1° (6.1°)</td>
<td>0.0° (8.0°)</td>
<td>-0.8° (6.0°)</td>
<td></td>
</tr>
<tr>
<td>Wrist Flexion</td>
<td>61.7° (12.3°)</td>
<td>65.3° (6.7°)</td>
<td>74.5° (6.0°)</td>
<td>70.3° (6.8°)</td>
<td>3 &gt; 1 (p&lt;0.001) and 2 (p=0.005); 4 &gt; 1 (p=0.010)</td>
</tr>
<tr>
<td>Wrist Extension</td>
<td>57.2° (8.2°)</td>
<td>64.2° (7.5°)</td>
<td>64.4° (7.5°)</td>
<td>62.7° (7.3°)</td>
<td>1 &lt; 2 (p=0.031) and 3 (p=0.023)</td>
</tr>
</tbody>
</table>

However, there are limitations to discuss. First, differences found in the results of this study could be due to a few variables such as clothing differences, patient posture and patient joint position sense. The subjects who volunteered for this study were not instructed about the type of clothing to wear as they were a sample of convenience. Loose fitting blouses or patterns could have hindered the clinicians view of the joint making it difficult to find the same landmarks consistently. Likewise, a tight-fitting shirt could have hindered that patient’s ability to fully achieve range of motion. Second, posture could have played a role in differences of measurement. Subjects were not instructed to stand in anatomically correct (or ideal) posture, nor were any postural differences between subjects corrected for. For instance, a patient with forward rounded shoulders might have less range of motion than a patient whose posture is more anatomically correct or more ideal. However, for shoulder extension, subjects who noticeably altered trunk position to gain more extension were immediately told to remain in each person’s “typical” posture, but no other corrections were applied. Finally, joint position sense could have played into any difference in results. Still photos for all motions (shoulder flexion/extension, elbow flexion/extension and wrist flexion/extension) were all taken prior to range of motion testing. Conversely, subjects were asked to repeat the same motions multiple times for the in-person measurements which could have increased flexibility throughout the testing (or created fatigue due to repeated positioning for four examiners) and therefore changed the end position of motion as the patient progressed through the testing.

CONCLUSION

Measuring range of motion both in-person and virtually had good to excellent test/re-test reliability suggesting either method is acceptable to use clinically. Capturing screenshots during a virtual exam to measure range of motion is recommended and is supported by the higher percentage of ICCs being ranked as excellent for the virtual measurements. Using a goniometer can provide an objective component to assessment and diagnosis of upper extremity injuries in the virtual examination for more thorough and accurate clinical decision making.

CONFLICTS OF INTEREST

The authors report no conflicts of interest.

Submitted: February 02, 2023 CDT, Accepted: May 16, 2023 CDT

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-NC-4.0). View this license’s legal deed at https://creativecommons.org/licenses/by-nc/4.0 and legal code at https://creativecommons.org/licenses/by-nc/4.0/legalcode for more information.
REFERENCES


Normative Values of Isometric Shoulder Strength Among Healthy Adults

Helen Bradley¹, Lauren Pierpoint¹

¹ Howard Head Sports Medicine, Vail Health, ² Physical Medicine and Rehabilitation, University of Utah

Keywords: normative data, strength, shoulder, activity level, age

Background
Normative data is useful for comparing measured values of strength with population norms and can avoid the issues associated with limb symmetry index. The available normative shoulder strength values are limited by constraints on research designs and variability in subject groups which prevents this data being successfully extrapolated to the greater population.

Purpose
The purpose of this study was to establish normative isometric strength values for various movements of the shoulder that are specific to function and rotator cuff strength. A secondary goal of this study was to analyze the effect of age, gender, weight, height, activity level and arm dominance on shoulder strength.

Design
Observational cohort study

Methods
Subjects in four age groups (20-29, 30-39, 40-49, 50-59) were included in this study—200 males (40.0 ± 11.6 years, 179.1 ± 6.5 cm, 81 ± 13.0 kg) and 200 females (40.1 ± 11.5 years, 165.3 ± 7.4sm, 64.4 ± 11.6 kg). Bilateral isometric strength measurements were taken with a handheld dynamometer testing seven shoulder movements. Tables of normative strength data were constructed. Multivariate analyses were performed to analyze the effects of age, gender, weight, height and activity level on isometric shoulder strength.

Results
Men were stronger than women (p<0.001). Age was not associated with most strength measures with the exception of dominant arm abduction (p<0.004), non-dominant arm abduction (p=0.028) and non-dominant arm scapular plane abduction (p<0.004) which had a negative association with strength. Weight was positively associated with strength (p<0.001). Activity level was positively associated with all strength measures (p<0.05) except dominant sided abduction (p=0.056). There were no statistically significant differences between dominant and non-dominant sides.

Conclusion
This normative data may be useful to the clinician, as it permits a standard against which to compare shoulder strength for various age groups. Clinicians can have confidence that
the uninvolved limb, if symptom free, can be used as an adequate benchmark for strength measures.

Levels of Evidence
Level 3
©The Author(s)

INTRODUCTION

Strength of the shoulder complex musculature is necessary for activities of daily living, recreation and sport, and is an important physical parameter assessed by rehabilitation specialists. 1 Shoulder strength is often used as a benchmark for goal setting during rehabilitation, is helpful for evaluating progress, and is used to determine readiness to return to activity, work or sport. 2 The use of limb symmetry index (LSI) is often used in practice and is a beneficial way to compare limb strength, however it has several limitations including when an individual has bilateral pathologies or a limb dominance which can over- or underestimates strength values. 3 Normative data can be useful for comparing measured values of strength with population norms. Previous researchers have reported normative values for shoulder strength however some of this has been restricted to specialized groups of individuals, 4, 5, 6 conducted using small sample sizes or age ranges, 7-9 and tested limited shoulder movements. 7, 9 Activity level of subjects was not considered in some studies 6, 10, 11 which may influence strength. It is thought that an adequate balance in strength between ER (external rotators) and IR (internal rotators) helps maintain dynamic stabilization of the shoulder. 12 Studies that have investigated this metric typical involve throwing athletes with reported ER:IR ratios ranging from 0.72 to 1.42. 13-16 In elite swimmers this ratio is approximately 0.70 bilaterally. 17 There is a lack of information regarding what a normal ER:IR ratio is in a healthy population. Differences in normative strength outcomes can be affected by the type of test used to measure strength. Some studies used a “break test” instead of a “make test”, which registers higher strength values than the make test. 3, 8 During the break test, the examiner pushes the hand held dynamometer (HHD) against the subject until the subject’s maximal muscular effort is overcome and the subject is unable to maintain an isometrically held position, while the make test is characterized by the examiner holding the dynamometer stationary while the subject exerts a maximal force against the dynamometer and examiner. 18 Previous studies that have determined normative strength measurements of shoulder musculature included a limited number of testing positions, not testing all shoulder movements. 7-9 In other studies, movements that optimize recruitment of specific muscles, such as the belly press movement to isolate the subscapularis muscle, were not included. 19

Previous authors have reported normative values for shoulder strength using isokinetic equipment which provides valuable data but isokinetic dynamometers are not accessible for the majority of rehabilitation facilities. 9, 20 A HHD is an inexpensive tool that offers clinicians a means of objectively assessing muscle force production (as a measure of strength). Previous authors have shown that hand-held dynamometry used to measure shoulder strength is a reliable and valid tool. 21-25

The purpose of this study is to establish normative isometric strength values for various movements of the shoulder that are specific to function and rotator cuff strength. A secondary goal of this study is to analyze the effect of age, gender, weight, height, activity level and arm dominance on shoulder strength.

METHODS

This observational cohort study examined normative strength data using healthy subjects. Ethical approval was obtained from the Vail Health Institutional Review Board. Subjects were recruited using a convenience sample including hospital employees, local health fair participants, community events attendees, and via word of mouth. Subjects were included if they were free of any upper extremity impairments (discerned via health questionnaire and range of motion screen). They had to be able to stand for 30 minutes during testing and understand the instructions provided to them. They were excluded if they had prior history of shoulder surgery including clavicle, any radicular symptoms of the upper extremities, or pain in the shoulder, elbow or wrist in the preceding three months.

Subjects were recruited according to age group (20-29, 30-39, 40-49 and 50-59 years) with a total of 50 men and 50 women in each age group (400 subjects total). Each subject completed a health questionnaire, informed consent, and the Shoulder Activity Scale (SAS) to determine the subjects overall shoulder activity level and whether they participated in overhead or contact sports. Previous studies using the SAS have shown its reliability, validity and responsiveness. 24, 25

Maximum isometric strength was collected using a calibrated, MicroFET 2© handheld dynamometer device (Hogan Health Industries). Seven movements were used to test isometric strength of the dominant and non-dominant shoulder musculature of each subject. These movements included, external rotation at the side (ER90), internal rotation at the side (IR0), abduction at 90° of shoulder abduction (ABD), external rotation at 90° of shoulder abduction (ER90), internal rotation at 90° abduction (IR90), belly press (BP) (the subject was asked to push his or her hand against a solid surface, such as a goniometer or clipboard, held at their abdomen) and scapular plane abduction (SCAP) measured at 90° of shoulder elevation in the scapular plane, with neutral rotation (see Appendix 1). Subjects were asked to produce a five second maximal isometric 'make' contraction against the HHD. Three trials were completed bilaterally in each position, the average of the trials
was used for normative data and analysis. A five second break was given between each repetition, a 30 second break was given between movements and the testing order was randomized using a random number generator. Verbal encouragement was provided as the subject performed each isometric push.

Muscle force was normalized to body mass (Newtons of force/body mass in kg) for between subject comparisons of strength. Hurd and colleagues evaluated the effects of normalizing muscle strength using a spectrum of anthropometric parameters and concluded body weight was the most effective parameter. The ratio of ER:IR was calculated for all subjects.

STATISTICAL METHODS

Interrater and inrater reliability were calculated for all the strength tests across two raters among a sample of five subjects (outside of the 400 subjects in this study), using a two-way random effects model and is reported as the intraclass correlation coefficient (ICC). This was conducted to determine reliability of the testing protocol used in this study. Means with SD or 95% CI are presented along with $5^{th}$, $25^{th}$, $50^{th}$, $75^{th}$, and $95^{th}$ percentiles stratified by age decade (20-29, 30-39, 40-49, 50-59), arm dominance, and gender. Numbers and percentages are presented for categorical variables. Independent t-tests or chi square analysis were used for comparisons between groups (e.g., gender). Paired t-tests were used for comparisons between subjects (e.g., dominant versus nondominant shoulder). Multiple linear regression was performed for each shoulder position with covariates of interest chosen a priori. Predictor variables for modeling of strength included age, gender, weight, height, and activity level. Effect modification between gender and age and activity level (e.g., does the relationship between age and strength depend on activity level?) was assessed via the inclusion of statistical interaction terms. Separate models were run for dominant (DOM) and non-dominant (ND) arms. $R^2$ values for each model are reported. Results were considered statistically significant at $p<0.05$ or if 95% CIs did not contain 1.00. All analysis was performed in SAS V 9.4 (SAS Institute Inc., Cary, NC).

RESULTS

Interrater reliability for strength measurements was good to excellent for most arm positions (ICC range 0.828-0.958) and moderate for ND ER90 (0.545), DOM ER90 (0.556), and DOM SCAP (0.694). For Rater 1, inrater reliability was good to excellent for most arm positions (ICC range 0.762-0.990) with moderate reliability for ND ER90 (0.597) and poor for DOM ER90 (0.409). For Rater 2, most ICCs were also good to excellent (ICC range 0.766-0.974) with moderate agreement for DOM SCAP (0.624) and ND SCAP (0.552) and poor for DOM ABD (0.480). Inter-rater and inrater reliability coefficients are presented in Table 1.

Participant demographics are shown in Table 2. Males had significantly higher activity level scores compared to females ($p<0.001$) and were more likely to participate in contact sports at higher levels ($p=0.002$). BMI increased with age ($p=0.01$), and younger participants were more likely to participate in contact and overhead sports at higher levels ($p=0.005$ and 0.001, respectively). Ten point three percent of the subjects tested were left hand dominant.

Table 3 displays normative data for mean shoulder strength for each of the movements, stratified by gender and arm dominance. For all muscle tests males had higher strength values than females ($p<0.0001$). Strength between DOM and ND limbs was not significantly different for most positions except for IR0 ($p=0.03$) and IR90 ($p=0.04$) for females and IR90 for males ($p=0.05$), in which the DOM arm was stronger.

The results of the general linear model showed that age, gender, weight, height and activity level explain 37-55% of the variation in shoulder strength (range of $r^2$ values in Table 4). Males were stronger than females. There was not a significant decline in strength with increasing age for most shoulder movements except ABD and SCAP.

Activity level was positively associated with strength such that those with higher activity scores had greater strength ($p=0.000$ to 0.056 depending on movement tested). Although interactions between gender and age were tested, they were not significant and therefore not included in the results. The interaction between age and activity level was significant for some arm positions (ER90 and IR0 both DOM and ND) (see Table 4). ER:IR strength ratios were higher in females compared to males (DOM $p=0.002$; ND $p=0.001$) and ER:IR ratio significantly declined with age (DOM $p=0.015$; ND $p=0.002$). Ratios are presented in Table 5.

<table>
<thead>
<tr>
<th>Shoulder position</th>
<th>Inter-rater reliability</th>
<th>Intra-rater reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Rater 1</td>
</tr>
<tr>
<td>DOM ABD</td>
<td>0.955</td>
<td>0.762</td>
</tr>
<tr>
<td>ND ABD</td>
<td>0.970</td>
<td>0.984</td>
</tr>
<tr>
<td>DOM BP</td>
<td>0.908</td>
<td>0.986</td>
</tr>
<tr>
<td>ND BP</td>
<td>0.920</td>
<td>0.985</td>
</tr>
<tr>
<td>DOM ER0</td>
<td>0.943</td>
<td>0.901</td>
</tr>
<tr>
<td>ND ER0</td>
<td>0.921</td>
<td>0.910</td>
</tr>
<tr>
<td>DOM ER90</td>
<td>0.556</td>
<td>0.409</td>
</tr>
<tr>
<td>ND ER90</td>
<td>0.545</td>
<td>0.597</td>
</tr>
<tr>
<td>DOM SCAP</td>
<td>0.694</td>
<td>0.938</td>
</tr>
<tr>
<td>ND SCAP</td>
<td>0.897</td>
<td>0.941</td>
</tr>
<tr>
<td>DOM IR0</td>
<td>0.879</td>
<td>0.935</td>
</tr>
<tr>
<td>ND IR0</td>
<td>0.953</td>
<td>0.932</td>
</tr>
<tr>
<td>DOM IR90</td>
<td>0.828</td>
<td>0.899</td>
</tr>
<tr>
<td>ND IR90</td>
<td>0.958</td>
<td>0.990</td>
</tr>
</tbody>
</table>

ICC = intraclass correlation coefficient; DOM = dominant arm; ND = Non-dominant arm; ABD = abduction; BP = belly press; ER = external rotation; SCAP = scapular plane abduction; IR = internal rotation
Table 2. Participant demographics and characteristics by gender and age decade

<table>
<thead>
<tr>
<th>Variable</th>
<th>Overall Sample</th>
<th>20-29 years</th>
<th>30-39 years</th>
<th>40-49 years</th>
<th>50-59 years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total (n=400)</td>
<td>M (n=200)</td>
<td>F (n=200)</td>
<td>M (n=50)</td>
<td>F (n=50)</td>
</tr>
<tr>
<td>Age, years [mean (SD)]**</td>
<td>40.0 (11.5)</td>
<td>40.0 (11.6)</td>
<td>40.1 (11.5)</td>
<td>25.5 (2.6)</td>
<td>25.6 (2.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>34.6 (3.1)</td>
<td>34.5 (3.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>44.6 (4.1)</td>
<td>45.3 (2.9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>55.2 (3.1)</td>
<td>55.1 (3.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Height, cm [mean (SD)]*</td>
<td>172.2 (9.6)</td>
<td>179.1 (6.5)</td>
<td>165.3 (7.4)</td>
<td>180.6 (7.7)</td>
<td>165.8 (8.4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>179.1 (6.3)</td>
<td>167.3 (6.6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>178.5 (5.7)</td>
<td>164.6 (6.4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>178.1 (6.0)</td>
<td>163.7 (7.7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>Weight, kg [mean (SD)]</td>
<td>72.7 (14.8)</td>
<td>81.0 (13.0)</td>
<td>64.4 (11.6)</td>
<td>77.8 (8.0)</td>
<td>63.3 (11.4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>80.3 (14.4)</td>
<td>65.8 (11.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>82.3 (14.2)</td>
<td>63.7 (11.9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>83.6 (14.0)</td>
<td>64.8 (12.2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>BMI, m/kg² [mean (SD)]*</td>
<td>24.4 (4.1)</td>
<td>25.3 (3.8)</td>
<td>23.4 (4.3)</td>
<td>23.9 (2.2)</td>
<td>23.0 (3.9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25.0 (3.8)</td>
<td>23.5 (3.9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25.8 (4.1)</td>
<td>23.5 (4.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>26.4 (4.3)</td>
<td>24.3 (5.2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>Arm length, cm [mean (SD)]**</td>
<td>31.5 (2.8)</td>
<td>32.7 (2.5)</td>
<td>30.4 (2.4)</td>
<td>33.4 (4.4)</td>
<td>30.0 (3.3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>32.7 (2.4)</td>
<td>30.1 (2.3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>32.6 (2.1)</td>
<td>30.4 (2.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>32.7 (2.8)</td>
<td>31.4 (4.3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>Activity Score [mean (SD)]**</td>
<td>12.4 (4.1)</td>
<td>13.5 (4.1)</td>
<td>11.4 (3.8)</td>
<td>13.8 (3.9)</td>
<td>10.6 (3.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14.3 (4.4)</td>
<td>12.1 (4.1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12.9 (4.4)</td>
<td>11.8 (3.4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13.0 (3.7)</td>
<td>10.9 (4.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td>Dominant hand (n, % left)</td>
<td>41 (10.3)</td>
<td>25 (6.3)</td>
<td>16 (4.0)</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participated in contact sports [n (%)]</td>
<td>3 (0.8)</td>
<td>1 (0.5)</td>
<td>2 (1.0)</td>
<td>0.002</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>Unorganized</td>
<td>303 (75.8)</td>
<td>137 (68.5)</td>
<td>166 (83.0)</td>
<td>30 (60.0)</td>
<td>38 (76.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>27 (54.0)</td>
<td>40 (80.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>38 (76.0)</td>
<td>42 (84.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>42 (84.0)</td>
<td>46 (92.0)</td>
</tr>
<tr>
<td>Organized</td>
<td>48 (12.0)</td>
<td>29 (14.5)</td>
<td>19 (9.5)</td>
<td>10 (20.0)</td>
<td>7 (14.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>10 (20.0)</td>
<td>6 (12.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5 (10.0)</td>
<td>3 (6.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4 (8.0)</td>
<td>3 (6.0)</td>
</tr>
<tr>
<td>Professional</td>
<td>46 (11.5)</td>
<td>33 (16.5)</td>
<td>13 (6.5)</td>
<td>10 (20.0)</td>
<td>5 (10.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13 (26.0)</td>
<td>4 (8.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7 (14.0)</td>
<td>3 (6.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3 (6.0)</td>
<td>1 (2.0)</td>
</tr>
<tr>
<td>Participated in overhand throwing/serving/swimming sports [n (%)]</td>
<td>3 (0.8)</td>
<td>1 (0.5)</td>
<td>2 (1.0)</td>
<td>0.2</td>
<td>0 (0.0)</td>
</tr>
<tr>
<td>Unorganized</td>
<td>273 (68.3)</td>
<td>128 (64.0)</td>
<td>145 (72.5)</td>
<td>30 (60.0)</td>
<td>32 (64.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>27 (54.0)</td>
<td>34 (68.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>35 (70.0)</td>
<td>38 (76.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>36 (72.0)</td>
<td>41 (82.0)</td>
</tr>
<tr>
<td>Organized</td>
<td>90 (22.5)</td>
<td>53 (26.5)</td>
<td>37 (18.5)</td>
<td>14 (28.0)</td>
<td>15 (30.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>16 (32.0)</td>
<td>6 (12.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12 (24.0)</td>
<td>8 (16.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11 (22.0)</td>
<td>8 (16.0)</td>
</tr>
<tr>
<td>Professional</td>
<td>34 (8.5)</td>
<td>18 (9.0)</td>
<td>16 (8.0)</td>
<td>6 (12.0)</td>
<td>3 (6.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12 (24.0)</td>
<td>10 (20.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3 (6.0)</td>
<td>2 (4.0)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2 (4.0)</td>
<td>1 (2.0)</td>
</tr>
</tbody>
</table>

*arm length missing for 7 subjects (1F, 6M)
**p value compares males and females
* *p value corresponds to change in demographic or survey response with age category, accounting for sex.
5. Gender based percentiles for shoulder strength are presented for each decade in Appendix 2.

