

Fiabilidade Intra-Observador, Erro de Medida e Mudança Mínima Detectável do *Weight-Bearing Lunge-Test* e do Teste de Deslizamento Posterior do Astrágalo em Indivíduos com História de Entorse do Tornozelo

Maria do Paço¹, Eduardo Brazete Cruz²

Abstract

Objective: To evaluate the reliability and identify the standard error of measurement (SEM) and the minimal detectable change (MDC) of the weight-bearing lunge-test and the posterior talar glide test. **Methods:** A test-retest study with 16 individuals reporting history of at least one ankle sprain. Participants were recruited from a pool of university students to whom the above tests were applied, by the same examiner, in two different moments. **Resultados:** This study found an intra-class correlation coefficient (ICC_(3,3)) of 0,994 (95% confidence interval = 0,982 - 0,998, and a SEM of 0,18 cm) for the weight-bearing lunge test and an ICC_(3,3) of 0,995 (95% confidence interval = 0,987 - 0,998 and a SEM of 0,16°) for the posterior talar glide test. These results are lightly superior to others found in other studies (Collins et al. 2004; Vicenzino et al. 2001; Bennel et al. 1998) and show that these tests have high reliability and low systematic error. **Conclusions:** Our data suggests that both tests can be used as useful and easily applied instruments, to identify restrictions at the dorsiflexion range of motion and posterior talar glide. The MDC' results will help the clinical interpretation of treatment changes on subjects, allowing to distinguish the values that represent a true change from those due to test measurement errors.

Key Words: Ankle; range of movement; measurement; minimal detectable change

Escola Superior de Saúde do Vale do Sousa¹
Escola Superior de Saúde de Setúbal²

Correspondência:
ma_paco@hotmail.com

Introdução

A entorse do tornozelo é uma das lesões ortopédicas mais frequentes entre a população fisicamente activa (Hoch & McKeon, 2010; van Rijn et al., 2008; Hubbard & Hertel, 2006; Hertel, 2000). Os indivíduos com história de entorse referem frequentemente instabilidade do tornozelo e diminuição da capacidade funcional (Hertel, 2000), tendo sido identificados vários défices funcionais e mecânicos que poderão contribuir para os sintomas residuais, sequelas e perdas funcionais. Um desses défices é a limitação da amplitude de movimento (ADM) de flexão dorsal (FD) (Hoch & McKeon, 2010; Hubbard & J. Hertel, 2006; Denegar & Miller, 2002a; Green, Refshauge, Crosbie, & Adams, 2001; Vicenzino, Prangley, & Martin, 2001; Hertel, 2000), sendo esta um movimento essencial para actividades como a marcha, que requer uma FD entre os 10° e 20°, entre outras (Vicenzino et al., 2001; Tabrizi, McIntyre, Quesnel, & Howard, 2000).

Clinicamente, a restrição deste movimento está também associada a alterações mecânicas na articulação do tornozelo, sendo considerada um dos factores predisponentes para a ocorrência de nova lesão (Hadzic et al., 2009; de Noronha, Refshauge, Herbert, Kilbreath, & Hertel, 2006; Tabrizi, et al., 2000; Crosbie, Green, & Refshauge, 1999;).

Para avaliar estas restrições de movimento, bem como os efeitos do tratamento efectuado, é necessário que o Fisioterapeuta utilize instrumentos ou testes, objectivos e fiáveis. A literatura refere diferentes métodos para avaliar a ADM FD. Uns mais simples, embora apresentem também valores de fiabilidade mais baixos como o goniómetro universal ou a estimativa visual e outros métodos mais sofisticados que exigem materiais mais dispendiosos e nem sempre existentes em contexto clínico como a utilização do goniómetro eléctrico ou a fotografia de marcadores na pele com consequente medição de uma projecção da imagem, entre outros (Gatt & Chockalingam, 2011; Rome, 1996).

Pelo contrário, o *Weight-Bearing Lunge Test* (WBLT) é apresentado como uma medição de fácil aplicabilidade clínica, eficiente em termos de custo e tempo gasto, que requer um equipamento muito simples, e que é realizado em carga (Hoch & McKeon, 2011;

Bennel, Talbot, Wajswelner, Techovavanich, & Kelly, 1998). Em situações em que não existem contra-indicações, o facto de ser realizado em carga, poderá torná-lo mais representativo da ADM disponível na realização de actividades funcionais.