**DISCUSSION**

The primary purpose of this study was to establish normative strength data for relevant shoulder movements, across a broad age range, in individuals with healthy shoulders. Variables including age, gender, weight, height, activity level and limb dominance were evaluated to see how they related to shoulder strength. The most important findings of the present study were that gender and activity level were significantly associated with strength measures. Age and limb dominance were not associated with most strength measures.

**SHOULDER STRENGTH AND GENDER**

The outcome that shoulder strength is affected by gender is in agreement with other studies. Specifically, males exhibited greater muscle strength than females in all tested shoulder positions even when normalized to body weight. Comparisons of muscle strength between individuals necessitates data to be normalized in order for valid comparisons to occur. Hurd et al. showed body weight was the most effective scaling factor in terms of reducing variability. It is suspected that the findings of the current study are due to differences in muscle morphology (men have larger muscle fibers and longer fascicles than women) and differences in muscle mass distribution between males and females (men have a greater total skeletal muscle mass in the upper body compared to females).

**SHOULDER STRENGTH BETWEEN LIMBS**

Across all strength measures, male and female subjects demonstrated no significant difference between sides except in IR0 and IR90, with the DOM limb being significantly stronger. Westrick et al. also found similar differences between DOM and ND strength of IR while Riemann et al. reported stronger dominant IR in healthy subjects aged 20–40 years. The results from these studies differ with findings from normative strength studies which have shown a difference between limb dominance and strength. Those studies also used healthy volunteers, but had fewer subjects per age group which could explain the differences in data. The current study showed isometric strength measures were not statistically significantly different between limbs for most movements. Therefore, clinicians can have some confidence that the uninvolved limb, if symptom free, can be used as an adequate benchmark for strength mea-

### Table 3. Strength (Newtons) normalized to body weight (kg) by sex and arm dominance

<table>
<thead>
<tr>
<th>Strength measurement</th>
<th>Female</th>
<th>Male</th>
<th><strong>p-value</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>DOM ABD</td>
<td>1.76</td>
<td>0.44</td>
<td>2.18</td>
</tr>
<tr>
<td>ND ABD</td>
<td>1.78</td>
<td>0.44</td>
<td>2.18</td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td>0.19</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>DOM BP</td>
<td>0.87</td>
<td>0.21</td>
<td>1.12</td>
</tr>
<tr>
<td>ND BP</td>
<td>0.86</td>
<td>0.2</td>
<td>1.12</td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td>0.64</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>DOM ER 0°</td>
<td>1.31</td>
<td>0.28</td>
<td>1.53</td>
</tr>
<tr>
<td>ND ER 0°</td>
<td>1.32</td>
<td>0.28</td>
<td>1.55</td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td>0.58</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>DOM ER 90°</td>
<td>0.92</td>
<td>0.25</td>
<td>1.04</td>
</tr>
<tr>
<td>ND ER 90°</td>
<td>0.91</td>
<td>0.24</td>
<td>1.04</td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td>0.16</td>
<td>0.4</td>
<td></td>
</tr>
<tr>
<td>DOM SCAP</td>
<td>1.81</td>
<td>0.46</td>
<td>2.2</td>
</tr>
<tr>
<td>ND SCAP</td>
<td>1.79</td>
<td>0.45</td>
<td>2.18</td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td>0.35</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>DOM IR 0°</td>
<td>1.47</td>
<td>0.36</td>
<td>1.76</td>
</tr>
<tr>
<td>ND IR 0°</td>
<td>1.44</td>
<td>0.4</td>
<td>1.78</td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td>0.03²</td>
<td>0.22</td>
<td></td>
</tr>
<tr>
<td>DOM IR 90°</td>
<td>1.18</td>
<td>0.3</td>
<td>1.43</td>
</tr>
<tr>
<td>ND IR 90°</td>
<td>1.16</td>
<td>0.31</td>
<td>1.4</td>
</tr>
<tr>
<td><strong>p-value</strong></td>
<td>0.04³</td>
<td>0.05³</td>
<td></td>
</tr>
</tbody>
</table>

SD = standard deviation; DOM=dominant arm; ND = Non-dominant arm; ABD = abduction; BP = belly press; ER = external rotation; SCAP = scapular plane abduction; IR = internal rotation
*Row p-values compare dominant to non-dominant arms within sex, **columns p-values compare males and females for each strength measurements.

²Statistically significant difference, p<0.05
Table 4. Results of the multivariate linear regression modelling shoulder strength.

<table>
<thead>
<tr>
<th>Strength measurement (Newtons)</th>
<th>Intercept</th>
<th>SAS (Female vs male)</th>
<th>Age (years)</th>
<th>Height (cm)</th>
<th>Weight (kg)</th>
<th>Age x activity level</th>
<th>R&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>DOM ABD</td>
<td>38.004</td>
<td>0.827</td>
<td>-43.005</td>
<td>-0.437</td>
<td>0.491</td>
<td>0.671</td>
<td>-</td>
</tr>
<tr>
<td>p value</td>
<td>0.414</td>
<td>0.056</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.004&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.062</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-</td>
</tr>
<tr>
<td>ND ABD</td>
<td>31.749</td>
<td>1.645</td>
<td>-40.095</td>
<td>-0.338</td>
<td>0.403</td>
<td>0.766</td>
<td>-</td>
</tr>
<tr>
<td>p value</td>
<td>0.499</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.028&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.129</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-</td>
</tr>
<tr>
<td>DOM BP</td>
<td>49.171</td>
<td>0.629</td>
<td>-25.745</td>
<td>-0.016</td>
<td>-0.035</td>
<td>0.482</td>
<td>-</td>
</tr>
<tr>
<td>p value</td>
<td>0.051</td>
<td>0.007&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.843</td>
<td>0.805</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-</td>
</tr>
<tr>
<td>ND BP</td>
<td>67.888</td>
<td>0.519</td>
<td>-28.693</td>
<td>0.013</td>
<td>-0.170</td>
<td>0.551</td>
<td>-</td>
</tr>
<tr>
<td>p value</td>
<td>0.004&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.017&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.862</td>
<td>0.195</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-</td>
</tr>
<tr>
<td>DOM ER0</td>
<td>26.792</td>
<td>0.732</td>
<td>-24.492</td>
<td>-0.051</td>
<td>0.194</td>
<td>0.660</td>
<td>-</td>
</tr>
<tr>
<td>p value</td>
<td>0.335</td>
<td>0.000&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.577</td>
<td>0.217</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-</td>
</tr>
<tr>
<td>ND ER0</td>
<td>22.943</td>
<td>0.885</td>
<td>-25.086</td>
<td>-0.059</td>
<td>0.229</td>
<td>0.626</td>
<td>-</td>
</tr>
<tr>
<td>p value</td>
<td>0.403</td>
<td>0.001&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.514</td>
<td>0.141</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-</td>
</tr>
<tr>
<td>DOM ER90</td>
<td>44.554</td>
<td>0.675</td>
<td>-15.573</td>
<td>-0.131</td>
<td>-0.043</td>
<td>0.531</td>
<td>0.044</td>
</tr>
<tr>
<td>p value</td>
<td>0.091</td>
<td>0.006&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.127</td>
<td>0.775</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.035&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>ND ER90</td>
<td>22.064</td>
<td>0.706</td>
<td>-13.963</td>
<td>-0.061</td>
<td>0.052</td>
<td>0.552</td>
<td>0.046</td>
</tr>
<tr>
<td>p value</td>
<td>0.407</td>
<td>0.005&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.485</td>
<td>0.729</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.031&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>DOM SCAP</td>
<td>-6.310</td>
<td>1.410</td>
<td>-38.672</td>
<td>-0.318</td>
<td>0.672</td>
<td>0.690</td>
<td>-</td>
</tr>
<tr>
<td>p value</td>
<td>0.899</td>
<td>0.002&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.051</td>
<td>0.017&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-</td>
</tr>
<tr>
<td>ND SCAP</td>
<td>-7.828</td>
<td>1.484</td>
<td>-36.679</td>
<td>-0.324</td>
<td>0.622</td>
<td>0.790</td>
<td>-</td>
</tr>
<tr>
<td>p value</td>
<td>0.877</td>
<td>0.002&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.049&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.029&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-</td>
</tr>
<tr>
<td>DOM IR0</td>
<td>-47.536</td>
<td>1.386</td>
<td>-21.413</td>
<td>0.065</td>
<td>0.488</td>
<td>0.991</td>
<td>0.068</td>
</tr>
<tr>
<td>p value</td>
<td>0.276</td>
<td>0.001&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.646</td>
<td>0.048&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.049&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>ND IR0</td>
<td>-40.936</td>
<td>1.495</td>
<td>-27.756</td>
<td>0.227</td>
<td>0.476</td>
<td>0.855</td>
<td>0.074</td>
</tr>
<tr>
<td>p value</td>
<td>0.382</td>
<td>0.001&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.138</td>
<td>0.072</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.048&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>DOM IR90</td>
<td>-18.351</td>
<td>0.981</td>
<td>-18.658</td>
<td>0.098</td>
<td>0.198</td>
<td>1.004</td>
<td>-</td>
</tr>
<tr>
<td>p value</td>
<td>0.595</td>
<td>0.002&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.383</td>
<td>0.311</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-</td>
</tr>
<tr>
<td>ND IR90</td>
<td>-33.133</td>
<td>0.941</td>
<td>-17.404</td>
<td>0.198</td>
<td>0.263</td>
<td>0.969</td>
<td>-</td>
</tr>
<tr>
<td>p value</td>
<td>0.331</td>
<td>0.003&lt;sup&gt;a&lt;/sup&gt;</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.076</td>
<td>0.172</td>
<td>0.000&lt;sup&gt;b&lt;/sup&gt;</td>
<td>-</td>
</tr>
</tbody>
</table>

SAS = Shoulder Activity Scale; DOM = dominant arm; ND = non-dominant arm; ABD = abduction; BP = belly press; ER = External Rotation; SCAP = scapular plane abduction; IR = internal rotation.

*All models were statistically significant (p-value for F-statistic < 0.01).
P-values listed in table correspond to the beta coefficient for each variable. The interaction term of age x activity level was only included when the interaction term was significant (p < 0.05). Non-normalized strength was used as the dependent variable since weight was included as a predictor. For continuous variables, a positive Beta coefficient indicates a positive association between the variable and strength. For sex, a negative Beta coefficient indicates that females have lower strength than males. The interaction term indicates that the relationship between age and strength depends on activity level, such that higher levels of activity show higher levels of strength at older ages.

<sup>a</sup> Statistically significant difference p < 0.05
<sup>b</sup> Statistically significant difference p < 0.001

Table 5. External rotation/internal rotation ratios by gender and age

<table>
<thead>
<tr>
<th>Age</th>
<th>20-29</th>
<th>30-39</th>
<th>40-49</th>
<th>50-59</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td>Mean SD</td>
<td>Mean SD</td>
<td>Mean SD</td>
<td>Mean SD</td>
</tr>
<tr>
<td>DOM ER:IR 90</td>
<td>0.810.16</td>
<td>0.770.15</td>
<td>0.810.19</td>
<td>0.780.18</td>
</tr>
<tr>
<td>ND ER:IR 90</td>
<td>0.860.16</td>
<td>0.760.18</td>
<td>0.810.15</td>
<td>0.780.16</td>
</tr>
</tbody>
</table>

DOM = dominant arm; ND = non-dominant arm; ER:IR90 = external rotation/internal rotation at 90° abduction

*DOM arm p-value: Gender 0.002; Age 0.015; ND arm p-value: Gender 0.001; Age 0.002
sures, and the utilization of LSI may be of benefit when comparing shoulder strength, with the exception of IR.

Additionally, strengthening interventions for the musculature surrounding the shoulder joint should continue to focus on symmetrical strength performance.

ER:IR RATIO

An adequate balance between the strength of the ER’s and IR’s helps maintain dynamic stabilization of the shoulder.\textsuperscript{12} Ratios varied from 0.71 to 0.86 which are similar to those seen in other studies and provide good reference values for a healthy, active population.\textsuperscript{4,6,8} With age this ratio decreased, either due to a decreasing ER strength or increasing IR strength. The data from this study does not allow determination of which of these changes caused the decline, but it is suspected that a combination of changes may have occurred.

SHOULDER STRENGTH AND AGE

In this healthy population, age was not a predictor of strength for most muscle groups surrounding the shoulder, except for ABD (DOM p=0.004 and ND p=0.028) and SCAP (ND p=0.049). These were negative correlations, indicating older individuals were less strong in these movement patterns. Standing elevation in the scapular plane has been shown to be an optimal position to recruit the supraspinatus muscle for strength testing.\textsuperscript{30} Wickham et al.\textsuperscript{31} confirmed that the position for supraspinatus to reach peak muscle activity was 89° shoulder ABD. Given the positive association between asymptomatic rotator cuff tears and age,\textsuperscript{32} especially in individuals over 50 years,\textsuperscript{33} this outcome could be related to asymptomatic rotator cuff pathology. Further research is needed to confirm this possibility.

The finding that age did not affect strength measures, contradicts the study by Hughes et al.\textsuperscript{5} who showed age was negatively associated with all strength measures. Participants in their study were aged between 20-78 years. The current study evaluated strength in individuals up to 59 years, which may not have been an age that demonstrates significant age-related declines. Andrews et al.\textsuperscript{5} collected normative data on five movements of the shoulder in patients between the ages of 50-79 years old. Moderate to high correlations were found between isometric strength and height, weight and gender, and a weak but significant negative correlation between age. Studies of age and strength have reported conflicting results which could be related to how active the population involved in the study is.\textsuperscript{6,10,27} Age-related decreases in muscle mass from 30-50% have been described in both males and females between the ages of 40-80 years.\textsuperscript{34,35} The decrease in muscle mass is accompanied by at least an equal decrease in strength.\textsuperscript{36} However, there is evidence that age-associated atrophy and weakness can be slowed by staying active and exercising as is seen in Master athletes.\textsuperscript{35,38}

SHOULDER STRENGTH AND ACTIVITY LEVEL

There was a significant association between activity level and strength in all movements tested except dominant ABD. The fact that there were no significant differences in activity level between age groups (Table 2) could explain why minimal associations between strength and age were seen in the data. People who stay active as they age are able to maintain good shoulder strength into later life. Harlinger et al.\textsuperscript{27} saw no significant difference in strength in subjects of 20-64 years (except for a decline in external rotators in men). They interviewed participants to assess activity level and noted the majority of people tested did participate in regular exercise including swimming, weightlifting and physically demanding employment. In comparison when authors examined shoulder strength, age, and activity level, a weak or inconsistent relationship between activity level and strength and a significant regression of strength has been found with advancing age.\textsuperscript{5,9} These studies evaluated activity level by recording metabolic equivalent (MET) over a 24-hour period\textsuperscript{9} or getting subjects to grade their work and leisure activity level according to a four point ordinal activity scale.\textsuperscript{5} Neither measure was specific to the shoulder. The current study, which utilized the SAS, revealed that over 99% of individuals reported participating in contact sports and overhead, serving or swimming activities, indicating a very active population. The SAS has been validated in a healthy population.\textsuperscript{24} Studies have evaluated the SAS and age and shown that among subjects with no history of shoulder symptoms or treatment for a shoulder condition, the SAS decreases with age and is lower in women than men.\textsuperscript{39,40} The SAS was lower in women compared to men in this study, supporting previous studies, however decreases with age were not observed. This suggests the cohort used in the current study were more active than individuals seen in the general population.

SHOULDER STRENGTH, AGE AND ACTIVITY LEVEL

Older adults who were active showed less decline in strength across two arm positions (ER90 and IR0 both DOM and ND) compared to older adults who were not as active. This may indicate that staying active is important for maintaining strength, although this observation was not present in all shoulder movements. This may be due to the relatively active population comprising the study or indicate that this relationship was also movement-dependent.

LIMITATIONS

The primary limitation of this study was the potential for selection bias of the subjects. The subject population was drawn primarily from Colorado. Based on analysis of Center for Disease Control health data, Colorado is ranked highly for individuals who participate in physical activity so extrapolating this normative data to other regions that are less active could be difficult.\textsuperscript{41} Future studies should include recruitment of a wider range of subjects who are less physically active to reflect the population at large. The upper age range measured in this study was 50-59
years and this does not represent a senior demographic age group. Inclusion of individuals of 60 years and over would have added more information about strength changes with age. Additionally, this study used pain and a subject reported health history questionnaire to determine if an individual had a healthy shoulder. A physical examination which could identify asymptomatic rotator cuff pathology was not undertaken and could have helped verify that the study population was healthy prior to collecting normative strength values.

CONCLUSION

The findings of this study provide normative data regarding shoulder strength for a healthy, active population across four decades. Clinically this provides preliminary evidence that healthy individuals without injury have relatively symmetrical isometric strength regardless of limb dominance. This suggests that following shoulder injury or surgery, clinicians can have some confidence that the uninvolved limb can be used as an adequate benchmark for strength measures if the uninvolved arm is healthy. The normative data presented provides data regarding shoulder strength measures which are gender and age specific. It may assist in setting rehabilitation goals, monitoring progress in patients with shoulder injuries and allow clinicians to make better informed return to sport decisions. The data also highlights the positive association between activity level and strength; the more active an individual, the stronger they are, regardless of age.

CONFLICTS OF INTEREST

The authors report no conflicts of interest.

ACKNOWLEDGEMENTS

The assistance of the Physical Therapy Residents at Howard Head Sports Medicine for their time and dedication to data collection as well as Sofija Phillips, Thomas Olson, Dirk Kokmeyer and Eric Dube, their contributions to this project are gratefully acknowledged.

Submitted: January 19, 2023 CDT, Accepted: May 24, 2023 CDT

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-NC-4.0). View this license's legal deed at https://creativecommons.org/licenses/by-nc/4.0 and legal code at https://creativecommons.org/licenses/by-nc/4.0/legalcode for more information.
REFERENCES


SUPPLEMENTARY MATERIALS

Appendix 2

Appendix 1
Validity, Reliability, and Efficiency of a Standard Goniometer, Medical Inclinometer, and Builder’s Inclinometer

June Hanks1, Betsy Myers1

1 Physical Therapy, University of Tennessee at Chattanooga

Keywords: intrarater reliability, joint motion, goniometry, inclinometry, cost-effectiveness

International Journal of Sports Physical Therapy

Background
Joint range of motion (ROM) is an important assessment to aid diagnostic and clinical decision-making for persons with a wide variety of neuromusculoskeletal conditions. The current clinical standard for assessing ROM is the standard goniometer (SG).

Purpose
The purpose of this study was to investigate the validity, reliability and time required to assess ROM using a standard goniometer (SG), medical inclinometer (MI), and builder’s digital inclinometer (BI).

Study Design
Cross-sectional study.

Methods
Fifty participants with no current shoulder, elbow, or forearm pain limiting movement were assessed by a single tester. The tester measured three repetitions of passive forearm and shoulder rotation with an SG, MI, and BI. Device order was randomized. Time to complete assessment with each device was measured.

Results
BI and MI were significantly faster than the SG (p < 0.001) for all motions. Inclinometer measurements were more reliable (average ICC = 0.953 for MI and 0.919 for BI) than SG measurements (average ICC = 0.822). There was good correlation between MI and BI and mean differences between devices was less than 2°. Correlations between the SG and the inclinometers ranged from poor to fair and mean differences between devices was 4°.

Conclusion
The BI and MI were reliable for measuring forearm and shoulder rotation. The poor correlation between the SG and inclinometers indicates that clinicians should utilize the same device for testing. Because time can be a barrier to clinician assessment, the greater efficiency and reliability of inclinometers warrants consideration as the new measurement standard. Standard patient and inclinometer positioning is recommended to enhance reliability.

Level of Evidence
2

©The Author(s)
INTRODUCTION

Assessment of joint range of motion (ROM) aids diagnostic and clinical decision-making for persons with a wide variety of neuromusculoskeletal conditions. The current clinical standard for assessing ROM is the standard goniometer (SG), that costs about $10. Visual estimation is often used in clinical practice as a quick and free alternative to SG. However, studies of visual estimation have shown lower intrarater and interrater reliability than assessment using a SG.\(^1\)\(^2\) Although methodological differences challenge direct comparisons among studies, interrater reliability using the SG varies among studies and is weak.\(^1\)\(^3\)\(^-\)\(^8\) Use of a SG requires lining up the goniometer arms with standard bony landmarks. While some landmarks are superficial and easily found, others, such as the greater trochanter, are more challenging, with as much as 15 mm difference in location recorded between raters,\(^9\) leading to significant goniometric measurement errors. Maintaining SG alignment can be challenging in certain populations, such as in children with cerebral palsy\(^10\) or in people with orthopedic conditions.\(^3\) Additionally, keeping alignment generally requires two hands to reduce the risk of error.\(^11\)

The HALO® (HALO Medical Devices, Australia) is a hybrid between a goniometer and an inclinometer that utilizes a laser to intersect proximal and distal anatomical landmarks relative to the joint being measured. The HALO® device costs between $160 and $200.\(^12\) While the digitization of angular measurements removes one source of measurement error, the HALO® still requires proper alignment of the laser with proximal and distal landmarks. Intrarater and interrater reliability for the HALO® has been found to be similar or slightly less than the SG for shoulder rotation.\(^13\) Smart phone applications which utilize a magnetometer-based goniometer are also available for ROM assessment. Smart phone applications have greater intrarater and interrater reliability than the SG\(^14\) and they can be useful for patient self-assessment of ROM.\(^15\)\(^16\) However, smart phone applications have not been proven reliable across multiple joints or in all movements.\(^17\)\(^18\) They also require a dedicated clinic phone (or phones) to maintain proper infection control. The lack of consistent guidance regarding alignment may contribute to measurement error.\(^15\)

Bubble inclinometers are handheld devices used to assess ROM in rehabilitation settings. Bubble inclinometers have a fluid-filled circular face. The fluid is a combination of a colored fluid and clear fluid. The interface of these two fluids moves with gravity, and this movement is used to measure motion against a rotating 360° dial. Bubble inclinometers are typically used to measure spinal motion. This measurement requires the placement of two devices at specific landmarks, zeroing each device, having the patient move through a ROM, and then reading the end measurements once the fluid has stopped shifting within the device. Bubble inclinometers, therefore, require increased time, and, in the case of spinal motion, require the use of two hands. Inaccurate assessment using the bubble inclinometer includes misplacement of the inclinometer on the body part, failure to maintain constant pressure of the inclinometer against the body part and tilting of the device during reading of the face of the device.\(^19\)

The digital inclinometer is a handheld device placed against a body surface to measure angular position relative to the vertical or horizontal plane. The digital inclinometer does not require alignment with multiple reference points and requires only one hand, making it easier to use than the SG. Many digital inclinometers can store multiple measurements, improving efficiency by allowing multiple motions to be assessed without needing a break to record each motion. Unlike goniometry, there are no universally accepted procedures for using inclinometers to assess joint ROM. A digital inclinometer used in clinical and research settings, referred to as a medical inclinometer (MI), is expensive, with costs exceeding $400. A builder's digital inclinometer (BI) commonly used by construction workers operates under the same principle as a MI and is much less expensive, ranging in cost from $10 to $50.

PURPOSE

Studies examining the reliability and time efficiency of the SG, MI, and BI are not well represented in the literature. The purpose of this study was to investigate the validity, reliability and time required to assess ROM using a SG, MI, and BI. The research hypotheses were as follows: 1) measurements taken using the inclinometers would be faster than measurements taken with the goniometer; 2) inclinometer measurements would be more reliable than goniometric measurements; 3) measurements with the inclinometers would be highly correlated; 4) measurements taken with inclinometers would be correlated with goniometric measurements.

METHODS

Fifty healthy individuals (34 females and 16 males, 25 to 58 years of age) were recruited from a local university setting to participate in the study. To be eligible, participants were at least 18 years of age with no reported current shoulder, elbow, or forearm pain limiting movement. Participants wore short-sleeve shirts to allow identification of necessary bony landmarks for testing. The devices used for ROM assessment were SGs (360°, six-inch and 12-inch), MI (Acumar, Grayline Medical; Norwalk, CA), and BI (AccuRemote, San Clemente, CA) The study was approved by the Institutional Review Board of a public university. All participants provided written informed consent prior to participation.

All measurements were taken by the same tester with more than 30 years of clinical experience as a physical therapist. All measurements were performed according to the preferred position as described by Norkin and White.\(^20\) For this study, shoulder internal/external rotation was considered as glenohumeral internal/external rotation. A 12-inch SG was used for assessing shoulder internal/external rotation and 6-inch SG was used for assessing forearm pronation/supination. The inclinometers were placed on the dorsal surface of distal forearm for shoulder internal rotation.
shoulder internal rotation. The order of devices was randomly assigned for each participant. Immediately prior to assessment of each motion, the participant performed three repetitions of active ROM, followed by one repetition of passive ROM performed by the tester. With the participant in the standard test position, the participant performed a warm-up of three repetitions of active ROM, followed by one repetition of passive ROM performed by the tester to take the joint to end range of motion prior to passive ROM measurement. The warm-up was performed to minimize the increase in ROM with repeated measurements. The tester then positioned the participant in the neutral starting position for the given joint motion. For the first repetition with each device, the tester signaled the timer to begin timing the assessment and passively moved the participant to the end range of motion and aligned the measurement device. Once the device was properly aligned at the end of joint passive ROM, the tester signaled the timer to stop. The timer then read and recorded the amount of passive ROM as well as the time required to the nearest second.

The digital display on the MI ranged from 0–180° in 1° increments. The digital display of the BI ranged from 0.0° to 90.0° in 0.1° increments, with ROM beyond 90° displayed as −89.9° to −1.0°. For ROM greater than 90°, the recording researcher calculated the ROM (e.g. for a display reading of -80°, the recorded ROM was 100°). All information was recorded on a data recording form, transferred to a spreadsheet, and uploaded to the Statistical Package for the Social Sciences (SPSS) version 28.0 for analysis.

STATISTICAL ANALYSIS

The Shapiro-Wilk test indicated a non-normal data distribution; therefore, the Wilcoxon Signed Rank test with an alpha level of 0.05 was used to determine the difference in the time required for each motion with each device. Intrarater reliability for each device was calculated with the intraclass correlation coefficient (ICC) (3,k) two-way mixed model with absolute agreement using the three trials for each motion.21 Construct validity, determined by correlations between devices, was assessed using the ICC.22 Interpretation of correlation strength varies among authors.8,21–23 For this study, reliability was considered excellent for ICC values > 0.90; good for values 0.70–0.89; acceptable for values 0.60–0.69; fair for values 0.50–0.59 and poor for values <0.50.

The absolute reliability of each device was quantified using the standard error of measurement (SEM) and minimal detectable change (MDC).22 The SEM is a measure of how much test scores vary about a “true” score. The SEM was calculated using the formula: SEM = standard deviation * √(1-r).22 The MDC is the smallest change that can be considered a true change as opposed to change due to measurement error. The MDC was calculated using the formula: MDC = SEM * 1.96 * √2 to determine the magnitude of change that would exceed the 90% confidence interval.22 Given the normal distribution of the difference scores between devices, the level of agreement between the BI and MI was visualized using Bland-Altman plots with 95% lim-

Figure 1. Measurement of passive range of motion with inclinometers.
A. Forearm pronation with BI. B. Forearm supination with MI. C. Shoulder external rotation with BI. D. Shoulder internal rotation with MI. BI = builder’s inclinometer, MI = medical inclinometer.

and forearm pronation and on the ventral surface of the distal forearm for shoulder external and forearm supination (Figure 1).

The order of joint testing was as follows: forearm pronation, forearm supination, shoulder external rotation, and
Table 1. Descriptive Measurement Data for Range of Motion

<table>
<thead>
<tr>
<th></th>
<th>Goniometer</th>
<th></th>
<th></th>
<th>Builder</th>
<th></th>
<th></th>
<th>Medical</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Means±SD</td>
<td>SEM</td>
<td>MDC</td>
<td>Means±SD</td>
<td>SEM</td>
<td>MDC</td>
<td>Means±SD</td>
<td>SEM</td>
<td>MDC</td>
</tr>
<tr>
<td>Pro</td>
<td>80±6.5</td>
<td>3.1</td>
<td>2.5</td>
<td>87±7.9</td>
<td>2.8</td>
<td>2.2</td>
<td>89±8.8</td>
<td>2.8</td>
<td>2.2</td>
</tr>
<tr>
<td>Sup</td>
<td>85±5.3</td>
<td>2.7</td>
<td>2.1</td>
<td>89±8.3</td>
<td>2.9</td>
<td>2.3</td>
<td>89±9.3</td>
<td>3.2</td>
<td>2.6</td>
</tr>
<tr>
<td>ER</td>
<td>93±6.4</td>
<td>2.3</td>
<td>1.8</td>
<td>94±9.4</td>
<td>2.4</td>
<td>1.9</td>
<td>94±7.7</td>
<td>2.1</td>
<td>1.7</td>
</tr>
<tr>
<td>IR</td>
<td>67±9.0</td>
<td>3.5</td>
<td>2.8</td>
<td>66±10.8</td>
<td>2.5</td>
<td>2.0</td>
<td>62±10.7</td>
<td>2.3</td>
<td>1.9</td>
</tr>
</tbody>
</table>

All values are in degrees. Builder = builder’s inclinometer; medical = medical inclinometer; SEM = standard error of measurement; MDC = minimal detectable change; Pro = pronation; Sup = supination; ER = shoulder external rotation; IR = shoulder internal rotation.

Figure 2. Goni = goniometer; Builder = builder’s inclinometer; Medical = medical inclinometer; ER = external rotation; IR = internal rotation; * = goni significantly different than builder; † = goni significantly different than medical; ‡ = builder significantly different from medical.

The results of intrarater reliability with ICC and 95% confidence intervals for each device are listed in Table 2. For convenience, the SEM and MDC for each motion and each device have been included in Table 1. While all devices had good reliability (ICC > 0.70), SG measurements were less reliable than inclinometer measurements. Except for shoulder external rotation, the MI was the most reliable device for all motions, with all but one motion (supination) exceeding 0.90, equating to excellent reliability. While slightly better for the inclinometers, absolute reliability was good considering the low SEM across devices, ranging from 1.4° to 3.5°. The MDC and was low across devices, ranging from 1.7° to 2.8°. Therefore, the second research hypothesis was supported.

**RESULTS**

Table 1 provides the means and standard deviations of ROM measurements for each motion with each device. On average, the mean difference between the inclinometers was less than 2°. In contrast, the average mean differences between the SG and BI were 3° and 5° for the MI.

**RESEARCH HYPOTHESIS 1: TIME REQUIRED**

Figure 2 demonstrates the mean time required to perform each measurement with each device. The Wilcoxon Signed Rank test indicated that the BI and MI inclinometers required significantly less time than the goniometer (p < 0.001) for all motions. Assessment using the MI required significantly less time than the BI for both pronation and supination, p = 0.004 and p = 0.005, respectively. Therefore, the first research hypothesis was supported.

**RESEARCH HYPOTHESIS 2: INTRARATER RELIABILITY OF DEVICES**

As listed in Table 3, the correlations between the SG and BI ranged from a low of 0.434 for forearm pronation ROM to a high of 0.727 for shoulder internal rotation. The correlations between the SG and the MI were similarly low, ranging from 0.377 for forearm pronation to 0.647 for shoulder external rotation. Therefore, the fourth research hypothesis was not supported.

**DATA VISUALIZATION**

The Bland-Altman plots in Figure 3 and Figure 4 demonstrate the 95% limits of agreement between the BI and MI. More than 95% of measurements fall within these limits and are evenly distributed both above and below the mean throughout the ranges of motion. For pronation, supination, and shoulder external rotation, the mean differences between the BI and MI were less than 2°, with most measurements within ± 5°. For shoulder internal rotation, the mean differences were approximately 4°, with most measurements within ± 5°. 
Table 2. Intrarater Reliability of Devices

<table>
<thead>
<tr>
<th>Motion</th>
<th>Goni ICC (95% CI)</th>
<th>Builder ICC (95% CI)</th>
<th>Medical ICC (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pronation</td>
<td>0.770 (0.663 – 0.852)</td>
<td>0.878 (0.802 - 0.927)</td>
<td>0.901 (0.848 - 0.939)</td>
</tr>
<tr>
<td>Supination</td>
<td>0.748 (0.611 - 0.844)</td>
<td>0.876 (0.812 - 0.923)</td>
<td>0.878 (0.814 - 0.924)</td>
</tr>
<tr>
<td>Shoulder ER</td>
<td>0.870 (0.802 – 0.919)</td>
<td>0.933 (0.890 - 0.960)</td>
<td>0.925 (0.883 – 0.954)</td>
</tr>
<tr>
<td>Shoulder IR</td>
<td>0.849 (0.772 – 0.905)</td>
<td>0.946 (0.916 - 0.967)</td>
<td>0.952 (0.925 – 0.971)</td>
</tr>
</tbody>
</table>

Goni = goniometer; Builder = builder’s inclinometer; Medical = medical inclinometer; ER = external rotation; IR = internal rotation

Table 3. Inter-Device Correlation

<table>
<thead>
<tr>
<th></th>
<th>Goni-Builder ICC (95% CI)</th>
<th>Goni-Medical ICC (95% CI)</th>
<th>Builder – Medical ICC (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pronation</td>
<td>0.434 (-0.034 – 0.708)</td>
<td>0.377 (-0.080 – 0.676)</td>
<td>0.711* (0.539 – 0.826)</td>
</tr>
<tr>
<td>Supination</td>
<td>0.557 (0.212 – 0.755)</td>
<td>0.534 (0.278 – 0.714)</td>
<td>0.718* (0.551 – 0.829)</td>
</tr>
<tr>
<td>Shoulder ER</td>
<td>0.664 (0.477 – 0.794)</td>
<td>0.647 (0.453 – 0.783)</td>
<td>0.680 (0.497 – 0.805)</td>
</tr>
<tr>
<td>Shoulder IR</td>
<td>0.727* (0.565 – 0.835)</td>
<td>0.430 (0.145 – 0.641)</td>
<td>0.704* (0.427 – 0.843)</td>
</tr>
</tbody>
</table>

Goni = Goniometer; Builder = Builder’s Inclinometer; Medical = Medical Inclinometer; ER = external rotation; IR = internal rotation. * = ICC > 0.700

Figure 3. Bland Altman plot of the differences between the builder’s and medical inclinometer for A. Pronation and B. Supination.

The solid black line is the mean difference between devices and the segmented lines demonstrate the 90% limits of agreement.

Figure 4. Bland Altman plot of the differences between the builder’s and medical inclinometer for A. Shoulder internal rotation and B. Shoulder external rotation.

The solid black line is the mean difference between devices and the segmented lines demonstrate the 90% limits of agreement.

DISCUSSION

Accurate, reliable, and efficient measurements are important to detect motion deficits and changes over time. The SG is readily available in most clinics, has standardized methodology, and low cost. Use of the SG is taught in many medical professions relying on joint ROM assessment to determine baselines and the impact of interventions. However, the time efficiency and comparison of the SG to multiple types of inclinometers has not previously been fully described.

In this study, the inclinometers were more than twice as fast as the SG for measuring range of motion, equating to 3.2 to 6.0 seconds per measurement. Consider a clinician performing pre/post treatment ROM measurements for two motions, for example shoulder internal and external rotation, this could mean saving half a minute of time. When considering the course of a full-time work week, the time required equates to over 30 minutes that could be spent on patient management, clinical documentation, or limiting the need for overtime. The saving of time should entice both clinicians and clinic managers to incorporate inclinometers into clinical practice. While there was a statistically significant difference between the BI and MI for measuring pronation and supination, the nearly half second difference is not clinically relevant, such that, the adoption of either inclinometer device would improve efficiency.

Intrarater reliability for each device was good to excellent with the highest reliability demonstrated with the MI (average ICC value of 0.914), followed closely by the BI (average ICC value of 0.908), and the lowest ICC found with the SG (average ICC value of 0.809). In this study, the SG was the least reliable assessment tool; a result consistent with other studies that found inclinometry to be more reliable than goniometry.25,26 Nonetheless, the SG is the most used clinical tool to assess ROM. Cools et al.27 found high SG intrarater reliability for shoulder rotation (ICC ranging from 0.850 to 0.990). However, for that study, two individuals performed the measurements: one providing stabilization and one performing the measurement. In clinical practice, it is rare to have the assistance of another individual when performing ROM assessment.

International Journal of Sports Physical Therapy
Measurement error in the current study was low, with an MDC from 1.8° to 2.8 for the SG and from 1.7° to 2.0° for the inclinometers. Cools et al. reported a greater MDC for shoulder rotation, ranging from 4.4° to 8.0° for the SG and 4.0° to 6.4° for the inclinometer. Likewise, both the SEM and MDC for shoulder rotation with the SG and inclinometers in the current study were lower than those found by Correll et al. when using a SG and the Halo® device.

As expected, measurements taken with the BI and MI were more highly correlated with each other than with the SG. The average ICC value for BI and MI was 0.703 indicating good, but not excellent, consistency between devices. One reason for the small differences between the BI and MI could be due to the different device contact surfaces. The flat contact surface of the BI may allow for unintentional tilting of the device and thus align differently with body contours compared with the central rounded cutout of the MI. While the correlation between inclinometers was lower than expected, the mean differences between devices were quite small (0° to 4°). Clinically, these differences are unlikely to be relevant. However, given individual measurements varied by as much as 5°, it is recommended that clinicians use the same type of inclinometer when measuring a patient to ensure a change, or lack of change, has occurred. These suggestions are consistent with recommendations of other studies. Given the low cost and high reliability of the BI, clinicians and clinic managers may consider use of a BI rather than investing in the more expensive medical inclinometer.

STRENGTHS

This study included participants of a broad range of ages without regard to prior injury or surgery, theoretically resulting in a wider range of scores than if only young, healthy individuals were tested. Having more variability in ROM is expected to prevent a misleadingly low reliability coefficient due to low score variability. This study utilized one tester who was blinded to measurement results, reducing the potential for variations in passive force application to achieve end ROM.

LIMITATIONS

This study included participants who were symptom free. Therefore, it is unclear if the results might apply to individuals who experience pain with testing. Since this study utilized only one tester, it is not possible to determine intratester reliability. Additionally, the tester was an experienced clinician and results may not generalize to an inexperienced clinician. The BI used in this study is typical of builder’s digital inclinometers, in that the device cannot record measurements greater than 90°. Angles greater than 90° required the recorder to calculate the true measurement which could have been a source of error. It is not known if other types of BI would perform similarly to the one used in this study.

CONCLUSION

The BI and MI were found to be reliable for measuring pronation, supination, and shoulder rotation. The low correlation between the SG and inclinometers indicates that clinicians should utilize the same device for testing. Because time can be a barrier to clinician measurement and remeasurement during an episode of care, the greater efficiency and reliability of inclinometers warrants consideration as the new measurement standard. In contrast to the SG, inclinometer assessment requires only one hand to manage the device. The MI cost may be prohibitive for routine clinical use. However, the lower BI cost makes the device clinically affordable. Future studies should examine the reliability of a BI capable of displaying angular measurements greater than 90°. Assessment of a greater number of joint motions using the BI and on various patient populations would be beneficial. Given the current lack of standardization, creating and utilizing a standard for patient and inclinometer positioning is recommended to enhance reliability.

DISCLOSURES

The authors report no conflicts of interest.

Submitted: March 15, 2023 CDT, Accepted: June 23, 2023 CDT

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-NC-4.0). View this license’s legal deed at https://creativecommons.org/licenses/by-nc/4.0 and legal code at https://creativecommons.org/licenses/by-nc/4.0/legalcode for more information.
REFERENCES


2. van de Pol RJ, van Trijffel E, Lucas C. Inter-rater reliability of measurement of passive physiological range of motion of upper extremity joints is better if instruments are used: a systematic review. J Physiother. 2010;56(1):7-17. doi:10.1016/s1836-9553(10)70049-7


A Novel Intrinsic Foot Muscle Strength Dynamometer Demonstrates Moderate-To-Excellent Reliability and Validity

Jennifer Xu, Dante D Goss, Susan A Saliba

1 Kinesiology, University of Virginia

Keywords: intrinsic foot muscles, intrinsic foot muscle strength, reliability, foot health, foot injury, handheld dynamometer, strength testing

https://doi.org/10.26603/001c.84310

Background
Intrinsic foot muscle (IFM) weakness can result in reduced foot function, making it crucial for clinicians to track IFM strength changes accurately. However, assessing IFM strength can be challenging for clinicians, as there is no clinically applicable direct measure of IFM strength that has been shown to be reliable and valid with the foot on the ground.

Purpose
The purpose was to investigate the intra-rater and inter-rater reliability of a novel, budget-friendly IFM dynamometer and determine its agreement with a handheld dynamometer (HHD). The researchers also examined correlations of foot morphology and activity level to IFM strength.

Study design
Descriptive Laboratory Study

Methods
Two assessors measured IFM strength of 34 healthy volunteers (4 male, 30 female; age=21.14±2.57, height=164.66±7.62 cm, mass=64.45±11.93 kg) on two occasions 6.62±0.78 days apart with the novel dynamometer to assess intra- and inter-rater reliability. The HHD was used to measure IFM in the first session in order to assess validity.