Para além deste teste, e uma vez que alguns autores atribuem a limitação na ADM FD a uma alteração mecânica do astrágalo, defendendo que uma restrição na artrocinemática do astrágalo poderá conduzir a uma limitação no movimento fisiológico de FD (Hoch & McKeon, 2011; Denegar & Miller, 2002a; Green, Refshauge, Crosbie, & Adams, 2001), a utilização de um método que permita avaliar o deslizamento posterior do astrágalo (DPA), verificando assim a contribuição desta estrutura para a restrição do movimento encontrada, pode igualmente constituir-se uma ferramenta clínica de grande utilidade.

O WBLT tem sido objecto de estudos de fiabilidade tendo revelado valores de fiabilidade intra-observador elevados e valores de erro standard de medição baixos (Collins et al. 2004; Vicenzino et al., 2001; Bennel et al., 1998). Também nos estudos de fiabilidade intra-observador do DPA têm sido encontrados valores de fiabilidade intra-observador razoáveis a elevados (Vicenzino et al., 2006; Denegar & Miller, 2002a), e de SEM reduzidos, não se conhecendo no entanto, os valores destes testes em estudos realizados na população portuguesa.

Para além destas propriedades e para a interpretação dos efeitos de tratamento é igualmente importante identificar a mínima mudança detectável (MMD), de forma a distinguir os valores obtidos no teste do erro de medida. A mudança mínima detectável é definida como a quantidade mínima de diferença que não é atribuível à variação na medida (Haley & Fragala-Pinkham, 2006).

Assim, por ser sugerido na literatura que a identificação destes factores na avaliação clínica é essencial para que estes possam ser revertidos, e uma vez que a identificação da MMD indica a verdadeira melhoria da amplitude de FD ou de deslizamento posterior do astrágalo, respectivamente, e não a um erro de medida, e não existindo, do nosso conhecimento, qualquer estudo realizado em indivíduos com história de entorse do tornozelo nesse âmbito, propusemo-nos a avaliar a fiabilidade intra-observador, o erro padrão de medida e a mudança mínima detectável para o WBLT e para o DPA.

Metodologia

Amostra

Para a realização do estudo utilizou-se uma amostra de indivíduos saudáveis seleccionados a partir do universo de estudantes da Escola Superior de Saúde do Vale do Sousa e da Escola Superior de Saúde do Vale do Ave, segundo um método de amostragem não probabilístico, de conveniência. Com base na revisão da literatura estabeleceram-se os seguintes critérios de inclusão: Idade mínima de 18 anos (por conveniência); História de pelo menos uma entorse unilateral por inversão (Reid, Birmingham, & Alcock, 2007); Ausência de lesão no tornozelo nos últimos 3 meses, uma vez que se pretendia garantir a não presença de sinais inflamatórios que pudessem condicionar a avaliação (Gribble, Taylor, & Shinohara, 2010; Hopper et al., 2009); Ausência de história de lesão, fractura ou cirurgia no membro inferior de forma a garantir-se que não existiriam outros factores a contribuir para as alterações/ limitações que os indivíduos pudessem apresentar, e que pudessem igualmente condicionar a intervenção (Gribble, Taylor, & Shinohara, 2010; Delahunt, O'Driscoll, & Moran, 2009; Reid, Birmingham, & Alcock, 2007; Monaghan, Delahunt, & Caulfield, 2006; Vicenzino, Branjerdporn, Teys, & Jordan, 2006; Denegar, Hertel, & Fonseca, 2002b). Os potenciais participantes tomaram conhecimento do estudo através de anúncios colocados nas instituições referidas, anúncios esses que continham a informação relativa aos critérios de inclusão e exclusão do estudo, bem como uma descrição generalizada dos procedimentos e tempo médio que despenderiam com a participação no estudo. Do universo de estudantes das referidas escolas, voluntariaram-se 86 indivíduos para participar no estudo. Todos preencheram um questionário de forma a verificar se cumpriam os critérios de inclusão, com base no qual se seleccionaram 16 indivíduos, cujas características demográficas se apresentam na tabela abaixo (tabela 1).

Tabela 1. Características demográficas dos participantes.