Results
For the novel dynamometer, intra- and inter-rater reliability was moderate-to-excellent (ICC = 0.73 – 0.95), and the majority of the strength tests were within the 95% limits of agreement with the HHD. Wider foot morphology and a higher number of days walking over the prior seven days had small but significant correlations with IFM strength (dominant foot r = 0.34, non-dominant foot r = 0.39; r = -0.33, -0.39 respectively).

Conclusion
This novel IFM dynamometer is a budget-friendly ($75) tool that was shown to be reliable and valid in a healthy population.

Levels of evidence
Level 3

©The Author(s)

a Corresponding author:
Jennifer Xu
Department of Kinesiology, University of Virginia, Charlottesville, Virginia 22903, United States of America
Email: tnd9vf@virginia.edu
Telephone number: 215-264-3222
INTRODUCTION

Muscle strength is one of the most important factors when assessing an individual's recovery from injury. 1 For the intrinsic foot muscles (IFM), there is no cost-effective or reliable way to measure their strength in a clinical setting. 2 The IFM dynamically support the medial longitudinal arch, 3, 4 aid in performing functional movements, 4 and contribute to balance. 5, 6 Given this, assessment of IFM is particularly important in patients who may suffer from IFM weakness, including older adults at risk for falls, 7, 8 people with diabetic neuropathy, 9 and individuals with plantar fasciopathy, 10–12 among other pathological conditions of the lower extremities. For patients with IFM weakness, it is important to have a reliable IFM strength measurement in order to document impairments and track strength progression over the course of a rehabilitation program.

Researchers have approached IFM strength assessment in multiple ways, but these assessment methods are neither cost-friendly nor easily accessible by most clinicians. 2 For example, MRI, 13–15 ultrasound, 16–18 and EMG 19 have previously been used to assess IFM volume, cross-sectional area, and muscle activity respectively. These measures are considered a proxy for strength as no actual force output is obtained. 2 Furthermore, those methods are time-consuming and expensive for general clinical use. To approach this problem further, researchers have used custom-built toe dynamometers, 19–24 pressure mats, 25, 26 and a variety of handheld dynamometers (HHD). 27, 28 While these measures do provide a force output, they do not discriminate between intrinsic and extrinsic muscle groups and increase the overall cost of assessment.

The Toe/Hiip Strength Dynamometer (Human Locomotion, Newton, MA) is a novel IFM dynamometer that has been recently developed for clinicians. Using this tool, the test is conducted similarly to the paper-grip test, which was initially created as a dichotomous screening tool for patients with leprosy to assess if they were able to grip a business card with their toes or not, due to the neuropathy that is common with the disease. 29, 30 This dynamometer is similarly made of a plastic card attached to a hanging scale. For the test, individuals press their toes onto the card as hard as possible without toe curl to resist the assessor pulling the card out, and a peak force output is provided once the card slides out. Although it is still difficult to conclusively isolate the IFM, this device has a lower cost ($75) and size compared to other methods and may be useful in clinical settings. Furthermore, the tool is designed for the measurement to occur in a closed-chain position, which is the functional position of the foot.

There are also some individual factors that may affect IFM strength, such as foot morphology and recent activity level. Foot width in sitting and standing positions has been shown to significantly positively correlate with IFM strength when measured using other tools. 31 Additionally, a longer foot has a longer lever arm that renders musculature able to generate a higher force. 31 Recent activity level could also play a role in force output, given that muscular fatigue is known to decrease contraction intensity. 32

The purpose of this study was to investigate the intra-rater and inter-rater reliability of a novel, budget-friendly IFM dynamometer and determine its agreement with a handheld dynamometer (HHD). The researchers also examined correlations of foot morphology and activity level to IFM strength. It was hypothesized that the novel IFM dynamometer would have good reliability (ICC = 0.75–0.9), have good agreement with the HHD, and that foot width and length, and activity level would significantly influence an individual’s force output.

METHODS

In this laboratory-based study, two novice assessors measured participants’ IFM strength on two occasions 6.62±0.78 days apart to assess intra-rater and inter-rater reliability of the novel dynamometer. Other demographics and IFM strength using the HHD were also assessed at the first visit.

PARTICIPANTS

Healthy individuals between the ages of 18–30 were recruited. Individuals were excluded if they had any previous history of foot or ankle surgery, and any foot or ankle neuromusculoskeletal injuries or fractures within the prior three months. G-Power's two-tailed paired t-test for two dependent means determined that a minimum sample size of 34 participants were needed for adequate power of 80% and α = 0.05. All participants read, signed, and approved of the informed consent form, and the study was approved by the Institutional Review Board.

INSTRUMENTATION

Arch Height Index Measurement tool (AHI): The AHI (Figure 1a) measures foot morphologic characteristics (length, width, height), and has excellent inter-rater and test-retest reliability (ICC = 0.98–1.00). 27 The AHI can provide measurements to calculate foot volume (foot length, foot width, and dorsal arch height at 50% of the total foot length). 33

Novel IFM dynamometer: The novel IFM dynamometer consists of a plastic card attached to a hanging scale (American Weigh Scales, Georgia, USA), pictured in Figure 1b and Supplementary Video 1. There is no previous reliability or validity data on this device.

Before using the device, individuals were taught how to perform the motion. In this study, participants performed three isometric toe flexion repetitions while seated in a chair with the feet on the ground for ease of learning. They were shown a video of isometric toe flexion, where the toes were first extended to raise the medial longitudinal arch (MLA) as seen in Figure 2a, placed back onto the floor without losing arch tension, then pushed down onto the floor for three seconds when instructed on a “3, 2, 1, push” countdown (Figure 2b, Supplementary Video 3). Instructs were provided to “keep the heel and ball of the foot on the ground”, “imagine bringing the ball of the foot and heel together”, and to avoid curling their toes. This “short
foot* exercise mimics the setup of the IFM test which assesses the ability of the IFM to support the MLA,\textsuperscript{34} which was asserted to be more important than the ability of IFM to produce toe flexion.

Then, participants were shown a video of the assessment being conducted, with the great toe and lesser toes on each foot tested separately. The assessor instructed the participant to extend their toes and slide the plastic card under the toe(s), ensuring that the card did not touch the underside of the metatarsal head(s), as seen in Supplementary Videos 2 and 4. Participants placed the toes back on the card, and the assessor placed their hand on top of the participant’s foot to keep it stable. The assessor counted ”3, 2, 1, push”, with participants instructed to push “as hard as they could” on that command. Then, assessors counted to three silently while they pulled the handle of the device, increasing the force of the pull slowly so that the card slid out on “3”, keeping the device on the ground with the line of pull in the sagittal plane. Participants were instructed to ”press the toes on the card as firmly as possible to keep it in place”, and to press with all toes on the foot regardless of the toe condition being tested. This process is explained in Supplementary Video 3.

All assessments were conducted on the same low-pile carpet, as friction of different materials could affect the force output. Participants lay supine with knees bent at 90 degrees of flexion, with the feet flat on the floor and their arms crossed on their chest, as pictured in Figure 2c and Supplementary Video 2. This position reduced the influence of bodyweight on plantar pressure, as increased postural demand has been shown to increase IFM recruitment\textsuperscript{35,36} and the amount of ankle plantarflexion theoretically allowed for extrinsic muscles to be involved to a smaller degree.\textsuperscript{37} Further, it was easier to maintain a standard testing position between patients compared to a seated or standing position.

Handheld dynamometer (HHD): Fraser et al. (2017) demonstrated good-to-excellent reliability (ICC = 0.66 – 0.92)\textsuperscript{27} of a HHD in measuring strength of the great toe and lesser toes (Figure 1c). The present study used the same microFET2 HHD (Hoggan Health Industries, West Jordan, UT). For this test, participants lay supine on an examination table with knees bent at 90 degrees of flexion, with their toes hanging off the edge so that the HHD transducer pad could be placed under the toes to assess their pushing force. The assessor counted ”3, 2, 1, push”, with participants instructed to push “as hard as they could” on that command.

International Physical Activity Questionnaire-Short (IPAQ-Short): The IPAQ was developed in 1998 as an objective self-
Figure 2. a. beginning the test by lifting the toes. b. placing the toes back down onto the card. c. the testing position for the novel IFM dynamometer and d. measuring range of motion of the great toe.

reported measure of physical activity across a variety of countries, demonstrating acceptable test-retest reliability, acceptable agreement between the long and short versions, and fair to moderate agreement between the IPAQ and an accelerometer. It is sensitive to the specific research population (country, age, region) and the intention behind the survey, as the long-form may be more appropriate for domain-specific activity (job-related, transportation, housework, and recreational).39

TESTING PROCEDURE

Procedures were as follows (Figure 3).

Foot morphology: The dominant leg of participants was obtained by asking participants which leg they would use to kick a soccer ball. Foot length and width of the dominant leg was measured using the AHI while barefoot; and arch height and foot girth (using a soft measuring tape) were obtained at the midfoot.39 Great toe range of motion was obtained passively (Figure 2d).

Novel dynamometer testing: After learning the isometric toe flexion motion required for the test, participants underwent two familiarization trials per toe condition in the supine hook-lying position. Both the familiarization and assessment trials had four toe conditions, randomized via a Latin square formation: 1, dominant foot great toe (DGT); 2, dominant foot lesser toes (DLT); 3, non-dominant great toe (NGT); 4, non-dominant lesser toes (NLT).

After familiarization, five minutes of rest were provided, where the participant remained seated. Then, each assessor conducted three trials of each toe condition, in the same order as the familiarization trials. There were 30 seconds of rest in-between each toe-pushing repetition within each foot, though there was no rest when transitioning between feet (Supplementary Video 4). Five minutes of rest were provided between each assessor, with each assessor blinded to the others’ results for all sessions. The order of the two assessors was also alternated for each participant to rule
out practice or fatigue effects, though assessor 1 instructed all participants in the test and conducted all the familiarization trials.

**Handheld dynamometer measurement:** After the IFM strength assessment with the novel dynamometer, participants rested five minutes, then underwent the strength assessment with the HHD (Figure 1c) for only the great toe and the lesser toes on the dominant foot.

**Activity level:** In the five minutes of rest between each assessor, participants filled out the International Physical Activity Questionnaire-Short (IPAQ-Short) to indicate their recent activity levels over the prior seven days. At the end of the first session, a second assessment was scheduled five to seven days later. The same procedure took place, excluding initial demographics, foot morphology, and seated practice repetitions.

**Data analysis:** For the novel IFM dynamometer, means and standard deviations (SD) per assessor and session were calculated, along with standard error of measurement (SEM) to indicate measurement variation. Intra-class correlation coefficients (ICC) were calculated using SPSS (Version 28, Chicago, IL); ICC(3,1) indicates intra-rater reliability for each assessor and ICC(2,1) indicates inter-rater reliability between assessors, according to previous reporting guidelines by Koo and Li (2016). ICC values were interpreted as poor (< 0.5), moderate (0.5-0.75), good (0.75-0.9), and excellent (> 0.9).

For the HHD, means and SD for each toe condition were calculated. To determine agreement between the novel IFM dynamometer and the HHD, a Bland-Altman plot was created with 95% limits of agreement (R Statistical Software, v4.1.2; R Core Team 2021).

Pearson correlation coefficients were used to determine the correlation between IFM strength values, foot morphological characteristics, BMI, and activity level in the prior seven days (R Statistical Software, v4.1.2; R Core Team 2021). Measurements of great and lesser toes were added together on each foot to perform the correlations. The significance level was set $a$ priori at $\alpha = 0.05$. Correlation coefficient $r$: $|r| > 0.8$ indicates high correlation; $0.5 > |r| < 0.8$ denotes moderate correlation; $0.3 > |r| < 0.5$ suggests low correlation; and $|r| < 0.3$ implies weak correlation.\textsuperscript{31}

**RESULTS**

Initially, 37 healthy individuals volunteered to participate in this study; three participants were lost to follow-up for the second session (Figure 3).

Participant demographics and foot morphology data are presented in Table 1. For the novel IFM dynamometer, intra-rater and inter-rater reliability was moderate-to-excellent across all conditions (Table 2). For agreement with the HHD, the lesser toes of the dominant foot had all data points within the 95% limits of agreement (Figure 4a). For the great toe of the dominant foot, only two data points were outside of the 95% limits of agreement (Figure 4b). Certain variables had small but significant correlations with IFM strength (Table 3), including age, foot width and girth, and number of days walking over the previous seven days.

### Table 1. Participant characteristics

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>21.14</td>
<td>2.57</td>
</tr>
<tr>
<td>Sex</td>
<td>4 males, 30 females</td>
<td></td>
</tr>
<tr>
<td>Mass (kg)</td>
<td>64.45</td>
<td>11.93</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>164.66</td>
<td>7.62</td>
</tr>
<tr>
<td>BMI (kg/m(^2))</td>
<td>23.74</td>
<td>3.76</td>
</tr>
</tbody>
</table>

| Foot morphology | |
|-----------------|------|----|
| Foot length (cm) | 24.44 | 1.35 |
| Foot width (cm)  | 8.91  | 0.53 |
| Arch height (cm) | 5.96  | 0.53 |
| Foot volume (cm\(^3\)) | 683.4 | 111.24 |
| Foot girth (cm)  | 21.94 | 1.26 |
| Great toe ROM (°) | 39.29 | 12.68 |

| IFM Strength - HHD | |
|--------------------|------|----|
| Great toe (kg)     | 16.31 | 7.43 |
| Lesser toes (kg)   | 14.81 | 6.52 |

<table>
<thead>
<tr>
<th>Activity level – IPAQ-Short</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Activity level: # days over past week</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vigorous</td>
</tr>
<tr>
<td>Moderate</td>
</tr>
<tr>
<td>Walking</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Activity level: minutes on 1 day over past week</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vigorous</td>
</tr>
<tr>
<td>Moderate</td>
</tr>
<tr>
<td>Walking</td>
</tr>
<tr>
<td>Sitting</td>
</tr>
</tbody>
</table>

* ROM, range of motion; HHD, handheld dynamometer

**DISCUSSION**

The novel IFM dynamometer demonstrated moderate-to-excellent intra-rater reliability and good-to-excellent inter-rater reliability among all measures between and within investigators. The values obtained from the novel IFM dynamometer data was in agreement with the values obtained from the HHD. Certain variables including a wider foot and higher reported days spent walking in the past week had small but significant correlations to IFM strength values using the novel tool.

A variety of IFM strength measurement methods have been explored, though with certain limitations. Previous methods include MRI, ultrasound, and EMG to obtain muscle volume,\textsuperscript{13-15} cross-sectional area (CSA),\textsuperscript{16-18,41} and muscle activity,\textsuperscript{19} respectively. These characteristics have demonstrated relationships to muscle strength\textsuperscript{42} and can provide insight about IFM function. However, without a quantitative force output,\textsuperscript{2} these measures should be considered an indirect strength assessment. Further, these methods have a high cost and are difficult to access for some clinicians. MRI is considered the gold standard and is

---

\textsuperscript{1} A Novel Intrinsic Foot Muscle Strength Dynamometer Demonstrates Moderate-To-Excellent Reliability and Validity

International Journal of Sports Physical Therapy
Table 2. Toe condition means, SD, SEM, and ICC values

<table>
<thead>
<tr>
<th>Toe condition</th>
<th>Toe Condition Means (SD), in kg</th>
<th>ICC values (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Session 1 – Baseline</td>
<td>Session 2 – Reassessment</td>
</tr>
<tr>
<td>Assessor 1 (kg)</td>
<td>Assessor 2 (kg)</td>
<td>Assessor 1 (kg)</td>
</tr>
<tr>
<td>1</td>
<td>2.64 (1.78)</td>
<td>2.27 (1.7 kg)</td>
</tr>
<tr>
<td>2</td>
<td>1.77 (1.19)</td>
<td>1.65 (1.13)</td>
</tr>
<tr>
<td>3</td>
<td>1.81 (1.37)</td>
<td>1.71 (1.13)</td>
</tr>
<tr>
<td>4</td>
<td>1.85 (1.34)</td>
<td>1.83 (1.28)</td>
</tr>
</tbody>
</table>

* 1, great toe dominant foot; 2, lesser toes dominant foot; 3, great toe non-dominant foot; 4, lesser toes non-dominant foot
† ICC values were interpreted as poor (< 0.5), moderate (0.5–0.75), good (0.75–0.9), excellent (> 0.9)

Figure 4. Bland-Altman plots to demonstrate limits of agreement between two methods of IFM dynamometry capable of isolating individual IFM, but can be expensive and time-consuming, and does not actually measure functional strength. IFM assessment via ultrasound requires specific clinician experience in diagnostic ultrasound, and should be conducted in a weight-bearing position as their function is to support the arch while weight-bearing, but it can be difficult without specially constructed lab equipment. Ultrasound findings could also be dependent on recent activity and time of day, as IFM may be more engorged at the end of the day because of increased blood flow with more muscle use. Lastly, though surface and intramuscular EMG can detect muscle activity, they do not provide an actual force output.

IFM strength can also be measured directly via custom-built toe dynamometers, pressure platforms, and force plates. The HHD has good-to-excellent reliability (ICC 0.66 – 0.92 and 0.82 – 0.88 in two separate studies), and does not allow for flexion of the interphalangeal joints of all toes compared to some other methods, which is viewed as ideal when assessing IFM. Though these quantitative methods provide a force output, there is no conclusive isolation of individual IFM, which may be necessary as individual IFM have different functions. Further, some of these methods are still not cost-friendly or accessible to most clinicians. Building custom dynamometers can require expertise in biomechanics and materials that clinicians do not generally possess. Pressure mats and force plates can be costly, and it can be difficult to isolate the toe pushing force from the rest of the foot.

Various methods exist to assess IFM strength in laboratory settings, but assessing IFM quantitatively in a clinical setting has proven difficult due to lack of clinic space, cost of, and access to appropriate instruments. However, being able to accurately and efficiently assess IFM strength in standard physician offices, physical therapy clinics, or athletic training facilities is necessary to inform clinical decision-making when treating foot and lower leg pathologies. The novel dynamometer in this study will be applicable in clinical settings, given its low cost and small size.

Though the IFM function mostly in a weight-bearing position and performing the assessments in a standing po-
Table 3. Correlations

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Dominant</th>
<th>Non-dominant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.29</td>
<td>0.40*</td>
</tr>
<tr>
<td>BMI</td>
<td>0.14</td>
<td>0.17</td>
</tr>
<tr>
<td>Foot length (cm)</td>
<td>0.17</td>
<td>0.18</td>
</tr>
<tr>
<td>Foot width (cm)</td>
<td>0.34*</td>
<td>0.39*</td>
</tr>
<tr>
<td>Arch height (cm)</td>
<td>0.24</td>
<td>0.13</td>
</tr>
<tr>
<td>Foot volume (cm³)</td>
<td>0.31</td>
<td>0.28</td>
</tr>
<tr>
<td>Foot girth (cm)</td>
<td>0.25</td>
<td>0.33*</td>
</tr>
<tr>
<td>GT (1), range of motion</td>
<td>-0.11</td>
<td>-0.02</td>
</tr>
<tr>
<td>Activity level: # days over past week</td>
<td>-0.11</td>
<td>-0.14</td>
</tr>
<tr>
<td>Vigorous</td>
<td>0.28</td>
<td>0.31</td>
</tr>
<tr>
<td>Moderate</td>
<td>0.07</td>
<td>0.06</td>
</tr>
<tr>
<td>Walking</td>
<td>-0.33*</td>
<td>-0.39*</td>
</tr>
<tr>
<td>Vigorous</td>
<td>-0.12</td>
<td>-0.16</td>
</tr>
<tr>
<td>Moderate</td>
<td>-0.04</td>
<td>-0.10</td>
</tr>
<tr>
<td>Walking</td>
<td>-0.14</td>
<td>-0.10</td>
</tr>
<tr>
<td>Sitting</td>
<td>0.03</td>
<td>0.0</td>
</tr>
</tbody>
</table>

* significant at 0.05 level (2-tailed)
Dominant foot = great + lesser toes (Mean 9.55, SD 6.04); Non-dominant foot = great + lesser toes (Mean 7.99, SD 3.23)

sition could be considered more functional,25 the supine hook-lying position used in this study can ascertain if individuals could produce the toe-pushing testing movement without influence from other larger muscles. The ability to contract IFM in an isolated manner should indicate how much the IFM can contribute to supporting the MLA in functional movements. Conclusively isolating the IFM is still difficult, but the testing position chosen was intended to minimize extrinsic foot muscle involvement as much as possible, and matches the position used previously by Fraser et al.27 for testing with a HHID. Because the novel dynamometer had participants pressing their toes on the ground, a flat, standardized foot position was guaranteed to a greater degree than Fraser et al.'s procedure, where participants' toes hung off the edge of a treatment table to be placed on the transducer of the HHID.27 The positioning for the HHID strength test could result in a variety of ankle and metatarsophalangeal joint angles, which could alter force production capability.37 In a previous study where individuals had their toe flexion strength assessed using a grip bar that required toe curl, indicating extrinsic toe muscle involvement, placing the ankle in 120 degrees of plantarflexion had the lowest strength, while a neutral position at 90 degrees had the highest.19 This could indicate that the extrinsic toe flexor muscles are affected by the ankle joint angle, which is why a plantarflexed ankle position was chosen for this study.