	Frequência (n)	Percentagem (%)	Média	Desvio Padrão
Género				
Masculino	5	31,3		
Feminino	11	68,8		
Idade			21,69	3,38

Instrumentos

O WBLT mede indirectamente a FD através da medida do avanço da tibia sobre a zona posterior do pé com o indivíduo numa posição de carga Bennel et al. (1998). O teste apresenta valores de fiabilidade intra-observador elevados [coeficiente de correlação intraclasse (CCI)_(3,3)=0,97] e valores de erro padrão de medição baixos [erro padrão de medição (EPM)=0,6 cm], valor que revela um erro sistemático mínimo. Por sua vez, o teste de DPA é um teste que avalia indirectamente o movimento/ restrição do

movimento do astrágalo. Considera-se que o DPA pode ser estimado tendo por base a flexão passiva do joelho (Denegar & Miller, 2002a).

Procedimentos

Todos os participantes receberam informação acerca da natureza e procedimentos do estudo e preencheram uma declaração de consentimento informado. Foi garantido que não corriam quaisquer riscos e que poderiam desistir da participação no estudo a qualquer momento, sem que dessa decisão adviesse qualquer prejuízo para os próprios. Foi ainda garantida a confidencialidade e anonimato dos dados recolhidos.

Para a realização dos testes, e com base em estudos anteriormente realizados, foram seleccionados dois protocolos, um para cada teste. Para realizar o WBLT, o participante deveria colocar-se numa posição estandardizada em que a linha média do calcâneo, o *hállux* e o joelho estivessem alinhados e numa posição perpendicular à parede. Para garantir esta posição, a linha média do calcâneo assim como o *hállux*, foram alinhados pela fita métrica colocada no chão. Nesta posição o participante foi instruído a inclinar-se para a frente devendo o joelho tocar uma linha vertical desenhada na parede, de forma a manter-se o alinhamento do segmento. Mantendo o contacto do joelho com a parede e do calcâneo com o chão (que foi assegurado por comandos verbais e por contacto manual do avaliador) foi medida a distância máxima entre o *hállux* e a parede que o participante conseguiu atingir (Figura 1). Foi dada a indicação aos participantes de que poderiam apoiar-se na parede durante o teste e que poderiam ter o membro contralateral numa posição confortável (Bennel, Talbot, Wajswelner, Techovavanich, & Kelly, 1998).

Figura 1: Avaliação da ADM FD, através do WBLT



Para avaliar o DPA o participante foi instruído a sentar-se na marquesa e relaxar. De seguida foi colocado o inclinómetro no 1/3 proximal da perna, lateralmente à tuberosidade anterior da tíbia. Nesta posição o avaliador colocou a articulação subastragalina numa posição neutra, que foi mantida enquanto aplicava um deslizamento antero-posterior no astrágalo induzindo flexão dorsal, até encontrar um fim de movimento capsular firme. Assim que o avaliador encontrou uma restrição no movimento (fim do movimento capsular firme), registou o ângulo de flexão do joelho (Denegar, Hertel, & Fonseca, 2002b) (Figura 2).

Figura 2. Avaliação do DPA



Previamente à execução dos testes o avaliador realizou um período de treino de 10 minutos para se familiarizar com os procedimentos do estudo. As medições da amplitude de movimento de FD e do DPA foram efectuadas três vezes em cada momento (avaliação inicial e reavaliação), sendo calculado o valor médio das três medições. Dada a proximidade dos dois momentos de recolha de dados, imediatamente após o período de avaliação inicial realizou-se um período de wash-out em que os participantes eram livres de se movimentarem normalmente durante 10 minutos. Após esse período cada indivíduo foi novamente avaliado com os dois testes, pelo mesmo avaliador. Os dados recolhidos permitiram testar a consistência nas medições do examinador (CCI), quantificar o erro existente em cada medição (EPM) e conseqüentemente, calcular a diferença mínima que deverá ocorrer para que essa diferença não corresponda a um erro do examinador (MMD) (Portney & Watkins, 2009).