While standing, an individual's bodyweight or forward lean during the test could increase loading onto the toes and increase their force output, given that increased postural demand leads to increased IFM activation.35,36 Further, forward leaning ability correlates highly with ankle dorsiflexor and plantarflexor strength,46 which means that those who can lean forward more during this type of test could falsely inflate their output. In a sitting position, which has previously been used for IFM strength tests, joint angles can vary based on anthropometric measurements, and individuals are still able to lean forward to affect force production. Therefore, the force reading supplied by the device might be considered more of a "force output" than absolute "strength", given that fatigue, body positioning, and effort provided may affect the output.

In terms of agreement between the two dynamometers, the majority of comparisons were within the 95% limits of agreement, or two standard deviations of the mean, which is ideal. It also appears that as the individual's average strength value increases, the difference between the two methods tends to get larger. Thus, although the methods would not be interchangeable, they do appear to agree.

There were some individual factors that could have affected IFM strength, such as foot morphology. Foot width in both sitting and standing positions has been shown to significantly positively correlate with IFM strength.31 Additionally, a longer foot has a longer lever arm that could theoretically generate a higher force.31 Further, habitual use of minimalist shoes is associated with increased foot width47 and increased IFM strength48 compared to those in conventional footwear, which could indicate an association between foot width and IFM strength. In this study, foot width had a small but significant correlation with IFM strength of both dominant and non-dominant feet, which is in agreement with previous findings.

Activity level could play a role in IFM force output capability on a given day, as muscular fatigue is associated with a decrease in contraction intensity.32 IFM fatigue specifically has been shown to increase navicular drop,49 which could indicate a loss of stability, due to decreased strength from the fatigue. In this case, the number of days of walking was correlated with IFM strength, which could indicate that walking more fatigues the IFM more. This may be an indication for clinicians to keep testing time consistent for individuals, when possible, given they may be more fatigued in the evening after a day of activity, compared to the morning. It may also be prudent to ask patients about any changes in their activity levels when performing the testing. However, given that 29 out of 37 individuals in this study had moderate levels of activity, this consideration may not be as applicable for individuals with different activity levels.

There was an increase in IFM strength over time between these two sessions, even though no strengthening was conducted, likely due to a learning effect. To combat this, a familiarization session was used for both sessions prior to the testing. Some individuals had no prior experience with the testing movement in the first session, but all participants indicated that they had previous experience at the second session. It is possible that having previous experience with the toe pushing movement could make it easier to perform the test, thus leading to higher force output. This may be an inherent problem to strength testing, especially of the IFM. There is a high perceived mental and physical workload associated with initially learning IFM exercises that decreases
after two weeks,\textsuperscript{50} which should be considered for patients learning IFM exercises for the first time versus those who have previous experience. Clinicians should be wary that testing during the initial phase of learning the test may be slightly disrupted by this learning effect.

**Limitations:** Though multiple decisions were made in determining the testing procedure to limit a learning effect and isolate the IFM as much as possible, it still must be acknowledged that a learning effect likely exists and some of the extrinsic foot muscles could have played a role in IFM strength during the test. Further, this device has not yet been validated in injured individuals who could stand to benefit the most from a device to track IFM strength in injury recovery.

**CONCLUSION**

The results of the current study indicate that a novel, budget-friendly ($75), IFM dynamometer demonstrates acceptable intra- and inter-rater reliability when examining a healthy population. The device also demonstrates validity when referenced to a previously used IFM strength testing method with a separate handheld dynamometer. Further, the weak-to-small correlations between IFM strength, foot morphology, and activity level suggest that other factors may explain some of the variance in IFM strength, e.g., IFM activation, motor control, or activity level over a greater period of time. Future research should assess the validity of this device compared to other previously used IFM assessments to determine its feasibility in tracking strength changes in rehabilitation processes in pathological populations, or potentially identifying individuals at risk for injuries. Injured individuals could have different results as pain may inhibit force generation.

**DISCLOSURES**

The investigators received the novel IFM dynamometer from the company Human Locomotion (Newton, MA, USA) that is commercially available for $75.

Submitted: April 04, 2023 CDT, Accepted: July 04, 2023 CDT
REFERENCES


A Novel Intrinsic Foot Muscle Strength Dynamometer Demonstrates Moderate-To-Excellent Reliability and Validity

SUPPLEMENTARY MATERIALS

Video 4

Video 3

Video 2

Video 1
Clinical Commentary/Current Concept Review

A Model for Applying Situational Awareness Theory to the Return to Sport Continuum

Ke’La H Porter, Matthew C Hoch

Keywords: theory, return to sport, situational awareness

INTRODUCTION

Return to sport decisions can be challenging for clinicians and patients. While the phases of healing and clinical practice guidelines for specific injuries such as anterior cruciate ligament rupture or acute ankle sprains can be useful decision aids, return to sport decisions are often individualized and criteria are poorly defined for many patients. Ensuring clinicians have the knowledge and skills to make an informed return to sport decisions is vital for the health and safety of athletes.

The return to sport (RTS) continuum was developed to support the sports medicine team in return to sport decision making. The RTS continuum is comprised of three stages that are aligned with the rehabilitation process. However, numerous factors influence RTS besides physical and functional performance measures. Contextual factors such as psychological, environmental, social/contextual, personal, and cognitive factors may impact the injury, treatment, and outcomes. Disableness models, such as the International Classification of Functioning or Disability and Health, bring attention to these contextual factors and provide a patient-centered framework that accounts for the contextual factors. While disableness models highlight the importance of contextual factors, they may not offer sufficient guidance for transitioning patients back to participation. Combining the components of disableness models with the RTS continuum may help guide the individual and their sports medicine team in RTS decision making;

Background

Despite developing and implementing return to sport guidelines, high rates of re-injury remain. The return to sport continuum is a three-phase, criterion-based progression based on physical and psychological factors used to guide the sports medicine team in return to sport decision making. Situational awareness (SA) pertains to an athlete’s knowledge of the dynamic environment (i.e., their ability to perceive the components in the environment, comprehend the meaning of the perceived information, and predict future actions based on that comprehension). SA can be applied on a cognitive continuum that encompasses three levels, each stage becoming more challenging with additional time constraints and increased uncertainty. Integrating the cognitive continuum with the return to sport continuum may optimize the return to sport process and enhance the athletes’ preparedness for competition by incorporating cognitive challenges aligned with live competition. The purpose of this clinical commentary is to describe a return to sport model that integrates SA theory on the cognitive continuum with additional consideration for surrounding contextual factors.

Level of Evidence

5

©The Author(s)
SITUATIONAL AWARENESS (SA) THEORY

SA is an abstract concept that has been adapted and reconstructed over many years. There are many definitions of SA, though a commonly accepted definition is by Endsley in 1988 stating “situational awareness is the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future.”6 Endsley describes the three components of SA (perception, comprehension, and projection) as a state of knowledge pertaining to the dynamic environment.7 Further, SA, decision making, and action are depicted as distinct components of a linear process where SA dictates decision making and impacts how the individual responds.7 For example, a basketball athlete perceives where the ball, teammates, and opponents are in relation to a specific location on the court with attention to the game clock. Based on the disjointed knowledge they perceive from the environment; they piece together the information to paint a complete picture. The understanding developed from the position of the ball and other players allows them to predict the future movement of their opponent, decide how they want to respond, and act accordingly. Some argue against the linearity of this process, stating that perceiving an object or situation is both a decision made by the individual and an action; therefore, action both precedes SA and is a result of it.13

While SA, action, and decision making are linked, the view of the interaction between these three components has evolved.

Cognitive function influences the ability to achieve SA, though it is also an independent process.7 Cognitive constructs such as attention, perception, and memory should be considered when understanding SA.7 Attention allows individuals to direct their focus and dictate what aspects of the environment are vital and relevant. Attention has a limited capacity, so in complex environments where many events are happening simultaneously, such as the sporting environment, the increased attentional demand takes away from SA.7 Additionally, each individual’s perception of the environment, whether accurate or not, will influence expectations and the projection of the future (direct components of SA).7 Furthermore, manipulating short-term memory stores in conjunction with existing knowledge is used to adapt to the current situation, a process of working memory.5 Beyond working memory, long-term memory categorically stores information linked to relevant goals and scripts.7 It has been shown that long-term memory, rather than working memory, distinguishes between an expert and a novice.15 When an individual identifies specific cues in the environment, long-term memory allows them to draw on those scripts and make rapid decisions.7 Along with long-term memory stores, individuals may achieve automaticity. Automaticity is quick and effortless cognitive processing that requires little attentional demand.7 The individual is aware of their surroundings and the context of the situation; however, they can automatically retrieve the information from memory stores necessary to respond according to the circumstances.7 For example, in a basketball...
game numerous events are taking place at the same time. Suppose the athlete's attention is on the opposing player directly in front of them. In that case, there is less attention on the surrounding environment, so they may miss what is happening on the other side of the court (i.e., there is increased SA regarding the player in front of them and decreased SA of the surrounding environment). If the coach devises a new play and a teammate gesture to signal the play (perception), there are certain expectations on what that means and the immediate actions that will follow. The ability to remember and adapt the play to the present circumstances (working memory) is crucial for successful execution. Naturally, plans don't always work out perfectly. When plans go awry, individuals may revert to what they have practiced numerous times with little thought or planning (long-term memory and automaticity), such as going for the layup when there is an opening. Becoming more effective in these cognitive processes may aid in an individual's ability to achieve SA.

SITUATIONAL AWARENESS ON A COGNITIVE CONTINUUM

Varying degrees of SA and consequently, decision making, may differ based on the uncertainty of the situation and the time constraint. SA on a cognitive continuum describes three levels of decision making and control in response to complex situations: strategic, tactical, and reactive. Strategic control is when there is little to no uncertainty with no time constraints. Strategic control is primarily employed in controlled environments during rehabilitation where the athlete can coordinate movements at a self-designated pace. During the early rehabilitation phase, the athlete can form high-quality movements and establish a solid foundation upon which they can build. Tactical control transitions the individual from the controlled clinic settings to sport-specific drills in relatively controlled environments. Tactical control grows in complexity as uncertainty increases and/or the time allotted to complete the task decreases. The increased time constraint and uncertainty during the tactical level forces the athlete to rely on working memory and challenges the athlete's attention and perception; these cognitive processes aim to help facilitate SA. They begin to increase reliance on previous knowledge and experience to counteract the increased demand. During tactical control, the individual applies what was previously established during strategic control but in time-dependent and ambiguous situations.

As uncertainty and time constraints increase, tactical control progresses to reactive control. The environmental complexity in this phase is reflective of live gameplay. Due to the intricacy of a situation, athletes may exhibit a "panic" style of coordination instead of reactive control in which they are unable to achieve SA while simultaneously carrying out coordinated physical movements. During the reactive "panic" style of coordination, athletes may exhibit impulsive and high injury risk behaviors. As task complexity, cognitive load, and attention demand increase, the likelihood of attaining SA decreases. Due to limited attention capacity, some components of the environment will demand increased attention, taking attention away from other aspects of the environment. In areas of compromised attention there may be altered perceptions and expectations resulting in diminished SA. The culmination of altered cognitive processes that result in diminished SA may result in increased cognitive or movement errors, placing individuals at an increased risk of injury. Furthermore, where working memory falls short, there may be an increased reliance on long-term memory stores and automaticity during reactive control. Automaticity may combat the negative effects of increased uncertainty and time constraints by bypassing the cognitive processes. Automaticity may aid in the transition from exhibiting a panic style of coordination to reactive control.

A MODEL FOR APPLYING SITUATIONAL AWARENESS THEORY TO THE RETURN TO SPORT CONTINUUM

Applying SA theory to the return to sport continuum may provide additional considerations regarding an individual's ability to display control and make rapid decisions during the RTS process. In addition to functional performance, cognitive, psychological, personal, social/contextual, environmental, and physical factors should be considered before a patient initiates the return to sport phase. To promote the health and safety of the athlete, most individuals enter the model at pre-participation or return to participation, regardless of injury type or severity. This ensures that they meet the necessary criteria and contextual factors are considered before progressing through the RTS continuum. Strategic control should be established during pre-participation, tactical control should be established during return to participation, tactical-reactive (the transition from tactical to reactive) control should be established during the return to sport, and reactive control should be established during return to performance. While this model provides general guidelines, it is a fluid process where individuals may shift forward or backward between different levels of SA depending not only on the phase of RTS, but the specific task and environment as well. For example, an individual may exhibit proficiency at a reactive level of control during one task and a tactical level of control during another. Progressing through the phases of SA on the cognitive continuum should be a gradual process where there is successful task completion with minimal errors before adding additional challenges, constraints, or environmental complexity.

Furthermore, the RTS process should be individualized. Two individuals with the same injury and comparable physical function may exhibit different levels of control during the same task; one maintains reactive control while the other reaches failure and experiences a panic style of coordination under similar levels of constraint. The RTS process should hinge on an individual's ability to display movement control safely and successfully in uncertain and complex situations. For continuity, the authors maintain one example throughout the application of SA theory being applied.
to the RTS continuum. There are three example scenarios presented in Appendix 1 for application of the model; however, the purpose of this commentary is not to provide specific clinical guidelines.

Depending on the severity of the injury, athletes may be initially removed from sports participation prior to entering the RTS continuum to protect the health and safety of the athlete. During this time, the pre-participation phase, athletes typically undergo rehabilitation in a controlled clinic setting. Strategic control should be the primary focus during this phase. Regaining range of motion, building strength, and focusing on proper biomechanics are some appropriate tactics to improve strategic control. For example, following an acute lateral ankle sprain in a basketball player, the individual may be completing alphabet drills for range of motion, 4-way ankle exercises for strength, and working on returning to a normal gait pattern. When the player is ready to start dynamic load bearing exercises, they may complete plyometric tasks, working on proper jumping and landing mechanics. With no time constraints, the individual can become comfortable with the task at their own pace and form appropriate techniques in a safe environment.

As an athlete integrates rehabilitation in a clinical setting with sports-specific drills on the sideline, they enter the return to participation phase of the RTS continuum. During this phase, individuals should work toward gaining tactical control by adding decision making to movement competence. Cognitive challenges (i.e., choice-reaction time and working memory) can be added to jumping, landing, and linear movements to gain tactical control. To progress within this phase, increasing the complexity of activities to multidirectional cutting and agility tasks with cognitive challenges may be appropriate. For the lateral ankle sprain patient, it may start with a choice-reaction hopping task where the individual jumps and lands on a specific limb in rapid response to different visual or auditory cues. They should aim to maintain the proper landing mechanics they previously learned. When control during foundational tasks is established, they may progress to non-linear patterns of unanticipated cutting tasks requiring multidirectional movements that are unanticipated and in response to an external cue (i.e., a verbal command or visual cue).

When tactical control has been achieved and the athlete has met all other criteria (i.e., functional performance, cognitive control, and psychological readiness) they may transition to return to sport. While they may be ready to return to sport, it does not mean they have returned to their previous level of performance. Depending on the demands of their sport they may need to improve in aspects of speed, agility, accuracy, reaction time or decision making. During this phase, the increased uncertainty and time constraints shift from tactical control to tactical-reactive control. To achieve tactical-reactive control the athlete should be introduced to sports-specific reactive drills in which they are making rapid decisions in response to their environment (i.e., other players and obstacles). These tasks can be completed individually or integrated into practice, beginning with team drills and transitioning to live gameplay as they achieve greater tactical-reactive control. The lateral ankle sprain patient may increase the number of cues they have to remember and react to during the unanticipated cutting drills. Where previously only two scenarios could happen (cut right or cut left), now with increased cues (cut right, cut left, backpedal, forward sprint), it increases the amount of uncertainty within the task. With the team, they may complete exercises where they dribble through obstacles and are required to act in response to an environmental stimulus (i.e., shooting if an opponent comes from the left versus passing if an opponent comes from the right).

Traditionally, when an individual returns to competition that often signifies the end of formal rehabilitative care.
However, additional intervention may be necessary to return to previous levels of performance. During return to performance, the athlete is working to transition from a “panic” style of coordination where there is increased susceptibility of sustaining injury to achieving reactive control.\textsuperscript{9} As the lateral ankle sprain patient continues to work on unanticipated cutting, they may now have to increase the speed at which they are required to complete the task, giving them little time to process the external situation, predict future actions, and make decisions. Increased exposure to chaotic situations and training to gain reactive control will develop mental models in long-term memory that the individual can draw on, as well as form efficient automaticity where less attention is required and quicker processing occurs.\textsuperscript{7}

Return to sport is a complex and fluid process, therefore athletes may have setbacks and not progress through the continuum as expected. This may be due to numerous factors such as reaggravation of the injury, poor responses to loading, or psychological challenges. If an individual experiences a setback, they may return to the previous level in the continuum, return to the beginning of the model, or be removed from sport. They may return to a previous level due to one or more factors: failure to achieve or maintain the appropriate level of SA, inability to adapt to increased load, or surrounding contextual factors. This is highlighted in the figure with the use of bidirectional arrows between the stages (Figure 1). Following any obstacles, wherever the individual re-enters the RTS model they will continue to work on identified goals and progress through the model until they have achieved the desired level of performance.

There are innumerable factors that may hinder or facilitate in return to play decision making. Contextual factors include cognitive function, psychological factors, personal factors, social or contextual factors, environmental factors, physical factors, and functional performance. Cognitive performance may influence how quickly an individual is able to achieve SA and progresses through the phases of RTS. Psychological factors such as readiness or fear of reinjury may impact how effective an athlete is on the field.\textsuperscript{4} If the athlete is not ready psychologically they may more rapidly display a panic style of coordination or freeze in complex situations, predisposing them to reinjury. Personal factors such as sociodemographic factors, health behavior, or socioeconomic status may not all be modifiable, but are still important considerations in healing and RTS.\textsuperscript{4,5} There can be a wide range of social and contextual factors such as expectations of recovery, quality of life, level of competition, time of the season, and type of sport that may influence stakeholders’ decision to return to play.\textsuperscript{3,4} This includes, but is not limited to, whether they have realistic expectations of recovery versus impractical expectations, if it is preseason versus the conference championships, or if the athlete is a redshirt freshman versus a senior. Environmental factors such as living conditions, transportation, and support from community, family, and friends may create barriers to access and impact efficient RTS. Physical factors (e.g., muscle strength, swelling, and range of motion) and performance on functional tests (e.g., crossover hop, Y-Balance test, Agility t-test) should be within acceptable limits of the contralateral limb or pre-injury levels before attempting to return an athlete to sport.\textsuperscript{5,17} While physical and functional performance may not be direct indicators for RTS readiness, they are vital for the athlete’s safety. Contextual factors surrounding the individual interact with one another and, therefore, must be considered in conjunction rather than in isolation.

Research has explored the development of integrated neuromuscular-cognitive assessments. These assessments may be used to evaluate situation awareness in the RTS context. Some assessments have added cognitive challenges (reaction time, working memory, and inhibitory control) to traditional hop tests (single leg hop, single leg tripole hop, single leg crossover hop, and single leg six-meter hop).\textsuperscript{11} Additionally, other assessments evaluating upper and lower extremity reaction time and inhibitory control, reactive agility, and unanticipated cutting have been developed.\textsuperscript{18} Further, Walker et al.\textsuperscript{19} proposed a criterion-based progression in which tasks may increase in physical or cognitive difficulty when there is a low (0–1 errors) neuromuscular and cognitive error rate. If there is a high error rate (>5 errors) or significant errors the task may be too challenging.\textsuperscript{19} Guidelines from Walker et al.\textsuperscript{19} were utilized for exercise progression in the provided scenarios (Appendix 1). Further research is needed to determine the best approach for assessing SA during the return to sport process.

CONCLUSION

The proposed model for applying SA theory on the cognitive continuum to the return to sport continuum was developed to aid return to play decision making. This model identifies levels of SA that should be targeted during the varying stages of return to play. Additionally, the model considers the numerous contextual factors that influence RTS decisions. Applying this new model may further optimize disablement models by providing guidance on transitioning individuals back to participation. Conversely, neglecting to consider contextual factors and SA in RTS decisions could contribute to an increased risk of re-injury.

CONFLICTS OF INTEREST

The authors have no conflicts of interest.