Análise Estatística

No sentido de se verificar a existência ou não de diferenças entre os dois momentos de avaliação no WBLT e no DPA, realizou-se o teste paramétrico de T' Student para amostras emparelhadas e o teste não paramétrico de Wilcoxon, respetivamente. Para o cálculo da fiabilidade intra-observador, foi calculado o Coeficiente de Correlação Intraclasse $(_{3,3})$, comparando os resultados dos dois momentos de recolha de dados (avaliação inicial e nova avaliação após 10 minutos). Foi escolhido o $CCI_{(3,3)}$ uma vez que foram realizadas 3 medições em cada momento e pelo facto de ser um único elemento a avaliar e que constituiu o único elemento de interesse (Russek, 2004). Foram ainda calculados os intervalos de confiança de 95%, sendo que valores de fiabilidade acima de 0,70 são considerados excelentes (Portney & Watkins, 2009). De seguida procedeu-se ao cálculo do Erro Padrão de Medida através da fórmula: $EPM = DP * \sqrt{(1-CCI)}$, em que DP corresponde ao Desvio Padrão (Portney & Watkins, 2009). Para além do cálculo da MMD, a utilização do EPM permite ainda que os resultados sejam expressos na medida original do teste (neste estudo centímetros e graus), facilitando assim a sua interpretação (Anastasi & Urbina, 1997). Foram igualmente calculados intervalos de confiança de 95% com base na fórmula, $95\%IC = 1,96 * EPM$ (Haley & Fragala-Pinkham, 2006). No final foi calculada a MMD através da fórmula: $MMD = z \text{ score level of confidence} \times DP_{\text{baseline}} \times \sqrt{(2[1-\text{retest}])}$, em que z representa o score para o intervalo de confiança de uma distribuição normal, (neste estudo, $IC = 95\%$: $z = 1,96$) (Haley & Fragala-Pinkham, 2006). Para o tratamento dos dados estatísticos, utilizou-se o programa *Statistical Package for the Social Sciences – Versão 19.0* (SPSS Statistics 19.0, Chicago, Estados Unidos da América). O nível de significância para todos os testes estatísticos foi fixado em $\alpha = 0.05$ (intervalo de confiança de 95%).

Resultados

Na tabela 2 apresenta-se o valor médio e desvio padrão dos dados recolhidos no 1º momento (avaliação inicial) e no 2º momento (avaliação final). Não foram encontradas diferenças significativas entre os dois momentos de avaliação, quer no que diz respeito ao WBLT quer ao DPA. Após a realização dos testes de normalidade, e por se ter verificado que a distribuição era normal no WBLT e a não normalidade para o DPA, realizou-se o teste paramétrico de T' Student para amostras emparelhadas e o teste não paramétrico de Wilcoxon, respetivamente.

Tabela 2. Valores Médios e desvio padrão obtidos na avaliação inicial e final.

Testes	Avaliação Inicial		Avaliação Final		Valor p
	Média	Desvio Padrão	Média	Desvio Padrão	
Weight-bearing lunge-test (cm)	8,29	2,35	8,17	2,35	0,197*
Deslizamento posterior do astrágalo (graus)	7,36	2,30	7,38	2,19	0,794†

* Valor de p do teste T' Student para amostras emparelhadas. † Valor de p no teste não paramétrico de Wilcoxon.

Nas tabelas 3 e 4 estão apresentados os resultados do CCI e do EPM para o WBLT e para o DPA.

Tabela 3. Resultados do Coeficiente de correlação intraclasse (CCI) para o WBLT e para o DPA.

Testes	$CCI_{(3,3)}$	IC 95%
Weight-bearing lunge-test	0,994	(0,982 - 0,998)
Deslizamento posterior do astrágalo	0,995	(0,987 - 0,998)

Tabela 4. Resultados do Erro padrão de Medição (EPM) para o WBLT e para o DPA.

Testes	EPM	IC 95%
Weight-bearing lunge-test (cm)	0,18	± 0,35
Deslizamento posterior do astrágalo (Graus)	0,16	± 0,31

Tal como podemos verificar pelas tabelas anteriores os valores do CCI são bastante elevados e os valores do EPM, bastante baixos. Estes valores são indicativos de uma boa fiabilidade intra-observador (Portney & Watkins, 2009) revelando uma baixa probabilidade de erro aleatório e sistemático (tabela 4) (Steffen & Seney, 2008). Os resultados da tabela 5 mostram a MMD do teste que não ocorre devido a erro. Este valor não sendo relevante para a verificação da fiabilidade do teste, é no entanto de extrema utilidade clínica. Assim, e de acordo com os dados obtidos, um valor de mudança observado num utente numa situação de pós-tratamento que seja inferior aos valores obtidos (referidos na tabela 5) não é distinguível do erro de medida, ou seja significa que não houve qualquer alteração no parâmetro avaliado. Da mesma forma, se o valor obtido for igual ou acima dos valores referidos na tabela isso significa que houve uma verdadeira alteração na amplitude de movimento avaliada pelo teste e não um erro de medida.