Submitted: February 16, 2023 CDT, Accepted: June 27, 2023 CDT

This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-NC-4.0). View this license’s legal deed at https://creativecommons.org/licenses/by-nc/4.0 and legal code at https://creativecommons.org/licenses/by-nc/4.0/legalcode for more information.
REFERENCES


SUPPLEMENTARY MATERIALS

Appendix 1

Download: https://ijspt.scholasticahq.com/article/83946-a-model-for-applying-situational-awareness-theory-to-the-return-to-sport-continuum/attachment/171407.docx?auth_token=gC9ZFpm0kc1GQAJVvHFz
INJURIES TO TOOL IN DIAGNOSING PATELLAR TENDON MUSCULOSKELETAL ULTRASOUND: AN ESSENTIAL VISUALIZATION

VISUALIZATION OF THE PATELLAR TENDON, PROVIDING CLINICIANS WITH CRUCIAL INFORMATION REGARDING THE EXTENT OF THE INJURY. IN THIS ARTICLE, WE EXPLORE THE ROLE AND BENEFITS OF THIS NON-INVASIVE IMAGING TECHNIQUE IN THE DETECTION AND EVALUATION OF PATELLAR TENDON INJURIES.

Abstract
Musculoskeletal (MSK) ultrasound has emerged as a vital tool in diagnosing patellar tendon injuries. Traditional imaging techniques, such as X-rays and magnetic response imaging (MRI), have certain limitations in assessing soft tissue structures or may not be easily accessible. MSK ultrasound, with its high-resolution, real-time imaging capabilities, offers a cost-effective, safe, and patient-friendly alternative. It allows for accurate visualization of the patellar tendon's structure and function, thereby facilitating the identification of pathological changes such as thickening, thinning, or disruption which aids in diagnosing conditions such as tendinitis, partial tears, and ruptures. Furthermore, MSK ultrasound serves as an invaluable tool for guiding interventions like injections, as it provides real-time imaging. This article explores the role and benefits of MSK ultrasound in diagnosing patellar tendon injuries, highlighting its cost-effectiveness, accessibility, real-time assessment capabilities, and reduced patient burden compared to other imaging modalities. Despite its numerous benefits, the need for ongoing research to enhance its utility is highlighted. As technology advances, MSK ultrasound is set to revolutionize the early detection and management of patellar tendon injuries.

Introduction
Musculoskeletal ultrasound is rapidly gaining recognition as a primary investigative tool for diagnosing a broad array of MSK disorders, particularly those affecting the tendons, ligaments, muscles, and joints. This advancement in medical technology provides a real-time, non-invasive, cost-effective, and dynamic approach to visualize soft tissue structures that might be challenging to investigate using traditional imaging modalities. A specific area where MSK ultrasound is making considerable strides is in the diagnosis of patellar tendon injuries. Patellar tendon injuries can be frustrating for both athletes and physicians to diagnose, as they often present with subtle signs and symptoms. Accurate diagnosis is important for proper treatment of the condition, as the symptoms associated with patellar tendinopathy are highly variable and easily missed. Because the signs and symptoms can be subtle, MSK ultrasound is a useful tool for helping to accurately assess patellar tendon injuries. MSK ultrasound allows for detailed visualization of the patellar tendon, providing clinicians with crucial information regarding the extent of the injury. In this article, we explore the role and benefits of this non-invasive imaging technique in the detection and evaluation of patellar tendon injuries.

Understanding Patellar Tendon Injuries
The patellar tendon, a crucial part of the extensor mechanism of the knee, plays a vital role in facilitating knee extension. Patellar tendon injuries, ranging from tendinopathies to partial or complete ruptures, can significantly impair knee function and quality of life. The successful management of these injuries hinges on an accurate and timely diagnosis. Traditional imaging modalities like plain radiographs (x-rays) can offer limited information, as they are primarily focused on evaluating bone structures. MRI, although excellent for soft tissues, might not always be accessible due to its cost, availability, and contraindications in some patients. One of the key advantages of MSK ultrasound in diagnosing patellar tendon injuries is its ability to offer high-resolution, real-time imaging of the patellar tendon. This non-invasive imaging technique makes it easier to detect changes in tendon tissue composition and smaller tears that may not be visible on other imaging modalities such as MRI or through the physical examination alone. Additionally, MSK ultrasound can be performed quickly and conveniently, eliminating the need for patients to schedule separate appointments for imaging. The dynamic assessment allows clinicians to visualize the tendon's structure and function in both static and active states. It can readily identify changes in the structure of the patellar tendon, such as thickening, thinning, or disruption, which could suggest pathology.

The Role of MSK Ultrasound in Diagnosis
When it comes to diagnosing injuries, accuracy is key. This is where MSK ultrasound comes in handy for patellar tendon injuries. This advanced imaging technique allows medical professionals to examine the internal structures of the body in real-time, producing detailed and accurate images of the patellar tendon. With its ability to identify pathologic changes over time and to help guide treatment, MSK ultrasound has become an integral part of the management of these types of injuries. By providing a detailed, real-time view of the anatomy and pathology within the tendon, the use of MSK ultrasound can help to ensure that
patients receive timely and appropriate treatment in order to achieve the best possible outcomes.

1. Detection of injuries: MSK ultrasound can effectively identify structural changes in the patellar tendon, such as thickening, thinning, or disruption. Hypoechoic regions within the tendon, a classic sign of tendinopathy, can be easily visualized. In cases of complete or partial tears, MSK ultrasound can delineate the size and location of the tear.

2. Assessment of vascularity: Using Doppler imaging, an essential component of modern ultrasound machines, MSK ultrasound can assess vascularity within and around the tendon. This capability can help differentiate normal from abnormal tendons, as neovascularity is often seen in tendinopathy and healing tendons.

3. Dynamic assessment: Unlike other imaging modalities, MSK ultrasound can assess the patellar tendon while the patient is actively moving their knee. This feature is particularly useful in cases of fat pad impingement or patellar instability, where the pathology might only be evident during specific movements or positions.

4. Guidance for interventional procedures: In addition to diagnostic uses, MSK ultrasound can guide clinicians during interventional procedures such as injections or aspirations, increasing the accuracy and safety of these treatments.

MSK ultrasound is also helpful for monitoring post-treatment progress since it can be used to track changes in tendon structure and composition over time. This is impor-
 tant for ensuring that the treatment plan is having the desired effect and can aid in determining if a different course of action may be necessary. In addition, MSK ultrasound can also help to identify potential areas of re-injury or tendinopathy which may require further attention or intervention. MSK ultrasound can therefore be an essential part of post-treatment management for patellar tendon injuries.

**Limitations and Considerations**

Despite its numerous advantages, it is important to acknowledge that the efficacy of MSK ultrasound is highly operator-dependent, requiring significant expertise and training for accurate interpretation. Specifically, one should ensure that the transducer is positioned correctly, as an inaccurate positioning can lead to incorrect diagnosis. Additionally, it is important to be mindful of the ultrasound settings being used, as incorrect settings can result in inaccurate results.

**Figures 3a and 3b: Distal Patellar Tendon in Long Axis View.**
The cortical reflection of tibial tuberosity is seen to the right and is the hyperechoic arc with shadowing deep to it. The hyperechoic, fibrous echotexture of the patellar tendon can be seen running in parallel fibers to insert on the tibial tuberosity. Deep to the tendon are mixed/marbled echoes of Hoffa’s Fat Pad.

**Figures 4a and 4b: Patellar Tendon in Short Axis View.**
The patella tendon is seen in cross section as well-defined, wide/broad, bristle-like structure superficial to Hoffa’s Fat Pad. The tibia is deep to the fat pad, but is a variable landmark based on probe location along the 5 cm length of the patella tendon.
in poor image quality. Tendons are also susceptible to anisotropy, an imaging artifact that may lead to an incorrect diagnosis of a tear. Finally, interpretation of the ultrasound image can also pose a challenge, as patellar tendon injuries can sometimes appear similar to other injuries. With these considerations in mind, practitioners can ensure that they are making the most of MSK ultrasound in diagnosing patellar tendon injuries.

**Conclusion**
In conclusion, MSK ultrasound is an important and invaluable tool in the assessment of patellar tendon injuries. With its advantages of high-resolution visualization, cost-effectiveness, accessibility, real-time assessment, and dynamic images of the patellar tendon, it is becoming a first-line imaging tool for many clinicians in the assessment of soft tissue structures. It can provide detailed information regarding the structure and composition of tendons that other imaging modalities cannot, as well as guide treatment decisions and monitor response to therapy over time. As such, MSK ultrasound should be considered as part of any comprehensive evaluation and management plan for patellar tendon injuries. By providing a real-time view of the anatomy and pathology within the tendon, MSK ultrasound can help to ensure that patients receive timely and appropriate treatment to achieve the best possible outcomes. While MSK ultrasound has become an essential tool in the diagnosis and management of patellar tendon injuries, continual research is necessary to further refine its use. Advancements in technology and imaging techniques may provide even greater resolution and clarity, aiding in our ability to diagnose and manage patellar tendon injuries effectively.

**PATELLAR TENDON TEAR IN LONG AXIS (LAX) AND SHORT AXIS (SAX)**

**Figures 5a and 5b: Patellar Tendon Tear in Long Axis and Short Axis Views.**
**Figure 5a** shows the cortical reflection of the patella to the left and the yellow arrow points to the cortical irregularities of the patella. These irregularities are secondary to insertional, chronic patella tendinosis. The patella tendon appears thickened as well (Yellow Line) compared to the normal view of a patella tendon. A moderate anechoic defect is seen in the proximal tendon indicating a patellar tendon tear (Red Arrow). **Figure 5b** shows a short axis view of the same patella tendon with the tear in the patella tendon (Red Arrow) viewed as an anechoic defect running perpendicular to the tendon fibers. (Photos courtesy of F. Clarke Holmes, MD.)
Abstracts from the IFSPT/BFSP Partnered International Congress September 22-23, 2023
The Belgian Federation of Sports Physical Therapy is pleased to announce that it will host an International congress on Sports Physiotherapy, partnered with the IFSPT, on 22th and 23th of September 2023.

It will be a unique opportunity to network with the global community of Sports Physiotherapists. The themes of the congress will be Evidence update on the rehabilitation of sports injuries and Sports-specific rehabilitation. Gain the latest insights in optimising your rehabilitation and return to performance strategies!

An IFSPT General Meeting and networking session will be held the afternoon of 21 September.

The location in Brussels is the Holiday Inn Brussels Airport, with easy access to the airport, trains and the historical city center of Brussels. Registration is now open and special rates for the hotel are currently available.

Join us in Brussels this fall!

TOPICS
• Promoting Physical Activity for Health: Strategies
• Ankle Rehabilitation
• Knee Rehabilitation
• Shoulder Rehabilitation
• Rehabilitation of Lower Extremity Tendinopathy
• Muscle Injury Rehabilitation
• The Olympic Athlete in 2024
• Strategic Considerations in Sport-Specific Rehabilitation

EVENTS
• Exposition hall
• Includes breaks and lunch
• Networking dinner and awards ceremony (extra)
• AlfaCare award ceremony

CONFIRMED SPEAKERS
Professor Clare Ardern
Professor Dr. Ann Cools
Suzanne Gard, PT, Msc
Professor Jean-Francois Kaux
Olav Spahl (COIB)
Professor Dr. Bruno Tassignon
Professor Dr. Jo Verschueren
Professor Erik Witvrouw
...and many more!

REGISTER NOW AND RESERVE YOUR ACCOMMODATIONS!
https://www.bfsp.be/nl/Congres/Registration
ABSTRACTS

Deficits in neurocognitive performance in patients with CAI performing a dynamic balance task.
Maricot A, Verschueren J, Meeusen R, Roelands B, Tassignon B.

Screening assessment, perceived training load and injury incidence in a young and preselected volleyball population: results from a 10-month observation period using a retrospective design.
Jaspers A, Houtmeyers K, Vereecken S, Minten T, Dams M, Bogaerts S, Staes F.

Upper limb functional testing: normative data in overhead and non-overhead athletes.
Tooth C, Schwartz C, Croisier JL, Gofflot A, Forthomme B.

The effect of blood flow restriction training on quadriceps activity after anterior cruciate ligament reconstruction.
Foredi F, Nguyen C, Mazeeas J, Hewett TE, Korakakis V, Rambaud AJM.

Biceps femoris proximal myotendinous junction size is associated with hamstring strain injury history and hamstring endurance in male soccer players – a case control study using Magnetic Resonance Imaging.
Schuermans J, Pieters D, Witvrouw E, VanDen Berghe S, Wezenbeek E.

Test-retest, intra- and inter-rater reliability of the reactive balance test in patients with chronic ankle instability.
Maricot A, Verschueren J, De Pauw K, Meeusen R, Roelands B, Tassignon B.

The influence of tibial rotation on hamstring muscles groups’ coordination during a dynamic knee flexion assessment.
Ferré A, Schwartz C, Bornheim S, Delvaux F, Croisier J-L.

Impact of stride frequency on the maximal isometric forces of the plantar flexor muscles and hip abductors muscles during an 156 km ultra-trail.
Ayoub RT.

Higher knee joint work is a risk factor for the development of patellar tendinopathy in volleyball: A prospective study.
De Bleecker C, Vermeulen S, Spanhove V, Segers V, Willems T, Vanrenterghem J, De Ridder R.

What do upper-extremity physical performance tests actually measure? Insights from an electromyographical study.
Borms D, Berckmans K, Lowie A, Lepla D, Maenhout L, Cools A.

Is there an association between the jump-landing profile of volleyball players and the viscoelastic properties of the patellar tendon?

Does baseline cognitive performance predict the response to mental fatigue in healthy individuals?
Schampheleer E, Habay J, Proost M, Roelands B.
The additional value of Ischemic preconditioning during warm-up on isokinetic strength and endurance.
Jacobs E, Schuermans J, Wezenbeek E, Witvrouw E.

Countermovement jump may determine performance alterations after anterior cruciate ligament reconstruction.
Forelli F, Nekhoufi B, Vanderbrouck A, Duffiet P, Ratte L, Hewett TE, Rambaud AJM.

Early open kinetic chain improves strength recovery and return to sport after anterior cruciate ligament reconstruction without graft laxity increasing.

Gastrocnemius muscles activity increasing may impair running patio-temporal parameters after anterior cruciate ligament reconstruction: A pilot study.
Forelli F, Penguè-Koyi A, Mazeas J, Hewett TE, Rambaud AJM.

Superior foot & ankle muscle strength in non-rearfoot endurance runners compared with rearfoot runners: implications for the management of running-related knee injuries.

Neurocognitive deficits related to ligamentous ankle injuries and chronic ankle instability.
Corluy H, Schampheleer E, Maricot A, Verschueren J, Roelands B, Tassignon B.

Peak patellar tendon force during heavy load single-leg squatting is influenced by a decline board but not by the external weight's mass.

Physiotherapists and physiotherapy students estimate patellar tendon forces during rehabilitation exercises equally well.

Does sleep quality predict V0\textsubscript{2}max in healthy adults?
Quinten E, Habay J, Roelands B.

Hamstring muscle fibre type distribution in football players having suffered an Anterior Cruciate Ligament Injury: harder, better, faster, stronger? A case-control study using MR Spectroscopy.
Denolf S, Witvrouw E, Tampere T, Schuermans J.

Fatigue-induced biomechanical risk factors for patellar tendinopathy in volleyball: a prospective study.
Deficits in neurocognitive performance in patients with CAI performing a dynamic balance task.

Maricot A,1 Verschueren J,1 Meeusen R,1,2 Bart Roelands,1 Bruno Tassignon,1

Introduction:
While the majority of patients recover from an initial lateral ankle sprain and become copers, approximately 40% develop chronic ankle instability (CAI), which includes a recurrent feeling of giving way.1 Central adaptations have been suggested to partly explain the aetiology and chronic character of CAI.2 Consequently, deficits in neurocognitive ability may impact musculoskeletal injury incidence rates.3 Balance tests, which are routinely utilised in clinical practice for injury prevention and return to sport decision-making, lack neurocognitive load seen in the sports context. Therefore, the primary aim was to assess neurocognitive function and balancing ability using the reactive balance test (RBT).4

Methods:
Patients and healthy controls in this study visited the lab twice. During the first visit, they familiarized themselves with the Y-balance test (YBT) and RBT test procedures. During the second visit, the YBT and RBT were performed on each leg using the same procedures as the first visit. The randomisation used for the Fitlight colours and inter-stimulus time during the RBT differed between the two trials to avoid recall bias and minimise possible learning effects. The inclusion criteria followed the IAC recommendations5.

Results:
This study included 27 patients with CAI and 22 healthy controls. Patients with CAI had similar YBT test scores regardless of stance limb or axis tested compared to healthy controls (p = 0.455; composite scores; CAI – most affected side: 87.23cm, CAI – contralateral side: 87.86cm, CON: 84.00cm). The RBT revealed deficits in accuracy in the patient group. However, there were no side-to-side differences for either RBT outcomes (accuracy and visuomotor reaction time) in the patient group (p = 0.538; most affected side: 83.12% ± 8.04%, contralateral side: 81.48% ± 8.65%). There were no significant changes between groups for VMRT (CAI - CON: 776.03ms ± 107.77ms, 739.47ms ± 98.43ms, p = 0.584).

Conclusion:
This study found that patients with CAI performed less accurately than healthy controls during a neurocognitive balance task but maintained similar VMRTs. Also, in this study, patients performed the YBT as well as healthy controls. Based on these findings, neurocognitive stimuli should be added to rehabilitation programs for patients with CAI.

References:
Screening assessment, perceived training load and injury incidence in a young and preselected volleyball population: results from a 10-month observation period using a retrospective design.

Jaspers A, Houtmeyers K, Vereecken S, Minten T, Dams M, Bogaerts S, Staes F

Introduction:
Screening for injury prediction in team sports has been questioned in adult populations, but there is limited research on the relationship between screening and injury incidence in youth volleyball athletes. Additionally, the relationship between perceived training load and injury incidence in this population is unexamined.

Methods:
This retrospective study analyzed data from a routine, standardized screening assessment in 46 youth elite athletes aged 12-16 years. Injuries, training participation, and perceived training load using Borg scores were recorded by the medical team during a 10-month follow-up period. Differences in screening tests between a group with chronic overuse injuries and a group without chronic overuse injuries were explored using a Mann-Whitney U test.

Results:
Of the 46 athletes, 29 (63.0%) reported a chronic injury. The injured group scored significantly lower (p < 0.05) on the Biering-Sørensen test for both absolute and relative values, but had significantly higher values on various isometric strength tests, including flexion for the left and right leg, and extension for the right leg. These higher values were observed for absolute but not relative values. No significant differences were observed between the groups for perceived training load.

Conclusion: Our findings show equivocal results regarding differences in screening tests between injured and non-injured youth athletes in this population, which is consistent with earlier research in adult athletes. Therefore, screening information at the group level cannot be used to predict future injuries. In addition, perceived training load does not indicate injury susceptibility. Other approaches, such as multifactorial analysis methods or personalized approaches, should be explored to better understand the complex and dynamic nature of injuries.

References:
Upper limb functional testing: normative data in overhead and non-overhead athletes.

Tooth C,1,2 Schwartz C,1 Croisier JL,1,2 Gofflot A,1,2 Forthomme B,1,2

Introduction:
Upper limb functional testing has become more and more popular over the last years for its reasonable cost, its speed of implementation and its close links with the sporting gesture. However, there is a lack of normative value for most of the tests described in literature. Therefore, the first objective of this study was to provide normative data for upper limb functional testing in handball and rugby players. The second objective was to determine the influence of age and sport (handball vs rugby) on the results obtained.

Methods:
A total of 81 healthy sportspeople (17.8 ± 3 years; 178.7 ± 7.7 centimeters; 77.2 ± 15.2 kilo's) were recruited. They were classified in two categories according to their age (14-18 years or 18-25 years) and their sport (handball or rugby). They performed a battery of upper limb functional tests, including the Single Arm Medicine Ball Throw (SAMBT), the Modified-Athletic Shoulder Test (M-AST), the Upper Limb Rotation Test (ULRT), the Closed Kinetic Chain Upper Extremity Stability Test (CKCUEST) and the Countermovement push-up (CMPU). Isometric shoulder rotators strength was also measured in a supine position, shoulder at 90° of abduction, with a handheld dynamometer.

Results:
Significant differences were highlighted in upper limb performance according to the sport practiced (p<0.05) and the age category (p<0.05). High and significant correlations were found between isometric shoulder rotators strength and the SAMBT (r = 0.71) (p<0.05) in adolescent handball players or the M-AST (r = 0.68-0.87) (p<0.05) in all athletes. No correlation was observed between the other tests and isometric shoulder rotators strength as well as between the functional tests themselves.

Conclusion:
Normative data and cut-offs were provided for the different functional tests in both populations (handball and rugby) and age groups (14-18 and 18-25 years). These data will help clinicians in interpreting the values in an objective of performance, primary prevention or return to play decision.