Tabela 5. Resultados da Diferença Mínima Detectável (DMD)

Testes	DMD
Weight-bearing lunge-test	0,51 cm
Deslizamento posterior do astrágalo	0,45°

Discussão

Os resultados obtidos nesta amostra de indivíduos com história de pelo menos um entorse do tornozelo indicam que o WBLT é um teste com elevada fiabilidade intra-observador, e com baixo erro sistemático. Os valores do CCI obtidos podem ser classificados como excelentes, de acordo com os critérios estabelecidos por Portney & Watkins (2009), o que é indicativo da consistência e reprodutibilidade das medições (Eliasziw, Young, Woodbury, & Fryday-Field, 1994). Para além disso, e de forma a podermos considerar as diferenças encontradas como uma alteração real e não devida a potenciais erros de medição, calculamos também o EPM e a MMD (Portney & Watkins, 2009), nos quais obtivemos igualmente bons resultados.

O WBLT tem revelado em diversos estudos valores de fiabilidade intra-observador elevados e valores de erro padrão de medição baixos, o que revela um erro sistemático mínimo. No estudo de Bennel et al. (1998) o CCI_(3,3) variou entre 0,97 e 0,98 enquanto que as variações no EPM foram entre 0,5 cm e 0,6 cm. Collins et al. (2004) e Vicenzino et al. (2001) referem um CCI_(3,3) de 0,99 e um EPM de 0,35 cm. Outro estudo realizado com o mesmo teste apresentou um CCI_(3,3) de 0,95 (I.C. 95%=[0,80 - 0,99]) e um EPM de 0,20 cm (Vicenzino et al., 2006). Os resultados encontrados neste estudo são consistentes com os valores da literatura (Collins et al., 2004; Vicenzino et al., 2001; Bennel et al., 1998). O CCI_(3,3) encontrado foi de 0,994, para um IC 95%=[0,982; 0,998], e o EPM foi de 0,18 cm. Relativamente ao EPM, os valores encontrados neste estudo são inferiores aos reportados por Bennel et al. (1998), Vicenzino et al. (2001) e por Collins et al. (2004). Estas diferenças podem estar relacionadas com o facto de se ter realizado previamente um período de treino que permitiu ao avaliador aperfeiçoar a técnica de medição, minimizando assim os erros inerentes ao avaliador, situação que apenas se encontra reportada no estudo desenvolvido por Collins et al. (2004). Quanto ao teste do DPA, os valores encontrados para o CCI_(3,3) foram de 0,995 (I.C. 95%=[0,987; 0,998]), sendo superiores relativamente aos encontrados por Denegar & Miller (2002a), que reporta valores de fiabilidade intra-observador CCI_(3,1) de 0,88 a 0,99. No estudo de Vicenzino et al. (2006) foi apresentado um CCI_(3,3) de 0,99 (I.C. 95%=[0,96; 0,99]). Importa no entanto referir que o modelo de CCI calculado no nosso estudo foi diferente do usado por Denegar & Miller (2002a). Ou seja, os estudos diferiram no número de medições que foram realizadas, em que no nosso estudo foi utilizada a média de três medições (CCI_(3,3)) enquanto que no estudo de Denegar & Miller (2002a) foi utilizada uma única medição (CCI_(3,1)). Por este motivo uma comparação directa não deverá ser feita. Ao compararmos com o resultado de Vicenzino et al. (2006) verificamos uma concordância no valor do CCI, que é considerado elevado (Portney & Watkins, 2009). Relativamente ao EPM calculado no nosso estudo (0,16°), este é menor do que o calculado por Denegar & Miller (2002a) e Vicenzino et al. (2006). Denegar & Miller (2002a) reportou valores para o EPM que varia de 0,34° a 2,00° (consoante se trata do membro lesado ou não lesado), enquanto Vicenzino et al. (2006) encontrou um EPM de 0,4°. Mais uma vez, esta diferença poderá dever-se ao facto de ter existido um período pré-experimental de adaptação aos equipamentos e treino das técnicas de medição (estudo piloto), o que não é referido pelos restantes autores.