References:
The Effect Of Blood Flow Restriction Training on Quadriceps Activity After Anterior Cruciate Ligament Reconstruction.

Forelli F,1,3,4 Nguyen C,1,2 Mazeas J,1,2 Hewett TE,4 Korakakis V,6 Rambaud AJM,4,7

Introduction:
The main objective of this study was to evaluate whether quadriceps strengthening with low load blood flow restriction (BFR) improves electromyographic (EMG) activity of the vastus medialis, vastus lateralis, and rectus femoris similarly to quadriceps strengthening using heavy load. The secondary objective was to assess within-quadriceps regional EMG differences among the three muscle heads.

Methods:
This case-control study included 27 patients with a primary non-contact anterior cruciate ligament injury reconstructed with hamstring graft 3 months after surgery (101.9 ± 18.4 days). The control group (n = 14) performed heavy load knee extensions at 80% of maximal voluntary isometric contraction (MVIC), while the experimental group (n = 13) performed the same exercise with low load (30% MVIC) combined with BFR with 80% of limb occlusion pressure. (1,2) Patients performed one set of 12 knee extension repetitions to measure surface EMG activity of the vastus medialis, vastus lateralis, and rectus femoris. After root mean square (RMS) treatment of the raw signal, RMS EMG results were normalized by MVIC activity to allow inter-subject comparability, and are thus shown as %MVIC.

Results:
Univariate between-group analysis showed a significantly increase in RMS EMG for the control group compared to the BFR group for the rectus femoris (50.5%MVIC ± 14.6 vs 36.7%MVIC ± 17.4, p = 0.01, Cohen's d = 0.87), the vastus medialis (56.3%MVIC ± 19.1 vs 31.0%MVIC ± 18.1, p = 0.002, Cohen's d = 0.87), and the vastus lateralis (59.8%MVIC ± 23.0 vs 29.5%MVIC ± 19.1, p = 0.002, Cohen's d = 1.35). No significant differences were observed between rectus femoris and vasti muscles in both the BFR (p = 0.89) and the control group (p = 0.12).

Conclusion:
These findings indicate that high load resistance training increased significantly quadriceps RMS EMG amplitude in all the three examined muscle heads, as compared to the low load BFR group. The results may be too preliminary to draw definitive conclusions about BFR and quadriceps activity after ACLR.

References:
Biceps femoris proximal myotendinous junction size is associated with hamstring strain injury history and hamstring endurance in male soccer players – a case control study using Magnetic Resonance Imaging.
Joke Schuermans,1 Dries Pieters,1 Erik Witvrouw,1 Sarah VanDen Berghe,1 Evi Wezenbeek1

Introduction:
Hamstring strain injuries (HSI) are very common in athletes exposed to repeated high speed running. Previous research established that architectural features might play a crucial role in the muscle's injury risk. Most muscle lesions occur at level of the Biceps Femoris Long Head's (BFLH) proximal myotendinous junction (MTJ), which has never been subject of morphological research in the light of HSI or athletic performance capacity before. This study intended to explore the role of the BFLH's proximal MTJ size in hamstring injury susceptibility and athletic performance by means of a case control study.

Methods:
Fifteen male soccer players with a history of hamstring strain injury (HSIH) and 16 matched controls were submitted to a comprehensive assessment protocol consisting of Magnetic Resonance Imaging (MRI; T1 sequences) to evaluate the morphology of the proximal MTJ of the BFLH and a hamstring strength analysis (isokinetic hamstring strength and hamstring endurance (Single Leg Hamstring Bridge test (SLHB) and isokinetic strength).

Results:
Athletes with a HSI history (HSIH) presented significantly smaller MTJ widths (6.90mm ± 0.89mm) compared to the controls (9.94 mm ± 0.93mm) (p = 0.042). HSIH was also associated with weaker bilateral strength endurance (26 versus 38 SLHB repetitions, p = 0.019), whereas no isokinetic strength differences could be identified based on HSIH. The athlete's SLHB score presented a significant positive correlation with the BFLH's Proximal MTJ geometry (p = 0.014, Spearman's Rho = 0.456), whereas no association with isokinetic strength was found.

Conclusion:
HSIH is associated with smaller proximal MTJ size in the BFLH and hamstring endurance deficits, potentially highlighting the importance of the magnitude of the MTJ's muscle-tendon contact surface, possibly providing the BFLH with more passive stiffness and strength contribution capacity, protecting it against HSI. The SLHB seems to be a valuable functional test and should be promoted in prevention.

References:

Department of Rehabilitation, Faculty of Medicine and Health Sciences, Ghent University, Ghent, Belgium
Correspondence e-mail: joke.schuermans@ugent.be
Test-retest, intra- and inter-rater reliability of the reactive balance test in patients with chronic ankle instability

Alexandre Maricot¹, Jo Verschueren¹, Kevin De Pauw¹,² Romain Meeusen¹,² Bart Roelands¹,² Bruno Tassignon¹

Introduction:
Chronic ankle instability (CAI) affects 40-45% of those who have had an ankle sprain and leads to recurrent episodes of the ankle "giving way" and neuromuscular deficits.¹ Clinicians use functional performance tests (FPTs) like the star excursion balance test (SEBT) and the Y-balance test (YBT) to identify those at risk of (re)injury and make better-informed return to sport (RTS) decisions.²,³ However, these tests are limited in on-field sports contexts. To address this, the reactive balance test (RBT) was developed.⁴,⁵ This study aimed to determine the reliability of the RBT in patients with CAI.

Methods:
Forty-three eligible patients were screened on the IAC inclusion criteria for CAI. Patients visited the lab three times for familiarization and two experimental trials during which they performed the YBT and RBT on both legs. The test procedures and protocol were identical to the study of Tassignon et al. The duration and range of the test-retest time frame was chosen to reflect a clinically relevant period and lasted 22 (± 10) days on average. Three raters independently rated the different types of reliability by scoring the RBT outcome measures: accuracy and visuomotor response time (VMRT).

Results:
Twenty-seven patients with CAI were included in this study. The ICC measures for the test-retest reliability were similar for accuracy (0.609) and VMRT (0.594). Intra-rater reliability had high correlations and ICCs for accuracy (r=0.816, ICC=0.815) and VMRT (r=0.802, ICC=0.800). Inter-rater reliability had a higher ICC for VMRT (0.868) than for accuracy (0.690).

Conclusion:
Test-retest reliability was moderate, intra-rater reliability was good, and inter-rater reliability showed moderate reliability for accuracy and good reliability for VMRT. The data indicates the VMRT performance was more robust than the accuracy measure across the trials. When the Limits of agreement were compared with the minimal detectable change, the data indicates the RBT is more precise and sensitive to changes than the raters’ score. Additionally, the RBT shows robust standard error of measurement and mean difference measures.

References:

¹ a Human Physiology and Sports Physiotherapy Research Group, Faculty of Physical Education and Physiotherapy, Vrije Universiteit Brussel, Pleinlaan 2, 1050, Brussels, Belgium
² Brussels Human Robotics Research Center (BruBotics), Vrije Universiteit Brussel, 1050 Brussels, Belgium

Correspondence e-mail: alexandre.maricot@vub.be
The influence of tibial rotation on hamstring muscles groups' coordination during a dynamic knee flexion assessment.

Ferré A,1 Schwartz C,1 Bornheim S,1 Delvaux F,1,2 Croisier JL,1,2

Introduction:
Hamstring muscle (HM) strength testing is often used as a helpful means to detect strength weakness and prevent injuries. However, conventional evaluations often fail to provide a detailed understanding of the involvement of the medial (MH) and lateral (LH) hamstring muscle groups. This might lead to underestimate an abnormal MH/LH recruitment following a persistent deficit.1-3 Previous research has suggested that MR and LR could increase the activity of MH and LH, respectively.4-5 However, no study has examined the influence of these rotations on HM groups' activity or knee flexion strength during dynamic modalities.

Methods:
We aimed to address this research gap by recruiting 36 amateur male athletes and assessed their muscle activity (using Delsys Inc., Natick, MA, USA) of semitendinosus (ST) and biceps femoris long head (BFllh), as well as the knee flexion force (using Cybex; Computer Sports Medicine Inc., Stoughton, MA, USA) during concentric and eccentric contraction modes at different speeds (concentric at 60°/s, 240°/s and eccentric at 30°/s and 120°/s). We compare the EMG and strength curves between LR and MR using two-way repeated-measure ANOVA, and analyzed the data using Statistical Parametric Mapping (SPM 30, v0.4, www.spm1d.org, Matlab).

Results:
Our results showed that tibial rotations have a significant effect on knee flexion strength (p < 0.001) across most of the different dynamic modalities tested, with lower strength results associated with MR. We observed that lateral rotation consistently increased the recruitment of biceps femoris muscle and conversely (particularly during eccentric contractions). However, rotations did not statistically influence the semitendinosus muscle activity.

Conclusion:
Lateral rotation could be employed to specifically target the biceps femoris muscle (and conversely) during dynamic knee flexion strength assessment. However, the validity of tibial rotations should be further investigated in the context of hamstring injury prevention to ensure that a specific HM group deficit is not overlooked. By understanding the influence of tibial rotations on HM groups' activities and strength, clinicians' ability to identify potential weaknesses and develop targeted interventions may be improved to prevent injuries and enhance performance.

References:
Impact of stride frequency on the maximal isometric forces of the plantar flexor muscles and hip abductors muscles during an 156 km ultra-trail.

*Ayoub RT*

**Introduction:**
The main objective of this study was to assess the impact of stride frequency on the evolution of the maximal isometric forces of the plantar flexor muscles and hip abductors during an 156 km ultra-trail. A secondary objective was to analyze the evolution of the maximal isometric force of those muscles groups during an ultra-trail.

**Methods:**
The “Trail scientifique de Clécy” took place on November 11, 2021. It was about a 156 km ultra-trail divided into 6 identical loops. Each of 26 km with a total positive elevation of 6,000 m. We collected strength and running biomechanics (stride frequency) data on 55 volunteer ultra-trailers (25 to 70 years old). The measurement of the maximal isometric force for both the plantar flexors and the hip abductors was carried out using a portable dynamometer. The running biomechanics data were collected with an Optogait system. Based on the stride frequency data, we divided the sample into 2 groups (low stride frequency and high stride frequency). Those measurements were performed on the field, before, during and after the race.

**Results:**
The maximal isometric force of the hip abductors and the plantar flexor muscles decreased significantly between the beginning and the end of the race (p<0.001) as well as during the race (p<0.001). No significant differences could be observed between the mean value of the variations in the force of the hip abductors (p=0.722) nor in the mean value of the variations in the forces of the plantar flexor muscles (p=0.266) between the 2 groups.

**Conclusion:**
To our knowledge this study is the first to analyze the impact of stride frequency on the variation of the hip abductors force and the plantar flexor force during an ultra-trail. Our results suggest that the stride frequency does not impact the variation of force during such race. However, the maximal isometric force varies throughout the race. This was also reported in other scientific references studies. Therefore, it would be interesting to take the strength into account in the arising of the tiredness and the injuries and in the performance.

**References:**
Higher knee joint work is a risk factor for the development of patellar tendinopathy in volleyball: A prospective study.

De Bleecker C,1,2 Vermeulen S,1,2 Spanhove V,1 Segers V,2 Willems Tine,1 Vanrenterghem J,2 De Ridder R,1

Introduction:
Patellar tendinopathy (PT) is a common injury in sports with repetitive landing tasks such as volleyball. Impaired landing biomechanics might play a role in the development of this overuse injury. Therefore, the aim of the study was to investigate biomechanical risk factors for PT in volleyball players during sport-specific jump-landing tasks.

Methods:
In this prospective study, 82 healthy male volleyball players were evaluated during the season 2021-2022. Pre-season, three-dimensional full-body kinematics and kinetics of the push-off phase preceding the actual jump were collected during three different jump-landing tasks (spike jump, block jump and drop vertical jump). During follow-up, injury data were collected by using a weekly questionnaire and a 3-monthly retrospective control questionnaire. Univariate cox regression with competing risk analysis was used to identify significant contributors to the development of PT (p<0.05).

Results:
Of the 82 volleyball players, 10 developed PT during the follow-up period of one season (12%). The results of the study identified that increased concentric knee joint work during all jump-landing tasks (block jump p=0.01, spike jump p=0.03, drop vertical jump p=0.04) and increased eccentric knee joint work during block jump (p=0.04) are predictive parameters to develop PT in male volleyball athletes.

Conclusion:
The results of this study indicate that volleyball players with higher concentric knee joint work during different jump-landing tasks and higher eccentric knee joint work during block jump are prone to develop PT. Less knee joint work during jump-landing tasks might be beneficial for injury prevention.

References:
What do upper-extremity physical performance tests actually measure? Insights from an electromyographical study.

Borms D, Berckmans K, Lowie A, Lepla D, Maenhout L, Cools A

Introduction:
Shoulder injuries are common in overhead athletes with prevalence rates up to 42%. When sustaining an injury, the treatment approach must be based on the results of the clinical examination, including thorough analysis of possible functional impairments. In this view, physical performance tests (PPTs) focus on multijoint evaluations in which the athlete performs an activity that represents some aspects of athletic function. Evaluating the electromyographical (EMG) demands of those PPTs enables clinicians to select appropriate PPTs for their athletes.

Methods:
Thirty asymptomatic overhead athletes participated in this descriptive laboratory study. Four upper-extremity PPTs (Y-Balance Test - Upper Quarter (YBT-UQ), Closed Kinetic Chain Upper Extremity Stability Test (CKCUEST), Upper Limb Rotation Test (ULRT) and Shoulder Endurance Test (SET)), were evaluated using surface EMG on both dominant and non-dominant sides to measure muscle activity in upper (UT), middle (MT), and lower (LT) trapezius, serratus anterior (SA), infraspinatus (IS) and, posterior deltoid (PD).

Results:
During YBT-UQ performance on both sides, the supporting hand showed high SA activity levels (range: 51 – 94%MVIC) during all reach directions while IS was most active when supporting the superolateral reach (range: 92 – 129%MVIC). For the reaching hand, SA was most active (range: 46 – 83%MVIC). During the CKCUEST, all muscles were moderately to highly active, with SA (range: 64 – 87%MVIC) and IS (range: 42 – 85%MVIC) being the most active ones in both moving and supporting hand. Moderate to high activity was recorded for all muscles on both sides during the ULRT. For the SET, muscle activity progressively increased with increasing speed for both dominant and non-dominant performance.

Conclusion:
Our results provide specific EMG based information which allows clinicians to better understand PPT performance, enhancing selection of appropriate PPTs that match their patients’ needs to return to sport.

References:
Is there an association between the jump-landing profile of volleyball players and the viscoelastic properties of the patellar tendon?

Pieters D,1 De Bleecker C,1,2 Vermeulen S,1,2 Witvrouw E,1 De Ridder R,1 Evi Wezenbeek1

Introduction:
Patellar tendinopathy is highly prevalent in male volleyball players due to the repetitive jump-landing tasks. Landing kinematics have been denoted as a risk factor for the development of patellar tendinopathy. However, it is unknown whether these landing kinematics have an influence on the tendons' structural parameters such as stiffness. Previous studies showed that the viscoelastic properties of the patellar tendon are related to the injury susceptibility. Therefore, the aim of this study was to investigate a possible association between the jump-landing pattern of volleyball players and patellar tendon stiffness.

Methods:
Thirty-three male volleyball players, without a history of patellar tendon injuries and without any present lower extremity complaints, underwent a three-dimensional kinematic motion analyses of the jump-characteristics complemented by stiffness (shear wave elastography) and thickness (ultrasonography) measurements of the patellar tendon. The effect of the landing kinematics on the dominant patellar tendon stiffness was investigated with linear mixed models. The level of significance was set at $p = 0.05$.

Results:
The results of the study showed a significant association between knee flexion range of motion and tendon stiffness ($p = 0.022$). The lower the knee flexion range of motion, the higher the tendon stiffness.

Conclusion:
This study was the first to show an association between the jump-landing pattern of volleyball players and the stiffness of the patellar tendon. More specifically, the lower the knee flexion range of motion during landing, the higher the patellar tendon stiffness. These results suggest that the specific jump-landing pattern of volleyball players impacts the patellar tendon structure, co-determining the injury predisposition.

References:
Does baseline cognitive performance predict the response to mental fatigue in healthy individuals?

Schampheleer E, Habay J, Proost M, Roelands B

Introduction:
Mental fatigue (MF), a psychobiological state induced by prolonged demanding cognitive activity, is implied to cause an increased injury risk in a healthy population. However, the level of emergence of MF is highly variable between different subjects, making it difficult to screen which people are more sensitive to its effects. Meanwhile, almost no factors have reliably been identified that can predict the response to MF. Since cognitive processes, such as attention and decision-making, are crucial for physical activities, this study aimed to investigate the link between cognitive abilities and the response to mental fatigue.

Methods:
We employed a randomized single-blinded placebo-controlled counterbalanced cross-over design. First, participants completed three cognitive tests (sustained attention to response task, psychomotor vigilance task, and N-BACK task) to measure attention, working memory, and response inhibition. During the experimental and control trial, participants completed either a 45-minute modified Stroop task or watched a documentary of the same duration. Before and after the experimental and control trial, all 48 participants were asked to rate their feeling of MF on a visual analogue scale (M-VAS). After the Stroop task/documentary, participants performed a Go-NoGo task and a 15-minute time trial to assess cognitive and physical performance. Linear regression was used to evaluate the relationship between cognitive performance and MF effects.

Results:
Baseline cognitive functions did not significantly relate to differences between scores on the M-VAS (F=2.126; p=.094). No significant relationship was found between baseline cognitive functions and the extent to which MF affects physical performance (F=1.315, p=.286). Finally, a marginally significant relationship was found between baseline cognitive performance and accuracy scores on the Go stimuli of the Go-NoGo task when participants are mentally fatigued (F=2.485; p=.063) but not for reaction time (F=.664; p=.621) or accuracy on the NoGo stimuli (F=.629; p=.646). Conclusion: It is of the utmost importance to identify athletes who are more susceptible to MF, as this could help us in developing effective injury prevention strategies. The present study revealed no significant relationship between baseline cognitive performance and susceptibility to MF. More research is needed to fully elucidate the individual response to MF.

References:
The additional value of Ischemic preconditioning during warm-up on isokinetic strength and endurance

Jacobs E, Schuermans J, Wezenbeek E, Witvrouw E

Introduction:
Ischemic preconditioning (IPC) has become an upcoming topic within blood flow restriction (BFR) literature. A positive effect of IPC on muscle strength has already been suggested by several studies, but there is yet no conclusive evidence concerning the effect of IPC on isolated strength parameters compared to regular warm-up in a larger population.

Methods:
Thirty-three healthy participants attended two sessions of performing a maximal isokinetic strength test with a different warm-up. The control session comprised a five-minute warm-up on a stationary bike, whereas in the intervention session two bouts of five minutes of IPC preceded the strength test. Readiness to perform (RTP) was questioned after the warm-up. Quadriceps and hamstring strength parameters, including average and maximum peak torque/body weight (PT/BW) and work fatigue (WF), as well as rate of perceived exertion (RPE) and rate of perceived discomfort (RPD) were obtained after each test. The presence of delayed-onset muscle soreness (DOMS) was questioned 24 hours post-session.

Results:
The average and maximum PT/BW were significantly lower in the IPC-session for both the quadriceps (p = 0.024; p = 0.005, respectively) and hamstrings (p = 0.015; p = 0.007, respectively). Significantly lower scores were also found for RTP before, and DOMS after the IPC-session (p < 0.001). No significant differences between IPC and control were found for WF, RPE and RPD.

Conclusion:
Passive BFR applied as IPC does not appear to enhance muscle strength or endurance in terms of peak torque or work fatigue, on the contrary. Two 5-minute bouts of IPC-application 10 minutes prior to a maximal isokinetic strength test reduced the average and peak torque while simultaneously reducing the readiness to perform compared to a conventional cycling warm up.

References:
Countermovement jump may determine performance alterations after anterior cruciate ligament reconstruction.

Forelli F,1,2,3,4 Nekhoufi B,1 Vanderbrouck A,2 Duffett P,2 Ratte L,2 Hewett TE,5 Rambaud AJM,4,6

Introduction:
The main objective was to examine countermovement jump (CMJ) measures to identify which parameters can best distinguish between anterior cruciate ligament reconstruction (ACLR) and control participants. The secondary objective was to determine whether performance alterations between operated and non-operated limb exist during vertical two-legged activities after ACLR.