Os resultados encontrados devem no entanto ser lidos atendendo às seguintes limitações. Em primeiro lugar, a amostra utilizada foi reduzida, considerando as recomendações de Altman (1991) que refere um mínimo de 50 utentes para se testar parâmetros de concordância. Neste estudo apenas foi utilizado um avaliador, com bastante experiência na área. Com outros avaliadores com diferentes características poderia ser possível obter outro tipo de resultados. No entanto a decisão de utilizar apenas um avaliador teve por base o estudo de Bennel et al. (1998), o qual não encontrou diferenças significativas nos resultados quando o teste foi efectuado por fisioterapeutas com 10 anos de experiência ou por estudantes do curso de Fisioterapia do 2º ano, e sem prática na aplicação do teste. Também na aplicação dos testes, e particularmente no 2º momento de avaliação, não foi realizada a aleatorização da sequência de teste dos participantes. Este procedimento teria permitido minimizar o possível efeito de memória, dada a proximidade das medições.

Por fim, apesar dos resultados deste estudo demonstrarem que o WBLT é um teste fidedigno, importa salientar que o WBLT mede a FD mas não identifica qual a origem das limitações (ex. limitações articulares ou de tecidos moles).

Conclusões

Os resultados deste estudo sugerem que ambos os testes podem ser ferramentas credíveis e de fácil aplicabilidade clínica na detecção de limitações na amplitude de movimento de FD e do DPA, respectivamente (Vicenzino et al., 2006; Collins et al., 2004; Denegar & Miller, 2002a; Vicenzino et al., 2001; Bennel et al., 1998). Do ponto de vista clínico os resultados deste estudo indicam que resultados iguais ou acima de 0,51 cm (DMD para o WBLT) ou de 0,45º (DMD para o teste de deslizamento posterior do astrágalo) são devidos a uma verdadeira melhoria da amplitude de FD ou de deslizamento posterior do astrágalo, respectivamente, e não a um erro de medida.

Bibliografia

- Altman, D. (1991). *Practical statistics for medical research*. London: Chapman and Hall.
- Anastasi, A., & Urbina, S. (1997). *Psychological Testing (7th.)*. Upper Saddle River, NJ: Prentice-Hall.
- Bennel, K., Talbot, R., Wajswelner, H., Techovavanich, W., & Kelly, D. (1998). Intra-rater and inter-rater reliability of a weight-bearing lunge measure of ankle dorsiflexion. *Journal of Physiotherapy*, 44(3), 175-180.
- Collins, N., Teys, P., & Vicenzino, B. (2004). The initial effects of a Mulligan's mobilization with movement technique on dorsiflexion and pain in subacute ankle sprains. *Manual therapy*, 9(2), 77-82. doi: 10.1016/S1356-689X(03)00101-2.
- Crosbie, J., Green, T., & Refshauge, K. (1999). Effects of reduced ankle dorsiflexion following lateral ligament sprain on temporal and spatial gait parameters. *Gait & posture*, 9(3), 167-72. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10575077>.
- Delahunt, E., O'Driscoll, J., & Moran, K. (2009). Effects of taping and exercise on ankle joint movement in subjects with chronic ankle instability: a preliminary investigation. *Arch Phys Med Rehabil*, 90.
- Denegar, C. R., & Miller, S. J. (2002). Can chronic ankle instability be prevented? Rethinking management of lateral ankle sprains. *Journal of Athletic Training*, 37(4), 430-435.
- Denegar, C. R., Hertel, J., & Fonseca, J. (2002). The effect of lateral ankle sprain on dorsiflexion range of motion, posterior talar glide, and joint laxity. *The Journal of orthopaedic and sports physical therapy*, 32(4), 166-73. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11949665>.
- de Noronha, M., Refshauge, K. M., Herbert, R. D., Kilbreath, S. L., & Hertel, J. (2006). Do voluntary strength, proprioception, range of motion, or postural sway predict occurrence of lateral ankle sprain? *British journal of sports medicine*, 40(10), 824-8; discussion 828. doi: 10.1136/bjsm.2006.029645.
- Eliasziw M, Young SL, Woodbury MG, & Fryday-Field K. (1994). Statistical methodology for the concurrent assessment of interrater and intrarater reliability: using goniometric measurements as an example. *Physical Therapy*, 74(8):777-88.
- Gatt, A; Chockalingam, N: Clinical assessment of ankle joint dorsiflexion A review of measurement techniques. *Journal of the American Podiatric Medical Association* 101:1, 2011
- Green, T., Refshauge, K., Crosbie, J., & Adams, R. (2001). A randomized controlled trial of a passive accessory joint mobilization on acute ankle inversion sprains. *Physical therapy*, 81(4), 984-94. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11276181>.
- Gribble, P., Taylor, B., & Shinohara, J. (2010). Bracing does not improve dynamic stability in chronic ankle instability subjects. *Physical Therapy in Sport*, 11, 3-7.
- Hadzic, V., Sattler, T., Topole, E., Jarnovic, Z., Burger, H., Dervisevic, E., et al. (2009). Risk factors for ankle sprain in volleyball players : A preliminary analysis. *Isokinetics and Exercise Science*, 17, 155-160. doi: 10.3233/IES-2009-0347.
- Haley SM, Fragala-Pinkham MA (2006). Interpreting change scores of tests and measures used in physical therapy. *Physical Therapy*.86, 735-743
- Hertel, J. (2000). Lateral Ankle Sprain. *Sports Medicine*, 29(5), 361-371.

- Hoch, M. C., & McKeon, P. O. (2011). Joint mobilization improves spatiotemporal postural control and range of motion in those with chronic ankle instability. *JOURNAL OF ORTHOPAEDIC RESEARCH*, 29(3), 326-32. doi: 10.1002/jor.21256.
- Hopper, D., Samsson, K., Hulenik, T., Ng, C., Hall, T., Robinson, K., et al. (2009). The influence of Mulligan ankle taping during balance performance in subjects with unilateral chronic ankle instability. *Physical Therapy in Sport*, 10, 125-130.
- Hubbard, T. J., & Hertel, J. (2006). Mechanical contributions to chronic lateral ankle instability. *Sports medicine (Auckland, N.Z.)*, 36(3), 263-77. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16526836>.
- Monaghan, K., Delahunt, E., & Caulfield, B. (2006). Ankle function during gait in patients with chronic ankle instability compared to controls. *Clinical biomechanics (Bristol, Avon)*, 21(2), 168-74. doi: 10.1016/j.clinbiomech.2005.09.004.
- Portney, L., & Watkins, M. (2009). *Foundations of Clinical Research - Applications to Practice*.
- Rome, K: Ankle joint dorsiflexion measurement studies A review of the literature. *Journal of the American Podiatric Medical Association* 86:5, 1996
- Reid, A., Birmingham, T. B., & Alcock, G. (2007). Efficacy of Mobilization with Movement for Patients with Limited Dorsiflexion after Ankle Sprain : A Crossover Trial. *Physiotherapy*, 166-172. doi: 10.2310/6640.2007.00016.
- Russek, L. (2004). Factors affecting interpretation of reliability coefficients. *Journal of Orthopaedic & Sports Physical Therapy*, 34(6), 341-349.
- Steffen, T. Seney, M. (2008). Test-Retest Reliability and Minimal Detectable Change on Balance and Ambulation Tests, the 36-Item Short-Form Health Survey, and the Unified Parkinson Disease Rating Scale in People With Parkinsonism. *Physical Therapy*, 88 (6), 733- 746.
- Tabrizi, P., McIntyre, W. M., Quesnel, M. B., & Howard, a. W. (2000). Limited dorsiflexion predisposes to injuries of the ankle in children. *The Journal of bone and joint surgery. British volume*, 82(8), 1103-6. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11132266>.
- Vicenzino B, Prangley I, Martin D. The initial effect of two Mulligan mobilization with movement treatment techniques on ankle dorsiflexion. *Proceedings of the Australian Conference of Science and Medicine in Sport, Perth, WA, June 2001. Brisbane,Q: Sports Medicine Australia; 2001.*
- Vicenzino, B., Branjerdporn, M., Teys, P., & Jordan, K. (2006). Initial changes in posterior talar glide and dorsiflexion of the ankle after mobilization with movement in individuals with recurrent ankle sprain. *J Orthop Sports Phys Ther*, 36(7), 464-471.