Methods:
This case control study included 67 patients with hamstring graft at 6 post-operative months (203.5 days ± 32.2) and 47 healthy athletes with no knee injury history. Two groups were formed, an ACLR group (n=67) and a control group (n=47). An evaluation of CMJ by force plate was performed, to calculate vertical ground reaction force (vGRF), maximal power (MP) and eccentric rate force development (RFDe) during landing and limb symmetry index (LSI). (1–3) First analysis compared LSI vGRF, LSI MP and LSI RFDe between both groups during CMJ. Secondary analysis compared vGRF, MP and RFDe between operated/non-operated limb in the ACLR group and dominant/non-dominant limb in the control group.

Results:
CMJ measures in the ACLR group were significantly reduced compared to the control group for LSI vGRF (85.9% ± 9.6 vs 94.6% ± 5.3, p < 0.001, respectively), LSI MP (84.8% ± 8.4 vs 95.6% ± 4.1, p < 0.001, respectively) and LSI RFDe (68.0% ± 23.1 vs 76.7% ± 17.2, p < 0.001, respectively). Secondary analysis showed no significant result in control group between dominant/non-dominant limb. ACLR group showed significant results between operated / non-operated limb for PT (9.4 N.kg-1 ± 0.1 vs 10.8 N.kg-1 ± 0.13, p < 0.001, respectively), MP (17.7 W.kg-1 ± 4.0 vs 20.3 W.kg-1 ± 4.1, p < 0.001, respectively) and RFDe (825.2 N.s-1 ± 62.0 vs 1200.5 N.s-1 ± 87.8, p < 0.01, respectively).

Conclusion:
The results indicate significant torque, power and landing deficits and performance alterations on the non-operated limb during CMJ at time to return to sport after ACLR. The results may be too preliminary to draw definitive conclusions but double legged assessment should be considered in return to sport decision making after ACLR.

References:
Early open kinetic chain improves strength recovery and return to sport after anterior cruciate ligament reconstruction without graft laxity increasing

Forelli F,1,2,3,4 Wassim B,1,2 Kersante G,1,2 Vanderbrouck A,2 Duffett P,2 Ratte L,2 Hewett TE,3 Rambaud AJM,4

Introduction:
The main objective was to determine whether the early associated use of open kinetic chain (OKC) and closed kinetic chain (CKC) improved quadriceps and hamstring strength in the rehabilitation after anterior cruciate ligament reconstruction (ACLR). The secondary objective was to assess whether the early use of OKC had an influence on graft laxity at 3 and 6 postoperative months.

Methods:
This study included 103 patients with hamstring graft. Two groups were formed: OKC+CKC group (n = 51) vs CKC group (n = 52). OKC protocol which included exercises for quadriceps and hamstrings muscles, were introduced at 4 weeks after ACLR (31.4 days ± 7.6). At 3 months (101.9 days ± 18.4) and 6 postoperative months (199.2 days ± 28.1), an evaluation of muscle strength by isokinetic dynamometer was performed, to calculate peak torque-to-body weight ratio (PT/BW) for the quadriceps and hamstrings. The laxity measurement was performed by comparative measurements performed by GNRB.

Results:
At 3 and 6 postoperative months, quadriceps strength in the OKC+CKC group was higher than in CKC group for LSI (76.1 % ± 0.21 vs 46.9 % ± 0.21, p < 0.001 and 91 % ± 0.17 vs 61.8 % ± 0.26, p < 0.001, respectively) and PT/BW (1.81 Nm.kg⁻¹ ± 0.75 vs 0.85 Nm.kg⁻¹ ± 0.50, p < 0.001 and 2.40 Nm.kg⁻¹ ± 0.73 vs 1.39 Nm.kg⁻¹ ± 0.70, p < 0.001, respectively). There were similar findings for the hamstring strength: LSI (86.1 % ± 0.21 vs 64.3 % ± 0.24, p < 0.001 and 91.9 % ± 0.17 vs 82.4 % ± 0.24, p < 0.001, respectively) and PT/BW (1.09 Nm.kg⁻¹ ± 0.36 vs 0.69 Nm.kg⁻¹ ± 0.39, p < 0.001 and 1.41 Nm.kg⁻¹ ± 0.41 vs 1.06 Nm.kg⁻¹ ± 0.39, p < 0.001, respectively). At 3 months no difference was observed for laxity between OKC+CKC and CKC group (p = 0.48). At 6 months the laxity was greater in CKC group (p = 0.31).

Conclusion:
The results indicate that early associated use of OKC and CKC allow for enhanced correction of quadriceps and hamstrings strength deficits and readiness to return to sport without increasing graft laxity.

References:
Gastrocnemius muscles activity increasing may impair running patio-temporal parameters after anterior cruciate ligament reconstruction: A pilot study.

Forelli F,1,2,3,4 Pengue-Koyi A,1,2 Mazeas J,1,2 Hewett TE,5 Rambaud AJM,6

Introduction:
The main objective of this study was to evaluate whether running after anterior cruciate ligament reconstruction (ACLR) resulted in increased muscular activity of the gastrocnemius medialis and gastrocnemius lateralis compared to running in healthy participants. The secondary objective was to assess whether these changes in muscular activity correspond to changes in cadence, vertical stiffness, flight time and ground contact time while running.

Methods:
This pilot case-control study included 7 patients with hamstring graft at 6 postoperative months (208.7 days ± 34.6) and 8 healthy athletes with no knee injury history. Two groups were formed, an ACLR group (n=7) and a control group (n=8). After maximal voluntary isometric contraction (MVIC) assessment, both groups performed treadmill running assessment with Optogait®. After a 6 minutes warm up on a treadmill at 10 km.h⁻¹, (1) 30 sec were recorded to measure the surface electromyographical activity (EMG) of the GM and GL. (2,3) After root mean square (RMS) treatment of the raw signal, RMS EMG results were normalized by MVIC activity to allow inter-subject comparability.

Results:
Between-group analyses showed a significant increase in RMS EMG for the ACLR group compared to the control group for the GM (34.7%MVIC ± 11.0 vs 25.5%MVIC ± 13.0, p = 0.05, Effect Size = 0.52) and the GL (32.8%MVIC ± 10.6 vs 17.2%MVIC ± 6.30, p < 0.01, Effect Size = 0.78). Significant correlations were observed in the ACLR group with GL RMS EMG for ground contact time (r = 0.84; p = 0.02). However, there were no significant correlations with cadence (r = 0.50; p = 0.27), vertical stiffness (r = 0.50; p = 0.27) and flight time (r = 0.02; p = 0.97). No significant correlations were observed in the ACLR group with GM RMS EMG.

Conclusion:
These findings indicate that ACLR subjects presented with higher GM and GL activity while running compared to the control group. The overuse of these muscles may play a role in the alteration of spatiotemporal parameters of running after ACLR.

References:
Superior foot & ankle muscle strength in non-rearfoot endurance runners compared with rearfoot runners: implications for the management of running-related knee injuries.

Abran G,1,2 Schwartz C,1 Delvaux F,1,2 Borhneim S,2 Aguilaniu A,1,2 Croisier Jean-Louis,1,2

Introduction:
Transitioning to a forefoot strike pattern can be used as a component of a gait retraining intervention to manage running-related knee injuries. However, adopting a non-rearfoot strike induces a higher load on foot and ankle structures than rearfoot strike. In response to these biomechanical differences, non-rearfoot runners (NRF) appear to have a superior ankle plantar flexor strength and Achilles tendon cross-area compared with rearfoot runners (RF). Sufficient foot muscle strength is also necessary to prevent excessive longitudinal arch (LA) deformation when running with non-rearfoot strike. The aim of this study was to investigate the difference in foot-ankle muscle strength between RF and NRF.

Methods:
Forty RF and forty NRF were recruited. A navicular drop, a foot posture index and the maximal voluntary isometric strength (MVIS) of six foot-ankle muscles were measured. The footstrike pattern was determined using a high-speed camera during a self-paced run on a treadmill.

Results:
NRF had higher MVIS for ankle plantar flexor (+12.5%, p = 0.015), ankle dorsiflexor (+17.7%, p = 0.01), hallux flexor (+11%, p = 0.04) and lesser toe flexor (+20.8%, p = 0.0031). NRF also had stiffer plantar arches (p = 0.04) and less pronated feet (p = 0.02). There is a small positive correlation between MVIS of ankle plantar flexor with MVIS of ankle invertor (r = 0.22; p-value = 0.04), hallux flexor (r = 0.26; p-value = 0.01) and lesser toe flexor (r = 0.28; p-value = 0.01).

Conclusion:
The main finding of this research is the higher MVIS of hallux and lesser toe flexor in NRF compared with RF, despite a wide range of values. NRF also have a higher MVIS of ankle plantar flexor and dorsiflexor than RF. There is only a small correlation between ankle plantar flexor and foot muscle strength. Consequently, clinicians who practice gait retraining interventions including transition to a forefoot strike pattern should assess foot and ankle muscle strength separately. Then, they should implement, if necessary, a foot strengthening program to prevent excessive LA deformation.

References:
Neurocognitive deficits related to ligamentous ankle injuries and chronic ankle instability
Corluy H, Schampheleer E, Maricot A, Verschueren J, Roelands B, Tassignon B

Introduction:
The ankle is the most commonly injured body part in sports and is often subject to recurrent injury, especially ligament sprains. Up to 40% of ankle sprains may develop chronic ankle instability (CAI). Despite growing evidence that altering neurocognitive demands can affect lower limb biomechanics in individuals with CAI, the underlying mechanisms remain unclear. The aim of this systematic review was to summarise the current literature on the neurocognitive deficits linked with ligamentous ankle injuries and CAI.

Methods:
Five electronic databases were used, including PubMed, Web of Science, Scopus, PsychInfo and SPORTDiscus from their inception to February 22nd, 2023. Articles were eligible if they (1) were published in English, (2) were original research and (3) investigated neurocognitive functioning in 18-year-old or older patients with CAI or who experienced a lateral ankle sprain. The methodology followed the PRISMA guidelines. We grouped the neurocognitive functions into eight domains: executive functions, information processing speed, inhibitory control, attention, reaction time, visual spatial perception, motor control, and memory.

Results:
A total of 1221 results were identified, of which 18 studies met the inclusion criteria. The risk of bias assessment indicated an overall high risk of bias in the studies. Among these included studies, 335 individuals with CAI were included, 244 healthy controls, and 68 copers. Three studies highlighted that reaction time was significantly worse in individuals with CAI compared to copers and healthy controls. Regarding attention, three out of six studies suggest that these cognitive functions may be impaired in individuals with CAI. Also for memory, three out of six studies found deficits in individuals with CAI. The remaining five neurocognitive domains showed either inconclusive or no results in individuals with CAI.

Conclusion:
Overall, individuals with CAI appear to have neurocognitive deficits in reaction time, attention and memory, while other neurocognitive domains do not seem to be affected. Nevertheless, there are still few studies on the different neurocognitive subdomains in this population which highlights the need for further research to better map and understand this phenomenon and its underlying mechanisms.

References:
Peak patellar tendon force during heavy load single-leg squatting is influenced by a decline board but not by the external weight's mass


Introduction:
Heavy slow resistance training is an important stage in progressive exercise programs for the rehabilitation of patellar tendinopathy. An exercise frequently included this training, is the single-leg squat either on level ground or on a decline board. Unfortunately, the impact of variations of the heavy load single-leg squat on the peak patellar tendon force (PTF) has not yet been objectively quantified. The objective of this study was therefore to investigate the influence of the mass of an external weight and the use of a decline board on the peak PTF during a heavy load single-leg squat.

Methods:
Twelve healthy participants with at least one year of strength training experience were included. The participants performed single-leg back squats on a decline board and on level ground with external weights of 70%, 80% and 90% of their respective one-repetition maximum. Three-dimensional kinematics were collected using a passive marker-based motion capture system (100Hz; Vicon, Oxford, UK) and ground reaction forces were measured using a ground-embedded force plate (1000Hz; AMTI, Watertown, USA). Peak PTF was calculated in OpenSim (Stanford, USA) using the Catelli model and a static optimisation approach. A two-way repeated measures ANOVA determined the main effect for the mass of the external weight, the main effect for the surface, as well as for their interaction effect on the peak PTF.

Results:
Peak PTF values were significantly higher on the decline board compared to level ground (p(surface) = 0.025). Neither on the decline board nor on level ground did an increase in the mass of the external weight result in a significant increase of the peak PTF (p(mass) = 0.100; p(mass)*surface = 0.090).

Conclusion:
Progression in peak PTF during a single-leg squat can be obtained with a decline board. Increasing the mass of the external weight from 70% to 90% of the one-repetition maximum will not result in a higher peak PTF on either surface. Therefore, it can be concluded that it is possible to expose patients with patellar tendinopathy to high peak PTF loads during rehabilitation, even at lower heavy weights, simply by performing the single-leg squats on a decline board.

References:

(1) Department of Rehabilitation Sciences, KU Leuven, Leuven, Belgium
(2) Department of Movement Sciences, KU Leuven, Leuven, Belgium.
Correspondence e-mail: michiel.hagen@kuleuven.be
Physiotherapists and physiotherapy students estimate patellar tendon forces during rehabilitation exercises equally well

Introduction:
Exercise therapy with gradual progression of patellar tendon load is key for the rehabilitation of patellar tendinopathy. Unfortunately, objective quantification of patellar tendon forces (PTF) in clinical practice is impossible. Therefore, physiotherapists rely on clinical reasoning to estimate PTF for individual patients. The objective of this study was to examine whether physiotherapists and graduating physiotherapy students, specialised in musculoskeletal rehabilitation, can accurately estimate PTF during rehabilitation exercises.

Methods:
One healthy individual performed 19 rehabilitation exercises twice. Exercises included gait, squats, lunges, steps, calf raises, single-leg Romanian deadlift and jumps. The first time, marker-based motion capture data (100Hz; Vicon, Oxford, UK) and ground reaction forces (1000Hz; AMTI, Watertown, USA) were collected. PTF was calculated in OpenSim (Stanford, USA). The second time, videos of the exercises were recorded in a clinical setting. These videos were implemented in an online survey. Here, participants were asked to score each exercise on a 100-point numerical rating scale ranging from “absolutely no PTF” to “maximal PTF over all exercises”. The survey was completed by 50 physiotherapists and 50 second-master students. Two statistical analyses were performed. First, the differences between the participants’ scores and the objective scores were calculated. Both the average difference over all exercises and the difference for each individual exercise were analysed using one-sample t-tests (critical p-value = 0.01). Next, the scores were reduced to rankings, representing a more basic level of clinical reasoning. The correlation between the participants’ rankings and the objective rankings was analysed using Kendall’s tau coefficient.

Results: In both groups, the overall difference between subjective and objective PTF scores was not significantly different from zero (both p > 0.01). At the level of individual exercises, both groups estimated eight exercises correctly, underestimated five exercises and overestimated six exercises. Except for one exercise, the same exercises were over- or underestimated. Regarding the rankings, both physiotherapists (tau = 0.626, p < 0.001) and students (tau = 0.614, p < 0.001) had moderate to good correlations with the objective PTF rankings.

Conclusion: Physiotherapists and physiotherapy students have a moderate to good capacity to estimate PTF during rehabilitation exercises. The capacity to estimate PTF is not influenced by clinical experience as both groups scored equally well.

References:
Does sleep quality predict VO$_2$max in healthy adults?

Ezzy Q, Habay J, Roelands B

Introduction:
It has previously been discovered that the amount of daily vigorous activity is associated with an increased VO$_2$max in adults. Additionally, a relation was found between increased physical activity and overall sleep quality. However, only a limited amount of studies have explored how VO$_2$max affects sleep quality. This present study intends to builds upon these findings by investigating whether VO$_2$max predicts sleep quality, and whether sleep quality is affected by physical performance level in healthy adults.

Methods: Sixty-seven healthy adults (33 males and 34 females) aged 31.39 ± 8.92 years old, with a VO$_2$max of 47.29 ± 8.42 ml/kg^-1/min^-1, and performance level of 2,10 ± 0.82 were included in this study. Subjects performed a maximal incremental exercise test on a cycling ergometer, starting at a power output of 80W and increasing by 30W every three minutes. The Pittsburgh Sleep Quality Index (PSQI) was used to assess sleep quality. This questionnaire evaluates 19 self-reported items divided in seven subcategories: 1) subjective sleep quality; 2) sleep latency; 3) sleep duration; 4) habitual sleep efficiency; 5) sleep disturbances; 6) use of sleeping medication; and 7) daytime dysfunction. A simple linear regression was performed to investigate whether VO$_2$max significantly predicted sleep quality and a one-way ANOVA for investigating whether there were any differences in sleep quality depending on performance level.

Results:
The fitted regression model (Sleep quality = 0.017 * VO2max + 15,821) found no significant results (R$^2$ = 0.004, F = 0.258, p = 0.613) when investigating whether VO$_2$max predicts sleep quality. Additionally, no differences in sleep quality were found for groups of different performance levels (F = 0.043, p = 0.988).

Conclusion: This research shows that VO$_2$max does not predict sleep quality and that other factors will be more decisive at determining sleep quality. These findings further contribute to the understanding of factors determining sleep quality in healthy adults.

References:
Hamstring muscle fibre type distribution in football players having suffered an anterior cruciate ligament Injury: harder, better, faster, stronger? A case-control study using MR Spectroscopy

Denolf S,1 Witvrouw E,1 Tampere T,1 Schuermans J,1

Introduction:
Anterior cruciate ligament (ACL) Injuries are one of the most common acute sports injuries in football. Amongst others, architectural features as bone morphology, knee joint configuration and lower limb alignment have been associated with the ACL injury risk. Recently, the role of muscle fibre type distribution in the light of both athletic performance and hamstring muscle injury risk, has gained popularity in sports medicine research. As the ACL is a biomechanical synergist of the hamstring muscle unit and football players with a history of ACL injury are more prone to hamstring injuries and vice versa, this study intended to verify to what extent hamstring muscle fibre type distribution is related to the risk of ACL injury in football players, by conducting a case control study using magnetic resonance spectroscopy (MRS).

Methods:
21 amateur football players with a recent history of ACL injury and 45 matched controls were submitted to a MRS evaluation. The Semitendinosus muscle of the non-injured leg was used for carnosine content calculations in the ACL history group. In the control group dominant and non-dominant legs were chosen randomly based on respective distribution in the ACL history group. Muscle fibre type distribution was estimated based on the relative carnosine content (expressed in Arbitrary Units (AU)), and the associated Z-scores were used to classify participants having mostly Fast Twitch (FT), Intermediate Type (IT) or Slow Twitch (ST) fibre presence.

Results:
Mixed models analysis revealed that carnosine contents did not differ significantly based on ACL history presence, with average carnosine contents of 0.19 ± 0.054 and 0.17 ± 0.051 AU in the ACL injury and control groups, respectively (p = 0.145). FT, IT and ST dominances was found in 38%, 29% and 33% of participants in the ACL group and 24%, 40% and 36% in the control group, respectively.

Conclusion:
ACL injuries do not present any association with hamstring muscle fibre type. Nonetheless, given the large range and high variability in fibre distribution in this population, recovery needs might differ essentially, making certain players more prone to overload and fatigue related injuries than others.

References:
2. Schuermans J, et al. Hamstring muscle fibre typology is not associated with hamstring strain injury history or performance in amateur male soccer players: a retrospective magnetic resonance spectroscopy study. Biol Sport. 2023. Accepted for publication and in press
Fatigue-induced biomechanical risk factors for patellar tendinopathy in volleyball: a prospective study.

Vermeulen S,1,2 De Bleecker C,1,2 Spanhove V,1 Segers V,1 Willens T,1 Roosen P,1 Vanrenterghem J,2 De Ridder,1

Introduction:
Patellar tendinopathy (PT) is a highly prevalent overuse injury in volleyball. However, little is known about whether and how fatigue may increase the risk for developing PT through biomechanical alterations during repetitive jump-landing activities in volleyball. Therefore, the objective of this study was to identify fatigue-induced biomechanical risk factors for PT in volleyball during a spike jump-landing.

Methods: Eighty-two male volleyball players were tested pre-season in 2021 with subsequent prospective follow-up for one season. At baseline, three-dimensional full-body kinematics and kinetics were collected while performing a spike jump before and after a volleyball-specific fatigue protocol. Univariate cox regression with competing risk analysis was performed to identify significant predictors for the development of PT (p<0.05).

Results: During follow-up, 13 of the 82 players developed PT (16%). For the fatigued biomechanical variables, decreased hip flexion (-7.8°, p=0.026, HR=1.057), increased patellar tendon loading rate (+6.2 x body weight/s, p=0.043, HR=1.049) and increased length of the rectus femoris (+1.3 cm, p=0.005, HR=1.826) and vastus lateralis muscle-tendon unit (+0.3 cm, p=0.048, HR=4.034) were significant contributors for developing PT.

Conclusion: The results of this prospective study suggest that players who utilize a stiffer landing at the hip after fatigue may have an increased risk for developing PT due to the accumulation of tensile forces acting on the patellar tendon.

References: